
The perfect match: the role of categorical fit between venture capitalists and their startup investments

*Vom Promotionsausschuss der Technischen Universität
Hamburg-Harburg zur Erlangung des akademischen Grades*

Doktor der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

genehmigte Dissertation

von

Jan-Frederik ARNOLD

aus

Ankum

2018

Gutachter:

Prof. Dr. Christoph IHL

Prof. Dr. Sebastian SPÄTH

Tag der mündlichen Prüfung:

16. März 2018

Acknowledgements

I would like to thank the following people, as, without their support, I would not have been able to successfully finish this doctoral thesis.

Firstly, I would like to thank my "*Doktorvater*" Prof. Dr. Christoph Ihl, who provided guidance in all stages of the academic work. This includes the selection of the interesting and challenging topic, the embedding of the relevant theoretical literature, the data collection and analysis, and, finally, creating this dissertation. I am particularly grateful for Christoph's openness, approachability, and the great atmosphere that he created at the Institute of Entrepreneurship.

Furthermore, I have highly appreciated the further academic guidance from Prof. Dr. Sebastian Späth, who also accepted the role of second assessor of this thesis.

The academic activities at the Institute of Entrepreneurship, including both the formal discussions during the weekly doctoral seminars and the regular informal exchanges, were essential for the progression of my research. I thus want to thank my fellow doctoral students Giulio Barth, Viktoria Boss, Michael Engel, Dimitri Graf, Matthias Jacobi, Dr. Hannes Lampe, Benjamin Müller, Dr. Jan Reerink, and Jan-Niklas Wick for these inspiring and motivating discussions. They also made my visits to Hamburg very enjoyable and created lasting memories during the many social events we shared.

Birgit Grabi supported me in all the organizational matters of my dissertation – she is truly "*the heart and soul*" of the institute.

A special thanks goes to my family. My parents have always supported me unconditionally, on my path to starting the doctoral thesis and during the academic work towards it. In addition, I would like to thank my sister Henrike, who hosted me during my stays in Hamburg – I truly enjoyed the extra time spent together. I am also grateful for the continuous encouragement and exchange sessions with my brother Thorsten and his wife Ulrike.

Finally, my biggest thanks goes to my wife Susanne, whom I not only got engaged and married to during the time of my thesis, but who also continuously motivated and believed in me at all stages. Her support, especially when I faced challenges in my academic work, and sharing the good moments as I approached the finish line were crucial for the successful completion of my dissertation.

TECHNICAL UNIVERSITY HAMBURG-HARBURG

Abstract

The perfect match: the role of categorical fit between venture capitalists and their startup investments

by Jan-Frederik ARNOLD

Venture capitalists and startups consider a number of factors in their search for the perfect match before a venture capital funding round is closed. This dissertation sheds light on how the categorical fit between venture capitalists and startups affects this matching process.

It is generally accepted that organizations need legitimacy to be considered relevant by various audiences. In recent years, researchers have become increasingly interested in how venture capitalists differ from other audiences in the way they balance the trade-off between distinctive, boundary-crossing, and highly legitimized, conforming new ventures. A new venture's entrepreneurial identity is created by its cultural and structural embeddedness, which venture capitalists use in their assessment of potential investment candidates. Few attempts have been made to explain the role of market-category-based distinctiveness and categorical distance in the investment process.

Firstly, we theorize that venture capitalists with ample prior experience from investments in distinctive new ventures and with high-status prefer more distinctive new ventures, while venture capitalists with highly diversified portfolios prefer less distinctive new ventures. Secondly, we argue that categorical distance between new ventures and venture capitalists reduces the likelihood of an investment, but this effect is mitigated by a venture capitalist's experience, degree of portfolio diversification, and status. The hypotheses are tested in an empirical analysis of 11 years of venture capital investment data from the USA. The study uses measures of distinctiveness and categorical distance based on market categories. The analysis supports the majority of the hypotheses. Contrary to our expectations, the analysis shows support that portfolio diversification has a positive effect on preference for distinctiveness and does not support that experience with distinctive new ventures mitigates the relevance of categorical distance.

The contributions of this dissertation are threefold: Firstly, we help to explain which venture capitalists invest in which startups. Secondly, we add support to prior findings that venture capitalists are different from other audiences in their trade-off between legitimacy and distinctiveness. Thirdly, we make a methodological advancement by defining a measure of new ventures' distinctiveness and the categorical distance based on market categories. The dissertation is most relevant to researchers in entrepreneurship and management, sociologist, and practitioners such as venture capitalists or founders.

Contents

Acknowledgements	iii
Abstract	v
1 Introduction	1
1.1 Motivation	1
1.2 Research question and background	2
1.3 Methods	5
1.4 Results and contributions	6
1.5 Structure	7
2 Venture capitalists and startups: empirical research review	9
2.1 The venture capital industry	9
2.1.1 The startup - venture capitalist relationship	10
2.1.2 The effect of venture capital	12
2.1.3 Venture capitalists' performance	15
2.1.4 Syndication	16
2.1.5 Summary	17
2.2 The literature review framework	18
2.2.1 Unit of analysis	18
2.2.2 Network and related measures as explanatory variables	19
2.3 Literature-gathering approach	21
2.4 Review based on variables	23
2.4.1 A. Startup focus	23
2.4.2 B. Venture capitalist and syndicate focus	30
2.4.3 C. Venture capitalist - startup dyad focus	39
2.5 Summary and outlook	44
3 Theoretical framework and hypotheses	47
3.1 The importance of legitimacy and distinctiveness	48
3.1.1 Organizational legitimacy	48
3.1.2 Distinctiveness through boundary-spanning	49
3.2 The creation of entrepreneurial identity	51
3.2.1 Introduction to entrepreneurial identity	51
3.2.2 Cultural and categorical embeddedness	52
3.2.3 Structural embeddedness	53

3.3	Venture capitalists and entrepreneurial identity	55
3.3.1	Distinctiveness and novelty in venture capitalists' portfolios	55
3.3.2	Diversification of venture capitalists' portfolios	57
3.3.3	The combined effect of average distinctiveness and diversification	59
3.3.4	Status benefits from structural embeddedness	60
3.4	Categorical distance and venture capitalists' attributes	62
3.4.1	General effect of distance	63
3.4.2	The effect of portfolio distinctiveness on distant investments	67
3.4.3	Portfolio diversification through distant new ventures	68
3.4.4	Effect of status on distant investments	68
3.5	Summary and conceptual model	69
4	Methodology: data selection and variable definition	71
4.1	Introduction to methodology: data selection and variable definition	71
4.2	Data collection	71
4.3	Dependent variable	75
4.4	Independent and moderating variables	75
4.4.1	Startup's distinctiveness	75
4.4.2	Average portfolio distinctiveness	81
4.4.3	Portfolio diversification	81
4.4.4	The venture capitalist's status	84
4.4.5	Categorical distance	85
4.5	Control variables	86
4.5.1	Startup and funding-round level controls	86
4.5.2	Venture capitalist and syndicate level controls	88
4.5.3	Venture capitalist - startup dyad level controls	89
5	Methodology: empirical approach and analytical method	93
5.1	Empirical approach: the addition of unrealized ties	93
5.2	Analytical method: logistic regression and rare-events application	96
5.2.1	Logistic regression	97
5.2.2	Peculiarities in rare-events settings	100
5.2.3	Imputation for missing data	102
5.3	Specification of statistical models	103
6	Results	105
6.1	Descriptive statistics	105
6.1.1	Overview of independent variables	105

6.1.2	The dataset	111
6.2	Model results	112
6.2.1	Analysis part I: startup's distinctiveness	113
6.2.2	Analysis part II: categorical distance	118
6.3	Robustness test	119
7	Discussion	129
7.1	The effects on the preference for new venture's distinctive- ness	129
7.1.1	Portfolio distinctiveness	130
7.1.2	Portfolio diversification	130
7.1.3	A venture capitalist's status	132
7.1.4	Overarching discussion of new venture's distinctive- ness	134
7.2	The effect of categorical distance	138
7.2.1	Overarching discussion of categorical distance	138
7.2.2	Portfolio distinctiveness	139
7.2.3	Portfolio diversification	140
7.2.4	A venture capitalist's status	140
7.3	The effect of corporate venture capital	141
8	Conclusion	145
8.1	Synthesis of empirical findings	146
8.2	Theoretical contributions and implications	146
8.3	Practical implications	148
8.4	Limitations and future research	149
8.5	Closure	151
A	Detailed regression models	153
B	Software used in research project	157
C	Supplementary analysis corporate venture capital	159
	Reference list	161

List of Figures

2.1	Literature review framework	18
2.2	Literature review overview	25
3.1	Theoretical building blocks	47
3.2	Summary portfolio attributes and risk tolerance	59
3.3	Conceptual model	69
4.1	Filters to derive base dataset from the original Crunchbase file	73
4.2	Frequency distribution of categories per new venture in the sample	76
4.3	Comparing category sets of two startups in a geometric space.	82
5.1	Realized and unrealized ties in an investment network.	94
5.2	Plot of logistic function	98
6.1	Histograms and density plots with respective normal curves	107
6.2	Comparison of portfolio diversification and categorical dis- tance based on density plots and normal curves	109
6.3	Three dimensional scatter plot of the three attributes of the venture capitalist	110
6.4	Analysis by status quartiles	110
6.5	Frequency distribution by year	111
6.6	Frequency distribution of funding rounds per new venture .	112
6.7	Geographic distribution of startups	113
6.8	Graphical analysis of interactions	123

List of Tables

2.1	Overview of primary journals used for structured literature review	22
4.1	Overview control variables by unit of analysis	87
5.1	Overview of the creation of variables in unrealized ties	96
6.1	Descriptive statistics	106
6.2	Pairwise correlation matrix for main variables	108
6.3	Overview core information final dataset	111
6.4	Logistics regression output models part I: startup's distinctiveness*	117
6.5	Logistics regression output models part II: categorical distance*	121
6.6	Regression output models with two-way clustering of standard errors by startup and venture capitalist part I: startup's distinctiveness*	125
6.7	Regression output models with two-way clustering of standard errors by startup and venture capitalist part 2: categorical distance*	127
A.1	Regression output models with rare events correction*	156
C.1	Regression output models supplementary analysis corporate venture capital*	160

Chapter 1

Introduction

"It's almost always harder to raise capital than you thought it would be, and it always takes longer. So plan for that."

Richard Harroch¹

1.1 Motivation

The amount of growth capital to be invested in the United State of America (US) reached its second-highest value in 2015, with 58.9 billion US Dollar (USD) being invested in 4,380 deals. Both figures were topped only in the pre-crisis and dotcom-bubble year of 2000, in which 105 billion USD were invested in 8,041 deals, according to a recent report published by PricewaterhouseCoopers LLP and the National Venture Capital Association.² Growth capital, of which 71% was for early- and expansion-stage investments in 2015, typically made by venture capitalists, remains a major liquidity source for new ventures in the US. Venture capital has funded some of today's largest technology firms such as Apple, Google, and Amazon (Kaplan and Lerner, 2010), and survival rates are higher when startups are funded by venture capital (Sandberg, 1986). The funds committed to the venture capital industry have seen various cycles in the past, as has the performance of venture capitalists (Gompers and Lerner, 2001). The recent increase of growth capital shows the increasing interest of corporate and private venture capitalists in investing in new ventures to gain access to startups' technologies or to purely use venture capital as an asset class.

However, before a funding round takes place, a complicated and lengthy matching process between venture capitalist and startup takes place. Venture capitalists review many potential investment opportunities and usually have a diligent screening process before making an investment

¹In 2003, Richard Harroch joined VantagePoint Capital Partners (VPCP), a large venture capital fund established in 1996 and based in California. He today holds the position of managing director and global head of M&A. He is also the author of a number of books and founder of two companies that he successfully sold.

²Report retrieved on April 13, 2016 from <http://nvca.org/research/venture-investment/>. Data in the report was provided by Thomson Reuters.

(Kaplan and Strömberg, 2001). Venture capitalists look for new ventures that, among other factors, fit with their investment focus, have a promising business model and team, and provide a generally good investment opportunity. On the other hand, startups that decide to bring a venture capitalist on board search for an investor that fulfills not only their financial but also non-financial needs. Typically, venture capitalists work closely with their portfolio companies (Gorman and Sahlman, 1989), providing advice, access to their network partners, and other valuable resources. Consequently, new ventures also aim to find a suitable venture capitalist. Depending on a startup's life-cycle needs with regard to its development stage, industry, and investment time-horizon, a good fit adds value to the focal startup (Lungeanu and Zajac, 2016). The matching process before a funding round is thus simultaneously a crucial and tedious step before the start of the venture capitalist - startup relationship.

1.2 Research question and background

Not only is venture capital a driving factor for the economic development of young organizations, but the venture capital domain has also received a lot of attention in academia through a large number of scholarly articles. This stems from the fact that the venture capital domain is characterized by a high level of uncertainty, asymmetric information, and often few tangible assets in the target firms (Kaplan and Lerner, 2015). As most venture capital investments are made in a syndicate, this forms a useful domain in which to study network effects, further increasing the number of research articles in the venture capital context. One stream of research has focused on the decision criteria of venture capitalists to explain how investment decisions are made (e.g., Hall and Hofer, 1993; Zacharakis and Meyer, 1998; Shepherd, 1999; Dimov, Shepherd, and Sutcliffe, 2007; Monika and Sharma, 2015). However, despite these efforts, there still remains a gap in explaining which venture capitalists invest in which startups. Only recently have researchers started to explore the role of category labels in the matching process between venture capitalists and startups (Pontikes, 2012; Wry, Lounsbury, and Jennings, 2014), but a detailed cross-industry analysis thoroughly differentiating between various venture capitalists has not been performed. An important element is the role of categorical fit as a measure of cultural embeddedness in the matching process. We therefore aim to answer the following overarching research question:

How does categorical fit affect the search for the perfect match between venture capitalists and startups?

Derived from prior research, we focus on two sub-questions that serve as guides through this dissertation:

1. What affects a venture capitalist's choice regarding a startup's novelty as measured by its distinctiveness?
2. How does categorical distance affect the matching of startups and venture capitalists and what mitigates this distance's effect?

We aim to answer the questions above by treating venture capitalists as a heterogeneous group, differentiating on the basis of two portfolio attributes and the venture capitalist's status.

Firstly, organizations gain legitimacy from being embedded in their environment, which consists of generally accepted beliefs, norms, and categories (Scott and Meyer, 1983). It has been shown that organizations that are not considered legitimate are ignored. For instance, due to exclusion from analyst reports, less embedded securities trade at a discount (Zuckerman, 1999). Similarly, consumers prefer legitimized new ventures, i.e., organizations that are assigned clear category labels, as they use these categories when navigating their consumption behavior (Pontikes, 2012). Generally, legitimacy creates trust and meaningfulness, which can improve organizations' resource acquisition and survival rates (Meyer and Rowan, 1977; Suchman, 1995; Rao, 1994). The need for legitimization or to conform to the existing cultural environment can lead to the assimilation of organizations, so-called "isomorphism" (DiMaggio and Powell, 1983, p. 149). The need for legitimacy is particularly relevant for new ventures as they have little other means of proving themselves and being accepted (Aldrich and Fiol, 1994; Navis and Glynn, 2011). However, new ventures also need to demonstrate their novelty (Aldrich and Fiol, 1994) and often strive towards breakthrough innovations. Rosenkopf and Nerkar (2001) have shown that technological or organizational boundary-spanning leads to radical innovations, while highly conforming behavior only leads to incremental innovations. It indicates that, for new ventures, there exists a trade-off between being legitimate and being distinctive (the antagonist to legitimacy), and new ventures display heterogeneity regarding their level of distinctiveness (Navis and Glynn, 2011).

The way the trade-off between legitimacy and distinctiveness is perceived depends on the audience (Pontikes, 2012). While many audiences, particularly consumers, prefer highly legitimized organizations, scholars have recently placed attention to venture capitalists who instead prefer boundary-spanning, distinctive new ventures (Pontikes, 2012; Wry, Lounsbury, and Jennings, 2014). Distinctive new ventures demonstrate greater novelty and potentially stronger innovative power. This is paralleled with an increase in risk due to a startup's lower legitimacy, which should be

compensated by higher returns, which, in turn, make distinctive startups attractive to some venture capitalists. Pontikes (2012) and Wry, Lounsbury, and Jennings (2014) have treated venture capitalists as a homogeneous group, but the venture capitalist landscape is manifold: For instance, significant differences exist in terms of portfolio strategies, such as diversification versus specialization (Norton and Tenenbaum, 1993; Gupta and Sapienza, 1992; Matusik and Fitza, 2012), or with regard to their status gained from syndication networks (Podolny, 2001; Bothner, Kim, and Lee, 2015). We thus infer that venture capitalists are also heterogeneous with regard to their preference for distinctive versus legitimized new ventures.

A new venture's legitimacy and distinctiveness stems from the structural and cultural embeddedness of the new venture (Goldberg et al., 2016), which creates its entrepreneurial identity. Structural embeddedness describes ties to competitors, venture capitalists, and other network partners a startup may have. These ties can be a quality signal (Higgins and Gulati, 2003; Stuart, Hoang, and Hybels, 1999) or help to overcome lack of legitimacy (Aldrich and Fiol, 1994). Cultural embeddedness concerns the relation to existing beliefs, norms, and categories and is often viewed through an organization's categorization lens. For instance, consumers use categories as a navigation tool (Loken, Barsalou, and Joiner, 2008; Pontikes, 2012). Categories can come in many forms, like product categories (Meyers-Levy and Tybout, 1989), brand categories (Barone, 2005), cultural categories (Briley and Wyer, 2002), or market and industry categories (Sorenson and Stuart, 2001; Pontikes, 2012). New ventures either gain legitimacy by conforming to existing categories or demonstrate distinctiveness by crossing them, such that their entrepreneurial identity ideally shows a "legitimate distinctiveness" (Navis and Glynn, 2011, p. 482).

In order to shed light on the matching process of venture capitalist and startup, we take into consideration that startups and venture capitalists resemble heterogeneous groups. We differentiate between the venture capitalist's average portfolio distinctiveness, portfolio diversification, and status derived from its co-syndication network. For the first research question, we explore how the preference for distinctive and novel new ventures depends on these attributes. This differentiation is important, as researchers have previously treated venture capitalists as a homogeneous group, and said differentiation allows us to better understand why there is good reason for more and less distinctive startups to exist.

For the second research question, we are no longer interested in the startup's novelty in and of itself, but in relation to the portfolio companies of the venture capitalist. Similarly to industry and geographical distance, which reduce a venture capitalist's investment probability (Sorenson and Stuart, 2001), we expect that a larger market category distance also reduces

the likelihood of an investment. Even more importantly, we examine how the three attributes of the venture capitalist introduced previously, that is, portfolio distinctiveness, portfolio diversification, and status, mitigate the effect of categorical distance.

We use market categories for two main independent variables: Firstly, we measure a new venture's novelty or distinctiveness, i.e., its cultural embeddedness, and, secondly, the novelty of a new venture compared to the a venture capitalist's existing portfolio companies measured by categorical distance. We define a startup or new venture (using these terms interchangeably throughout this paper) as a company that has recently been founded by an entrepreneur. The focus of this paper is on startups that have surpassed the seed stage, in which they define their product and business model, and are in the early or expansion stage, in which venture capitalists are typically engaged, especially in cases of high capital requirements (Podolny, 2001). We focus on venture capitalists and the funding that they provide, as venture capitalists are the most important actors for the identification and financing of young businesses (Monika and Sharma, 2015).

1.3 Methods

We have built our empirical analysis on a dataset from Crunchbase³, a crowd-based startup database, including funding and investor information, that belongs to the technology blog TechCrunch⁴. After applying a number of filters, our dataset includes 5,826 US startups that have received funding in 10,576 funding rounds from 2,323 investors over the 11-year period from 2005 to 2015. Our unit of analysis is the venture capitalist - startup dyad per funding round, which totals 29,000 realized investments. We have constructed the final dataset by adding one non-realized venture capitalist - startup dyad for each realized investment, such that the final dataset consists of 58,000 observations. We have used market categories populated by contributors to the database to measure startup's distinctiveness, categorical distance, a venture capitalist's portfolio distinctiveness, and portfolio diversification. The methodology for the calculation of these network category measures is based on two measures of boundary-spanning developed by Hannan, Goldberg, and Kovács (2016). Status is measured with the standard eigenvector centrality score, so that the status of an investor depends on the status of his or her syndicate members and so forth (Bonacich, 1987). We have used logistic regression with a correction for rare-events application (King and Zeng, 2001) in order to validate our hypotheses and test the

³www.crunchbase.com

⁴www.techcrunch.com

robustness of our results with a multi-way clustering of standard errors by startup and venture capitalist.

1.4 Results and contributions

The first part of the analysis is concerned with the preference for new ventures' distinctiveness. Our results show that venture capitalists with experience in managing the risk of investing in distinctive startups are more likely to invest in distinctive startups in the future. Contrary to our expectations, we have found evidence that venture capitalists with highly diversified portfolios are more likely to invest in distinctive new ventures. Moreover, high-status investors can afford to invest in more distinctive, riskier startups. The second part of our analysis confirms that new ventures with large categorical distance to the existing portfolio companies are less likely to get funded by the focal venture capitalist. Highly diversified portfolios or high status mitigate this effect. We were not able to confirm our hypothesis that investors who have previously invested in distinctive startups are more likely to invest in startups with larger categorical distance.

Our empirical study makes a number of contributions: Firstly, we show that it might be beneficial for some organizations to be less legitimized and more distinctive, as they can attract certain audiences, such as high-status venture capitalists. This generally throws into question the widespread belief that organizations get discounted when they are boundary-spanning. Secondly, our analysis confirms prior studies that have shown that venture capitalists judge startups differently than other audiences, like consumers. Due to our cross-industry analysis, our results are more generalizable than the findings of prior research. As we do not treat venture capitalists as a homogeneous group, but differentiate them in terms of their portfolio strategy and status, we also have a more fine-grained understanding of which venture capitalists prefer distinctive startups and care less about categorical distance. Thirdly, the introduction of a measure of categorical distance, which assesses the degree of novelty of an organization compared to a group of other organizations, is highly relevant for future studies. Researchers can apply this distance measure to other settings to compare multiple organizations with each other. Fourth, we answer the research call made by Hannan, Goldberg, and Kovács (2016) to show that their measures of boundary-spanning can be transferred to the venture capital context. However, we come, in parts, to different conclusions regarding the preferences for boundary-spanning organizations. This is particularly relevant, as their paper has made significant contributions to sociology by developing a framework to measure boundary-spanning both on the object level and across objects. Fifth, our analysis is based on very recent data,

which has barely been analyzed in other studies, while many of the influential papers relevant to our analysis and to research on the venture capital industry in general are based on data that is at least 10 to 15 years old (e.g., Podolny, 2001; Sorenson and Stuart, 2001; Pontikes, 2012). We also show that some of the prior findings still hold true today, despite the fast changing startup and venture capital environment.

1.5 Structure

The remainder of this dissertation is structured as follows: We begin by reviewing the relevant empirical literature on and related to the venture capitalist - startup relationship. We initiate this review with an overview of the venture capital domain, including the mechanics at work, followed by a description of our approach to a structured literature review on the empirical network and related variables used in this domain. We then highlight the most relevant papers and classify them into our literature review framework. In chapter 3, we develop our theoretical framework based on the theoretical background of entrepreneurial identity and describe how entrepreneurial identity is created from cultural and structural embeddedness. We also develop our hypothesis from the theoretical framework and split the hypothesis development in line with the two main research questions by a startup's distinctiveness and categorical distance. Chapter 4 contains the filters that we applied to get our base dataset and the detailed description of the dependent, independent, and control variables. In the subsequent chapter, we outline our methodological approach of adding the non-realized ties to the base dataset and provide some background on the statistical models used to test our hypotheses. After presenting the descriptive statistics, the results of the logistic regression models, and the robustness test in chapter 6, we discuss our findings in relation to prior research in chapter 7. The concluding chapter focuses on the theoretical and practical implications of our studies, outlines limitations, and shows avenues for further research.

Chapter 2

Venture capitalists and startups: empirical research review

This chapter aims to put the remainder of this dissertation into a theoretical context. We start with a broad overview of the main elements and characteristics of the venture capital - startup domain. The section touches upon the most relevant aspects, highlighting major research and carving out unresolved discrepancies. The following section develops and introduces the framework for the structured literature review. After describing the approach to gathering literature, we summarize the main empirical studies in the venture capital - startup domain focusing on the analysis methods. Categorizing the relevant studies based on their units of analysis and the empirical network-based and related variables, we summarize how these variables were used and what the empirical findings were.

The two-step approach allows us to understand the venture capital - startup domain in general and to consider the empirical methods used separately. We conclude by placing our analysis in the research context highlighting the blind spots in prior research as presented in the literature review.

2.1 The venture capital industry

Venture capitalists provide financial capital to young firms that need capital to finance their growth. Venture capitalists collect money from their limited partners, who can be private, corporate, or institutional investors that aim to receive a return on their investment. In a limited partnership, which is the organizational form of 80% of venture capital funds (Gompers, 1995), the limited partners only have the responsibility to provide a pre-defined amount of capital to the general partners, the venture capitalists, that invest the capital into startups. In addition to the financial capital, venture capitalists can provide other types of resources like experience sharing,

network contacts, or direct involvement in important decisions through board membership and participation to the new venture.

2.1.1 The startup - venture capitalist relationship

The basics

Venture capitalists are intermediaries between financial investors and new ventures and they often, at least in the United States, use convertible securities to gain a significant equity share (Casamatta, 2003). Typically, new ventures require different amounts of capital throughout their lifecycle, and investors spread their investments into multiple funding rounds to reduce uncertainty (Sahlman, 1990). It is common to separate the following three phases with decreasing uncertainty regarding the quality and success of the new venture (Podolny, 2001): the seed stage, in which mainly an idea and a business plan exist; the venture capital stage, in which a working product is sold to customers even though, typically, no profit is made; and the expansion stage, in which the firms continue to expand and make profits. In the early stages, risk is mostly internal and firm-specific, while, in later stages, risk mainly derives from external factors, such as the market (Ruhnka and Young, 1991). The higher risk in venture capital compared to that in investments in more mature companies is compensated for by higher returns. The risk in the later funding rounds is lower than in the earliest round, and overall risk-return characteristics are similar to the smallest Nasdaq stocks (Cochrane, 2005).

Ferrary (2010) shows that each investor plays a certain role in a given syndicate. The author differentiates between pure venture capitalists, corporate venture capitalists, private equity investors, and investment banks. He shows in his empirical analysis that pure venture capitalists overtake the typical identification, screening, and monitoring roles, especially in the early stages of a startup, whereas the other types of investors focus on helping the new venture to develop their business and join syndicates in later stages.

Venture capitalists commonly have an investment horizon of five to seven years, after which they often aim to sell their stakes in the organization through an initial public offering (IPO) or in a private sale (Gorman and Sahlman, 1989; Bygrave and Timmons, 1992). Contrary to other types of financial investors, venture capitalists work closely with their investment portfolio companies. On average, investors spend over 80 hours per year with each portfolio company (Gorman and Sahlman, 1989). Through this involvement, venture capital ownership has a number of effects on the new venture's strategic development (e.g., Sapienza, Manigart, and Vermeir, 1996; Lerner, 1995; Sapienza, 1992; Gorman and Sahlman, 1989), mainly

in three areas, as detailed in section 2.1.2: innovation, internationalization, and overall performance.

The relationship

The venture capitalist - startup relationship is characterized by informational asymmetries leading to an adverse selection and moral hazard problem (Amit, Brander, and Zott, 1998). The first is created by the overstatement of the performance and quality of a new venture by its founders or management, making it hard for the investor to evaluate the new venture. The moral hazard problem describes the phenomenon of managers having a lower performance incentive when a large share of a new venture is owned by venture capitalists; however, new ventures need large amounts of capital to grow. The relationship is thus a typical principal-agent situation (Sahlman, 1990): The new venture is an agent that seeks financial and social capital, and the venture capitalists are the principals providing the desired capital. Due to the lack of control and high information asymmetry in the principal-agent relationship, diligent screening typically takes place, followed by strict contracting and continuous monitoring (for details, see Kaplan and Strömberg, 2001). The right contracts in the venture capitalist - startup relationship can (partially) mitigate agency problems (Admati and Pfleiderer, 1994; Bergemann and Hege, 1998). Splitting the investments into several stages is another prominent way to overcome the problems of informational asymmetries via closer monitoring (Sahlman, 1990).

In sum, venture capitalists typically invest in industries in which informational asymmetries exist, as they are better at overcoming these asymmetries than other forms of capital providers, such as banks; thus they provide capital to highly uncertain new ventures (Amit, Brander, and Zott, 1998; Gompers and Lerner, 2001). While venture capital is a relatively expensive form of financing, there is a threshold of entrepreneurial risk, up to which (debt) bank financing is optimal and after which venture capital is more attractive (Bettignies and Duchêne, 2015).

Vergara, Bonilla, and Sepulveda (2016) demonstrate in a theoretical model that, while the effort of an entrepreneur decreases with an increased equity share given to the venture capitalist, the optimum is a 50/50 split of the cash flows between venture capitalists and the entrepreneur. This is caused by the complementarity of the entrepreneur with his or her ideas and technological skills and the venture capitalist providing a broad contact network, experience, and other services.

The process of how venture capitalists and new ventures come together is a two-sided matching process. The venture capitalist aims to invest in new ventures that provide the greatest return, an improved reputation, the

best syndicate partners, the best strategic fit with the other portfolio companies or the portfolio strategy, and so on (e.g., Lungeanu and Zajac, 2016; Hallen, 2008; Sorenson and Stuart, 2001). Due to venture capitalists' ability to exert control over the company, they differ in their evaluations of new ventures from other audiences, such as consumers or employees (Pontikes, 2012). On the other hand, new ventures choose their investors according to their needs in terms of information access, network access, or reputational consequences.

For instance, a similar decision-making process between an entrepreneur and a venture capitalist improves the venture capitalist's evaluation of the focal new venture (Murnieks et al., 2011). Ethnicity also plays a role: Bengtsson and Hsu (2015) show that, if a venture capitalist and a new venture's founder have the same ethnicity, an investment becomes more likely and the investment size and involvement increases.

2.1.2 The effect of venture capital

Venture capital funding has been found to affect startups in a number of ways. Researchers have mainly focused on the effects on innovation and patenting, internationalization, and overall performance and growth. In this section, we outline the main findings including a few results outside of the three main areas.

Innovation and patenting

Overall, in the right configurations, that is, the right venture capitalists and new ventures, venture capital has a positive effect on a new ventures's innovativeness (e.g., Kortum and Lerner, 2000; Ueda and Hirukawa, 2008) despite some studies also finding negative effects (e.g., Lahr and Mina, 2016).

In general, venture capitalist increase the risk taking and innovative power of early-stage new ventures (Park and Tzabbar, 2016) especially in the two years after the venture capital investment (Arqué-Castells, 2012). A positive effect of venture capital on innovation in European countries has especially been found when these countries have a high-venture capital intensity and a venture-capital-friendly legal system (Popov and Roosenboom, 2012). Critical is the involvement of venture capitalists in innovative new ventures (Sapienza, 1992). Alvarez-Garrido and Dushnitsky (2016) found that new ventures financed by corporate venture capital are more innovative and benefit more from complimentary assets than new ventures financed by non-corporate venture capitalists. Their analysis was conducted on a sample of 545 new ventures in the US biotechnology sector. It was shown that venture capital not only drives innovation, but it also increases

the number of patent litigations up to the point that too much litigation reduces the venture capital invested (Kiebzak, Rafert, and Tucker, 2016).

Contrary to the findings of increased innovating and patenting activities, some researchers attribute greater innovativeness and patenting activity to a selection effect, that is, venture capitalists investing in more innovative new ventures. Lahr and Mina (2016) find that, when controlling for endogeneity, there is no positive effect of venture capital on the patenting activities of new ventures. The authors argue that venture capitalists are merely better at identifying the most innovative startups, but do not contribute themselves to increased innovativeness. Patenting, especially before the first round of financing, is thus instead used as a signaling activity to attract more venture capital (Hoenen et al., 2014; Audretsch, Bönte, and Mahagaonkar, 2012). These findings are confirmed by the work of Engel and Keilbach (2007) who find a significant selection effect in a German sample.

There is an agreement that patenting activity is used as a quality signal. Hoenen et al. (2014) find in their analysis of biotechnology companies that new ventures use patenting as a quality signal in the first round of financing but the effect diminishes in the subsequent rounds, as venture capitalists then have other means of monitoring and controlling the quality of the new venture. The authors measure the amount of venture capital received in the first two rounds of financing and analyze the effect of the number of patents and the economic value of the patents per new venture. They measure the latter by the number of citations of the focal patent.

Internationalization

Venture capital also has a positive effect on the internationalization behavior of new ventures especially when the venture capitalist has strong international knowledge and a good reputation (Fernhaber and McDougall-Covin, 2009).¹ Other authors have found that venture capitalists only increase the internationalization of young firms when they have an internationally experienced board member of the focal venture in place (Carpenter, Pollock, and Leary, 2003).

¹Other types of alliance networks also affect the internationalization of new ventures. In their analysis of 448 US startups, Fernhaber and Li (2013) find that new ventures increase their international sales faster when they have an internationally experienced partner network and internationally experienced, geographically close firms. The partner network has an even greater effect for more mature new ventures. They consider a company an alliance partner when the alliance or joint venture is officially recorded in a public database. The study shows that partner networks are important to new ventures, even though the study did not analyze the network positions by related measures and did not include venture capitalists in its analysis.

Overall performance and growth

The effect of venture capital on a new venture's performance has been widely researched, with mixed results ranging from positive performance impact (e.g., Brav and Gompers, 1997) to negative impact (e.g., Jain, Jayaraman, and Kini, 2008). For instance, Bertoni, Colombo, and Grilli (2011) aim to separate the selection effect from the actual effects of having a venture capital investor on a new venture's growth. They confirm their hypothesis that there is a positive effect on a new venture's employee and sales growth from having a venture capital investor. Among other factors, they control for geographical and industry density to separate these effects from the pure venture capital effect. Similarly, Davila, Foster, and Gupta (2003) found that venture-capital-backed companies grow faster than their non-venture-capital-backed peers.

However, by analyzing 76 different samples, Rosenbusch, Brinckmann, and Müller (2013) come in their meta analysis to the conclusion that venture capital does not improve the funded companies' performance if one controls for the industry effect. They find an overall positive effect of venture capital, which is stronger for studies that do not control for the industry. The positive performance effect is not statistically significant in the studies that control for industry effects. Venture capitalists are good at selecting the most promising industries, but, within an industry, on average, they do not significantly increase performance. Similarly, some authors have not found a positive effect of the information-sharing by venture capitalists with their portfolio companies in a longitudinal analysis of venture-capital-backed companies (Busenitz, Fiet, and Moesel, 2004).

A possible explanation is that the value added of venture capitalists also depends on the new venture's receptiveness and need for advice (Barney et al., 1996). The performance effect might also depend on the venture capitalist's involvement. Gerasymenko, Clercq, and Sapienza (2015) found support for the hypothesis that greater involvement of the venture capitalist increases the performance of a new venture, especially when the venture capitalist has strong experience. The authors have analyzed 163 French new ventures and used a combination of surveys, interviews, and secondary data to test their hypothesis.

Fitza, Matusik, and Mosakowski (2009) show in a large-scale empirically analysis that the new venture's performance variance between two rounds of financing is explained by the new venture (26.3 percent) and the venture capitalist (11.2 percent). That means that, while the overall effect of venture capital on a new venture's performance might be small, there is a high degree of variance that depends on selecting the right venture capitalist.

Other effects from venture capital

Venture capitalists can have effects on other areas of new ventures as well. While we cannot list all of them here, we want to highlight a few. For instance, in an analysis of 173 Silicon-Valley-based startups, Hellmann and Puri (2002) demonstrate that venture-capital-backed companies are more likely to professionalize their human resources, for instance, by appointing an external CEO or introducing new functions, than their non-venture-capital-backed peers. In an analysis of a survey of 93 responding new ventures from the Netherlands, Wijbenga, Postma, and Stratling (2007) show that a venture capitalist's monitoring and service activities affect a new venture's own control systems and its financial success. The authors differentiate between the monitoring activities and the resources, such as knowledge or network contacts, that a venture capitalist provides. They show that, if a venture capitalist focuses on monitoring activities, additional monitoring of the new venture harms its financial performance, whereas resource-providing activities by the venture capitalist make a new venture's cost-controlling systems improve their financial success. Due to the relatively small number of participants and the focus on the Netherlands, the generalizability of these results is questionable.

2.1.3 Venture capitalists' performance

The performance of venture capital funds is highly dependent on the experience and composition of the venture capitalist's management team. For instance, Bottazzi, Da Rin, and Hellmann (2008) found that venture capital firms that are run by managers with prior industry experience are more actively involved in the startup's decisions and have superior performance. Similarly, Sorensen (2007) found that more experienced venture capitalists are more successful, as measured by their higher IPO rates, compared to less experienced investors. Firstly, their greater experience improves their advising ability and thus the startup's prospects. Secondly, they gain access to higher quality startups, as they are chosen as an investor by more startups who expect more benefits from highly experienced investors.

In addition, Walske and Zacharakis (2009) found in their analysis that venture capital funds are more successful when their founders have a consulting background, have senior management experience, and have worked in the venture capital industry. Entrepreneurial experience in the fund, in contrast, has been found to reduce the performance of the venture capital fund. They define a venture capital firm's success as the ability to raise the next funding.

Rational decisions, the basis for good performance, are more commonly found in international investors than in local investors. Devigne, Manigart, and Wright (2016) differentiate between the behavior of international versus domestic investors when their portfolio companies underperform. They find that international investors are less emotionally attached and face less pressure than their domestic peers. They are thus likely to terminate their relationship with and commitment to the startup. This effect is reduced if the international investor has a branch office in the country of the portfolio company, but they remain more likely to reduce their commitments than local investors. The analysis of 1,618 funding rounds by 1,060 venture capitalists thus shows that international investors make more rational investment decisions than their local peers.

Another way to improve a venture capitalist's performance is by learning. Clercq and Sapienza (2005) find that, when venture capital firms have little experience, there is greater distance between knowledge areas and little trust in the venture capitalist - startup relationship, which reduces the learning effect. The results are based on 298 US-based venture capital firms that responded to a survey. All the main measures are based on the survey.

2.1.4 Syndication

Venture capitalists often invest in a syndicate, via which two or more investors provide the desired capital, and knowledge, risk, and information is shared (Lerner, 1994; Kogut, Urso, and Walker, 2007). Almost all syndicated investments have a lead venture capitalist that often chooses its co-investors (Cumming and Dai, 2013). Typically, the lead venture capitalist also demonstrates the greatest involvement and has the greatest decision-making power in the syndicate (Wright and Lockett, 2003). There are four reasons for syndicate investments (Lerner, 1994). Firstly, the investment selection process is improved so that more promising startups are selected. The decision-making is improved if more parties are involved (Lerner, 1994), and they benefit from the social capital in the syndicate (Ter Wal et al., 2016). Syndicates can help to overcome the competition among venture capital firms for the best investments (Bygrave, 1987; Casamatta and Haritchabalet, 2007). Secondly, venture capitalists need to have syndicate partners in later rounds in order to maintain a constant equity share.² Thirdly, venture capitalists allow other investors to join the syndicate in final rounds before going public so that they can advertise with their investment in the successful firm. They expect to be invited by their syndicate partners to join final venture capital rounds. Lastly, there is a risk-sharing component, as the risk is divided among the syndicate members (Wilson, 1968), and the

²If, in a later round, a venture capitalist buys all the shares – with no syndicate partners – he or she would increase his or her equity ownership share, which is typically not desired.

syndicate allows the members to diversify their portfolio, thus reducing unsystematic risk (Bygrave, 1987).

Furthermore, Cumming (2006) demonstrate another reason for syndication. He shows that the adverse selection risk, that is, the risk that low-quality startups choose equity finance while high-quality ones prefer debt finance, can be overcome by syndicating the funding round. In a syndicate, information and experience sharing improves the screening process and thus reduces the risk of low-quality startups in the first funding round. He proves his hypothesis in a sample of Canadian funds and startups, because of the use of more diverse types of capital in Canada. While the general assumption is that most of the time convertible preferred equity is used in venture capital financing, this does not hold for Canada, where common equity, straight debt, and convertible debt are more often used (Cumming, 2005).

The reasons for syndication also depend on the region. Manigart et al. (2006) show that, in Europe, contrary to the US, venture capitalists syndicate for portfolio management reasons and less for selection and monitoring reasons. Adding value, for instance, by filling in resource or knowledge gaps, is most important for early stage investments. The authors used a questionnaire in six European countries and the US to test their hypotheses. In addition, a better legal system increases the probability of syndicated investments by providing better investor protection (Cumming, Schmidt, and Walz, 2010).

While many researchers have focused their work on forces affecting the creation of syndicates and the composition of these, only a few researchers have placed attention on syndication breakups. Zhelyazkov and Gulati (2016) show that, after a withdrawal from a syndicate, an investor's reputation is harmed and he or she is less likely to invest with its prior co-investors again. More details on this analysis can be found in section 2.4.

2.1.5 Summary

In general, receiving venture capital is considered positive for a new venture, and it is also the choice of the entrepreneurs to try and raise venture capital. However, some researchers have also found negative effects on a new venture, especially a moral hazard problem, which could result in an early IPO or even business failure from risk-taking (Gompers, 1996; Wasserman, 2003; Fischer and Pollock, 2004).

		Unit of analysis		
		Startup	Venture capitalist and syndicate	Venture capitalist - startup dyad
Explanatory variable	Cultural and categorical
	Structural – General
	Structural – Status
	Reputation
	Distances and others

FIGURE 2.1: Literature review framework

2.2 The literature review framework

The following literature review and the framework developed in this section are additions to the overview of the venture capital industry in section 2.1. We use an empirical point of view, structuring and summarizing relevant empirical work that uses network-based and related variables to explain relationships on the startup, venture capitalist, and venture capitalist - startup dyad level. We develop and explain our framework for this review, which is displayed in figure 2.1, in the following.

2.2.1 Unit of analysis

From an empirical point of view, the literature on venture capital can be split according to the unit of analysis. Firstly, the startup level unit of analysis is concerned with the effects of venture capital on individual new ventures in terms of, for example, internationalization, innovation, and performance. Typically, the dependent variable describes a startup's attribute. Secondly, the venture capitalist or syndicate level includes studies that focus on either individual venture capitalists or the group of investors, a syndicate, as the unit of analysis. Dependent variables can include, for instance, the performance of a venture capitalist, attributes of their portfolios, or the composition of a syndicate. Typically, these studies do not focus on a startup's attributes or the relationship between startups and venture capitalists. Thirdly, the venture capitalist - startup dyad level includes studies

that focus on the creation of ties, for instance, trying to explain which investors invest in what kinds of startups. It is possible that the dependent variable describes a startup attribute, for example, whether or not the new venture received another round of financing. Nevertheless, the analysis might be concerned with the relationship formation between the startup and the venture capitalist.

2.2.2 Network and related measures as explanatory variables

The other dimension of interest in our literature review concerns the variables used in the analysis to explain the proposed relationship. Our aim is to understand, which network-based variables and related measures are used. We focus this analysis on the variables relevant to the empirical analysis in this dissertation. In this section, we key in on the variables used and empirical implementation, only briefly touching upon the theoretical concepts. A detailed discussion of the theoretical concepts behind these measures follows in section 3.2.

Cultural and categorical measures

Cultural and categorical variables measure cultural embeddedness, which is how startups or venture capitalists fit into their cultural surroundings, via, namely, the norms, beliefs, and categories around them (e.g., DiMaggio and Powell, 1983; Scott and Meyer, 1983; DiMaggio and Powell, 1991). Often, various types of categories are used to measure how the focal actor fits into its surroundings. Other applications include the comparison of similarities, for instance, the industry distances between two venture capitalists or between a new venture and its venture capitalists. The categories used can be market or industry categories, patent categories, or others distinct categories, such as the differentiation between technology and science.

General structural measures

The general structural measures include all variables that are based on the focal actors' connections with other actors in the network. For example, this can include connections to competitors, suppliers, and research institutions. Most often in our analysis, we have been interested in the venture capitalist network that is formed by the venture capitalist's investments in a syndicated funding round. Exemplary measures are the count of syndicate members, prior co-investments, and the closure of structural holes in the network (for the latter, see Burt (1992); more details in section 3.2). We include all network measures that are not focused on status. We separated

status as it is a theoretically distinct construct and is mostly defined in a common way.

Status structural measure

Status is a sociological concept that creates a hierarchical order of actors, given their embeddedness in their social structures (Podolny, 1993). Status can be considered a quality signal for consumers in their decisions between different producers or, in our context, different venture capitalists (Podolny, 2001). Status cannot be gained by an actor on his or her own; thus, status has a generally accepted positive relationship with quality, although, some factors mitigate this relationship (Lynn, Podolny, and Tao, 2009). The status measure that we are interested in and that is widely used in the venture capital context is that which is based on the structural network of the focal actor, typically the venture capitalist. The network measure used is the eigenvector centrality score, which was first introduced by Bonacich (1987). The distinctiveness of this measure is that the status of one focal actor depends on the status of his or her direct ties, which also depends on their ties and so forth. In the following, we refer to this centrality score developed by Bonacich only as *eigenvector centrality*.

Reputation

Another status-related concept is reputation. While status is a social ranking, reputation is often based on actual or perceived performance and quality (Washington and Zajac, 2005). However, there is no common agreement on a measure in the venture capital context. Some researchers have used the venture capitalist's age (e.g., Gompers, 1996; Lee and Wahal, 2004), the venture capitalist's size (e.g., Lee and Wahal, 2004), or, most often, the number of IPOs that the portfolio companies of a focal venture capitalist attained up to a certain point in time as a measure of reputation (e.g., Lee and Wahal, 2004; Dimov and Milanov, 2010; Gu and Lu, 2014). Lee, Pollock, and Jin (2011) developed a time varying index based on six key variables: the number of startups in the portfolio, funds invested, funds raised, the number of different funds raised, the number of successful IPOs, and the company's age. While these are mainly size and portfolio measures, reputation can also refer to the standing a venture capitalist has, that is, its reputation for ethical behavior (Drover, Wood, and Fassin, 2014).

Similarly to status, reputation can be a sign of a venture capitalist's quality. For instance, Rider (2009) shows that reputable brokers in the venture capital industry are more likely to represent high-quality venture capitalists. They measure quality with a combined measure of status, fund size, and experience.

Distances and other variables

In addition to the variable classes listed above, we included some papers with additional variables in our structured literature review when we considered them relevant for our empirical analysis. This especially includes a number of papers that are related to geographical distance, as this has been found to be an important factor in the venture capitalist - startup relationship.

2.3 Literature-gathering approach

The goal of this literature review is to demonstrate the need for the empirical analysis in the latter part of this dissertation, thus focusing on the venture capitalist - startup relationship and how it is formed. Within the literature on the venture capital industry, we aim to give a focused overview of the empirical work highlighting advantages, disadvantages, and discrepancies and finding blind spots. For this purpose, we categorize the relevant journal articles based on the two-dimensional framework developed in section 2.2. Firstly, we focus on the unit of analysis, and, secondly, we consider the high-level methods used in the major articles with a particular focus on network-related approaches.

The literature review aims to highlight the empirical work and empirical methods used, which is why we only include articles that have used relevant measures or have contributed with their findings to these measures. Relevant measures include those that are network-based, for example, the syndication or categorical networks, and related measures like reputation or distance measures. These variables measure the cultural or structural embeddedness of organizations, which is the pivotal point of this dissertation. We include in our descriptions only the venture-capital-related findings and emphasize methodological aspects if they are relevant for our analysis. We omit findings with regard to the applied theories, for instance, general contributions to organizational institutionalism. This way, we focus on the relevant findings about the venture capital industry. Relevant conceptual advancements in sociology are outlined as needed for our analysis in chapter 3.

Furthermore, we have excluded papers that focus on the peculiarities of corporate venture capital, business angels, or later-stage private equity, as we are interested in contributions about the general venture capital industry after the seed stage. We have further neglected articles on specific human-capital-related topics, for example, the relevance of the management team, as we focus on the organizational matters. Furthermore, we have excluded papers that are concerned with financial details in deal structures or other contracting-related topics.

Journal name	VHB-JOURQUAL3
Academy of Management Journal (AMJ)	A+
Administrative Science Quarterly (ASQ)	A+
Entrepreneurship: Theory and Practice (ET&P)	A
Journal of Business Venturing (JBV)	A
Journal of Management (JOM)	A
Journal of Management Studies (JMS)	A
Management Science	A+
Organization Science	A+
Organization Studies	A
Research Policy (RP)	A
Strategic Entrepreneurship Journal (SEJ)	A
Strategic Management Journal (SMJ)	A
American Journal of Sociology	A
Annual Review of Sociology	A
American Sociological Review	A+

TABLE 2.1: Overview of primary journals used for structured literature review

We aim to give a correct and concise overview of the relevant research in the venture capital domain to make it relevant for other researchers (Short, 2009). In order to focus our efforts, we are only using relevant high-quality journals. We selected the relevant journals mainly from the 50 publications used by the Financial Times (FT) to compile the FT business school research rank in 2016 (Ormans, 12. September 2016). The list is annually reviewed and adapted based on a survey among 200 universities. We selected the journals that publish empirical work in the fields of management, organization, and entrepreneurship. We thus excluded journals focusing on marketing, finance, human resources, accounting, and economics. We added the three major sociological journals, *American Journal of Sociology*, *Annual Review of Sociology*, and *American Sociological Review*, because this dissertation is built upon sociological theory, and thus sociological work anchored in the venture capital domain needs to be included. Table 2.1 lists all primary journals used in the structured literature review. As a reference, the table also includes the VHB-Jourqual ranking version 3.0.³ The VHB-Jourqual is a ranking of business-related journals conducted by the members of the *Verband der Hochschullehrer für Betriebswirtschaft e.V. (VHB)*, an organization with more than 2,000 mainly German-speaking members that are active researchers in the business administration domain.

We limited our analysis to articles published after 1990, as we are interested in the latest academic advancements and the most major contributions to the venture capital literature have been made since the 1990's

³For the complete list, please refer to <http://vhbonline.org/vhb4you/jourqual/vhb-jourqual-3/gesamtliste/>

(Gompers and Lerner, 2001). We used *Web of Science*⁴ as our database and searched for "venture capital" in our query on the research papers' titles, abstracts, and key words. This resulted in 415 articles, which we briefly reviewed and filtered based on the topic relevance and variables used. Next, we scanned the reference sections of the most relevant papers and then scanned, in an unstructured manner, further journals for relevant research. We added these additional articles.

2.4 Review based on variables

Based on the framework developed in section 2.2 and the literature-gathering approach described in section 2.3, we categorized all relevant papers so that each paper is presented and classified only once. The first dimension is a distinct classification based on the unit of analysis, typically the subject of the dependent variable, in the respective paper. In the second dimension, however, a paper could be classified into more than one category if the authors used measures that fall into multiple classes to test their findings. In these cases, we selected the field in the matrix that is most relevant for the paper. For instance, if the paper is anchored in the structural embeddedness literature but also uses reputation as a predictor, we would classify it into the structural variable category. The summary of the categorization into the framework is depicted in figure 2.2, in which we have also included the name of the most relevant independent variable. In the following, we briefly outline the main findings and detail the measures used in each of the fields of the matrix. We go through the matrix column by column.

2.4.1 A. Startup focus

i. Cultural and categorical

Some authors analyze cultural categories in a literal way while others focus on a more categorical perspective. For example, Dai, Jo, and Kassidieh (2012) analyze cross-border investments in Asia and demonstrate, that foreign venture capitalists typically invest larger funding amounts in later rounds, as they lack the monitoring and information-gathering abilities of their local peers. It follows that non-domestic investors prefer new ventures on which information is easily obtainable. The analysis shows that a successful exit is most likely if a new venture has cross-boarder and domestic investors. The authors measure uncertainty on three levels. The early funding rounds, the early development stages of the new venture,

⁴www.webofknowledge.com

and high-technology industries are considered the most uncertain investments from a venture capitalist's perspective. They also use a measure of "cultural distance" (Dai, Jo, and Kassiech, 2012, p. 670) between the country of the new venture and that of the venture capitalist based on a general four-category culture measure. The analysis shows that the higher the cultural distance between the venture capitalist and the new venture, the less likely a successful exit becomes.

ii. Structural – General

Stuart, Hoang, and Hybels (1999) demonstrate that venture capitalists facing uncertainty about the quality of potential new ventures use the structural embeddedness, or "prominence" (p. 327) as the authors call it, of a startup's existing partners as a guide in their assessment. They also show that these companies are typically more successful in IPOs, measured as time until IPO, and the valuation at that time, than companies with less prominent partners. Their sample includes 301 biotechnology new ventures. The authors use two network-based measures: The centrality in a patent co-citation network is used to measure technological prominence and commercial prominence is based on the centrality in the prior strategic-partner network.

Similarly, partner networks have been found to benefit new ventures. Baum and Silverman (2004) measure a new venture's performance by its revenue, research and development spending, employment growth, patenting activities, and survival rate. In a sample of Canadian biotechnology firms, the authors show that downstream and horizontal partner networks, that is, the "alliance capital" (Baum and Silverman, 2004, p. 421), for instance, to research institutions or customers, increase the amount of venture-capital funding, as these networks are considered a quality signal. Partner networks are measured by a simple count of alliances. A new venture's "intellectual capital" and "human capital" (p. 422) also increase the venture capital financing. In terms of a startup's performance, partner networks and intellectual capital have a positive performance impact, whereas human capital does not.

Clercq and Dimov (2008) find that the industry experience of venture capitalists increases a new venture's performance especially when the new venture lacks access to information. The "number of syndicate partners" and "number of prior interactions" (Clercq and Dimov, 2008, p. 595) of the syndicate members also have a positive performance effect, especially when the venture capitalist lacks industry experience. The authors measure the "industry knowledge" (p. 595) as the number of investments in the same industry category as the focal new venture. For structural network-related measures, they count the members of the syndicate and the prior

Unit of analysis	
Startup	Venture capitalist and syndicate
<p>"cultural distance" Dai, Jo, and Kassaciéh (2012), p. 670</p> <p>"prominence" Stuart, Hoang, and Hybels (1999), p. 327</p> <p>"alliance capital" Baum and Silverman (2004), p. 421</p> <p>"number of syndicate partners" Clercq and Dimov (2008), p. 595</p> <p>"number of first-order ties" Jääskeläinen and Maula (2014), p. 712</p> <p>"indirect ties" Pahnke, McDonald, et al. (2015), p. 1,343</p> <p>"measurement of the relational mix" Lechner, Dowling, and Welpe (2006), p. 529</p>	<p>"specialized industry knowledge" Dimov and Clercq (2006), p. 213</p> <p>"Geodesic length" Sorenson and Stuart (2008), p. 280</p> <p>Guler and Guillen (2010)</p> <p>"relational embeddedness" Meuleman et al. (2010), p. 1,006</p> <p>"direct relationships" Hopp and Lukas (2014), p. 647</p> <p>"giant component" Kogut, Urso, and Walker (2007), p. 1,184</p> <p>"overall withdrawal rate" Zhelyazkov and Gulati (2016), p. 285</p> <p>"brokerage advantage" Guler and Guillen (2010), p. 399</p> <p>"syndication frequency" Jääskeläinen, Maula, and Seppä (2006), p. 193</p> <p>"network embeddedness" Echols and Tsai (2005), p. 226</p> <p>"degree centrality", p. 256; "closeness", p. 257; "betweenness", p. 258; Hochberg, Ljungqvist, and Y. Lu (2007)</p>
<p>"high-status" Pahnke, Katila, and Eisenhardt (2015), p. 17</p> <p>"VC prominence" Ozmel, Reuer, and Gulati (2013), p. 857</p> <p>"Social defenses" Hallen, Katila, and Rosenberger (2014), p. 1,087</p>	<p>"status", "structural holes" Podolny (2001), p. 54</p> <p>"VC firm status" Dimov and Milanov (2010), p. 337</p> <p>"VCF status" Milanov and Shepherd (2013), p. 735</p> <p>"power source match" Ma, Rhee, and Yang (2013), p. 719</p> <p>"complementary status" Bothner, Kim, and Lee (2015), p. 592</p> <p>"status" Dimov, Shepherd, and Sutcliffe (2007), p. 491</p> <p>"status" Pollock, Lee, et al. (2015), p. 11</p>
<p>"VC firm reputation" Cumming and Dai (2013), p. 1,008</p> <p>"VC reputation" Fernhaber and McDougall-Covim (2009), p. 283</p> <p>"prior VC IPOs" Hsu (2006), p. 210</p> <p>"VC prominence" Gulati and Higgins (2003), p. 134</p> <p>"Prestigious venture capital firms" Pollock, Chen, et al. (2010), p. 13</p> <p>"VC reputation" Nam, Park, and Arthurs (2014), p. 11</p>	<p>"VC firm reputation" Petkova, Wadhwa, et al. (2014), p. 432</p> <p>"total reputation of other syndicate members" Clercq, Sapienza, and Zaheer (2008), p. 1,182</p> <p>"reputation" Gu and Lu (2014), p. 743</p> <p>"IPO success rate" Matusik and Fizza (2012), p. 420</p> <p>"the age of the lead venture capital firm" Compers (1996), p. 139</p>
<p>"Number of neighboring VCFs" Kolympiris, Kalaitzandonakes, and Miller (2011), p. 1,191</p> <p>"travel time" Bernstein, Giroud, and Townsend (2016), p. 1,599</p> <p>"sensegiving activities" Petkova, Rindova, and Gupta (2013), p. 873</p>	<p>"domestic, opportunistic international, and strategic international" LiPuma and Park (2014), p. 8</p> <p>"venture-level uncertainty", "country-level uncertainty" Liu and Maula (2016), p. 1,415</p>
<p>Cultural and categorical</p>	<p>Venture capitalist - startup dyad</p> <p>"fuzziness", p. 92, "leniency" p. 93, Pontikes (2012)</p> <p>"hybridization via patenting" Wry, Lounsbury, and Jennings (2014), p. 1,318</p> <p>"category dispersion", p. 123, "category similarity", p. 124, Wry and Lounsbury (2013)</p> <p>"network closure", "knowledge similarity" Ter Wal et al. (2016), p. 408</p> <p>"direct tie", "indirect tie" Hallen (2008), p. 698</p> <p>"social ties" Shane and Cable (2002), p. 372</p> <p>"relative strength of strong ties" Wuebbker, Hampf, and Wüstenhagen (2015), p. 174</p>
<p>Structural – General</p>	<p>"relative standing", p. 2,047, "status", p. 2,048, Ozmel and Guler (2015)</p>
<p>Structural – Status</p>	<p>"high industry deal experience" Hsu (2004), p. 1,824</p> <p>"ethical reputation" Drover, Wood, and Fassin (2014), p. 729</p>
<p>Reputation</p>	<p>"industry distance", p. 1,563, "geographic distance", p. 1,565, Sorenson and Stuart (2001)</p> <p>"industry fit" Lungeanu and Zajac (2016), p. 939</p>
<p>Distances and others</p>	

Explanatory variable

FIGURE 2.2: Literature review overview

investments of the focal venture capitalist with the other members of the syndicate.

Jääskeläinen and Maula (2014) demonstrate that direct and indirect ties to other venture capitalists serve different purposes. The "number of second-order ties" (Jääskeläinen and Maula, 2014, p. 712), that is, syndicate partners of directly connected venture capitalists, increases the access to information about startups and thus improves the screening and identification process. Direct tie information access, the "number of first-order ties" (p. 712), allows better evaluation of the quality of the new venture. It follows and the authors show that cross-border investments are thus more likely if respective ties exist. They analyze the effect of network proximity or distance on the probability of non-domestic exits in a sample of venture-capital-backed European new ventures. Their independent variables are based on the syndicate networks: the number of direct non-domestic ties from prior syndicated investments and the number of indirect ties to non-domestic investors through their direct connections.

Pahnke et al. (2015) focus on new ventures' ties to competitors through venture capitalists and how these ties affect the number of FDA approvals for products from the medical device companies in their sample. The authors determine whether or not a new venture has an indirect tie to a competitor through one of its venture capital investors, and they assign market categories (sub-categories in the medical device field) to each of the companies to determine whether or not they are competitors. Contrary to other researchers who find positive effects of venture capital on innovation, Pahnke et al. (2015) show that being venture-capital funded can reduce innovation. When competing firms are linked to each other by the same venture capitalist, that is, "indirect ties" (Pahnke et al., 2015, p. 1,343), their innovative power is lower than that of firms that are not linked to competitors. They argue that information is shared between competing firms in the network through the venture capitalist and that the firms cannot control this information leakage. The firms are then not able to benefit from their innovations in the way that their competitors can. Among other factors, they find partial support for their hypothesis that high-status venture capitalists are more likely to leak information than low-status investors. It is important to note that, as the authors themselves identify as a limitation, the analysis of only one industry highly limits the generalizability of the results of this study. It might be true that innovative power in other industries is increased when information is leaked.

Lechner, Dowling, and Welpé (2006) do not use network based measures, but they use interviews to support their hypothesis that there is a

positive effect of the various networks, such as family networks, organizational networks with competitors, and investor networks, on the performance and development speed of new ventures. However, the mix of these network partners, that is, the "measurement of the relational mix" (Lechner, Dowling, and Welppe, 2006, p. 529), is important and changes over time. The authors, using a questionnaire among 50 new ventures, also find support for their hypothesis that new ventures enter network connections with reputable partners to increase their reputation.

iii. Structural – Status

Pahnke, Katila, and Eisenhardt (2015) demonstrate that independent venture capitalists with "high-status" (p. 17) foster innovation in new ventures, whereas corporate venture capitalists' institutional way of doing things prohibit improving the innovative and patenting power.⁵ Innovativeness is measured by the number of patents (technical) and number of approved products (commercial). The authors thus strongly support the thesis that venture capitalists provide more than just capital, and the choice of the investor can significantly affect the prospects of a new venture. They test their hypotheses in the minimally-invasive surgical device industry.

Ozmel, Reuer, and Gulati (2013) demonstrate that young biotechnology firms have a positive signaling effect, as measured by creating research and development alliances via affiliation with central venture capitalists, similar to the positive effect a new venture receives from a position central to other companies in its alliance networks. The latter can substitute the effect of central venture capitalists: For example, if a firm already has a central position in the alliance network, the additional signaling effect of central venture capitalists is weak. They measure the "VC prominence", that is, the centrality of the venture capitalist, and the new venture's centrality, "alliance prominence" (Ozmel, Reuer, and Gulati, 2013, p. 857, 858), using the eigenvector centrality score that other authors use to measure status.

The relationship between venture capitalists and startups is not only characterized by the obvious resources given by the venture capitalist to the new venture. Another angle on the eigenvector centrality of independent venture capitalists is to consider it a "social defense" (Hallen, Katila, and Rosenberger, 2014, p. 1,087). It thus also serves as protection against threats from larger competitors, for instance, corporate venture capitalists who are considered a threat to a startup's intellectual property and growth. Hallen, Katila, and Rosenberger (2014) find that the central independent venture

⁵Contrary to these findings, Park and Steensma (2013) find that corporate venture capitalists increase innovativeness compared to startups with pure independent investors especially when the corporate investors are more reputable than the independent venture capitalists.

capitalist makes startups attractive for the corporate venture capitalist. If startups lack other defense mechanisms, like timing and secrecy defenses, the effect is even stronger. In addition, the greater the geographical distance between a startup and a venture capitalist, the higher the probability that a corporate venture capitalist invests. The authors use the funding round as the unit of analysis in their sample of 700 technology startups with 4,023 funding rounds. They measure the geographical distance between the startup and the venture capitalist and use the maximum value for centrality and the minimum value for distance in each syndicate. Furthermore, they consider heterogeneity among corporate venture capitalists in terms of research and development spent and whether or not the corporate venture capitalist is in the same industry as the startup.

iv. Reputation

As young companies mature, they need to set up their operations and choose their partners to suit their needs accordingly. Cumming and Dai (2013) find that it might be beneficial for a startup to change their lead investor. If they do, they are able to raise larger amounts of capital with better valuations. However, if they choose to switch to a more reputable venture capitalist, this comes at a cost. In these cases, funding rounds are smaller and subsequent valuations lower. As more reputable investors choose high-quality startups (Sorensen, 2007), the new venture benefits from a positive quality signal from the high-reputation investor. Cumming and Dai (2013) measure "VC firm reputation" (p. 1,008) based on historic IPO capitalization.

Besides the effect on innovation, a venture capitalist's reputation has been found to affect the strategic decisions regarding internationalization of new ventures. Fernhaber and McDougall-Covin (2009) show that the higher the venture capitalist's international knowledge and the stronger the reputation, the more international the new venture becomes. For the individual startup, it follows that, depending on the internationalization strategy, it is important to find an investor who can bring the required knowledge and experience to the startup. The researchers measure the degree of internationalization with a composite measure based on the share of international sales, share of international assets, and number of business regions. They measure "VC reputation" (Fernhaber and McDougall-Covin, 2009, p. 283) based on the venture capitalist's prior performance, experience, and media attention. The "international knowledge" (p. 284) of the venture capitalist is determined from the international experience of the board member representing the venture capitalist.

Highly reputable venture capitalists also have an effect on the partnerships and commercialization of new ventures. Hsu (2006) finds that new

ventures are more likely to have strategic partnerships and use licensing for commercialization when the new ventures are venture-capital funded and not funded by government programs. This effect is even stronger the higher the venture capitalist's reputation is. For the startups that are venture-capital backed, the author measures the venture capitalist's reputation based on "prior VC IPOs" and presents the eigenvector centrality score, "Bonacich VC centrality" (Hsu, 2006, p. 210), as an alternative reputation measure. Interestingly, contrary to most other researchers, the author does not distinguish between status and reputation here. Rather, he uses the standard status measure as an alternative to other more common reputation measures.

Gulati and Higgins (2003) show that new ventures have more successful IPOs, as measured by a combination of net proceeds and valuation measures, when they have prominent venture capitalists in unfavorable markets or prominent investment banks in favorable markets. While the authors consider the direct, structural ties of the biotechnology companies in their sample, they measure "VC prominence" not by standard network measures, but by a ranking of the dollars invested of the venture capitalists, and "Underwriter prestige" (Gulati and Higgins, 2003, p. 134) by a ranking based on tombstone information.

The additional benefit of highly prestigious venture capitalists based on fund size for IPO valuations is diminishing but still curvilinear with the number of other highly prestigious partners, like executives or underwriters (Pollock et al., 2010). Pollock et al. (2010) count the "number of prestigious venture capital firms" (p. 13) and consider a venture capitalist to have high prestige if he or she is among the top ten in a yearly ranking based on the size of new funds raised.

A venture capitalist's reputation can also affect a new venture's finance function. Nam, Park, and Arthurs (2014) show in their analysis that monitoring by reputable venture capitalists reduces earnings management of the funded companies. However, generally, the existence of a venture capital investor increases the likelihood of earnings management. The authors use working capital accruals as a measure of earnings management and analyze the abnormal returns at the point when the lock-up period ends. The authors measure "VC reputation" (Nam, Park, and Arthurs, 2014, p. 11) purely based on the number of prior IPOs of the focal venture capitalist.

v. Distances and others

The geographical distance between venture capitalists and the new venture is also an important factor for a new venture's development and success. Kolympiris, Kalaitzandonakes, and Miller (2011) show that, within a limited regional proximity of 10 miles between venture capitalists and

biotechnology startups, in so-called 'clusters', the focal new ventures receive larger amounts of venture capital. As explanatory variables, they use the "number of neighboring VCFs" (Kolympiris, Kalaitzandonakes, and Miller, 2011, p. 1,191) around a focal firm within a certain radius and control for fund size, fund age, and the number of syndicate partners of the venture capitalists invested in a certain new venture.

In a large-scale empirical analysis, Bernstein, Giroud, and Townsend (2016) show that easier travel routes resulting in shorter "travel time" (p. 1,599) between new ventures and venture capitalists lead to an increase in the number of granted and cited patents issued by the startup. This effect is strongest for lead venture capitalists. New travel routes also increase the likelihood that the startup will go public and the venture capitalist will have a successful exit. They confirm their findings with a survey among venture capitalists that shows the venture capitalist's expectation that a direct flight connection to the startup location would increase the frequency of on-site visits and lead to an improved understanding of the business. It is important to note that they restrict their analysis to a time-frame between 1977 and 2006. The impact of the widespread use of higher-quality and lower-cost video conferencing systems, for instance, *Skype*, and cloud-based working solutions to monitoring startups might not have been fully captured in the analysis.

Petkova, Rindova, and Gupta (2013) analyze one method of early-stage new ventures to gain legitimacy: the use of communication and "sense-giving activities" (p. 873). They demonstrate that the general use of these methods and the use of different activities increase general and industry media attention, with the latter also positively influencing the amount of venture capital received. Media attention is based on the average number of media articles per month after foundation. The authors analyze communication activities, for instance, press statements, or other dissemination formats, for example, presentations or events, and how diversely these are used.

2.4.2 B. Venture capitalist and syndicate focus

i. Cultural and categorical

Dimov and Clercq (2006) use industry categorization and investment stages to measure the impact of specialization on portfolio survival. In a sample of 200 US funds, they find that venture capitalists with an investment-stage focus have higher portfolio survival rates (a lower default proportion) than firms without such a focus, due to the increased expertise. The effect is not significant for industry specialization. Failure rates in

venture capitalists' portfolios are higher for venture capital firms that typically invest in a syndicate. The greater knowledge and experience available in a syndicate and the greater total involvement of the syndicate partners in their sample do not outweigh each individual investor's lower commitment to the funded startup. For instance, when a startup does not perform as expected, syndicate members reduce financial and other forms of involvement, and the startup becomes the "living dead" (Ruhnka, Feldman, and Dean, 1992, p. 138). Dimov and Clercq (2006) measure specialization strategy in two domains: "the specialized stage knowledge" and "the specialized industry knowledge" (p. 23). They use six stage and nine industry categories to calculate the Herfindahl-Hirschman index, that is, sum of the squares of the share of investments in a given stage or industry, for the complete distribution of the share of their first-time investments. Syndication is measured as the average number of co-investors in first-time funding rounds, as most of the deals are syndicated, but with a different number of investors.

ii. Structural – General

Sorenson and Stuart (2008) demonstrate that the setting of the investment also determines how syndicate structures form. Geographically- and industrially-distant venture capitalists are more likely to form a syndicate if the new venture is located in a region and industry with a lot of investment activity, and if the new venture is in a later development stage. With regard to the syndicate, it helps to mitigate the distances when potential syndicate partners already have ties from prior investments and are well-connected through indirect ties, and when the syndicate is large. From a strictly methodological perspective, one could argue that the authors use the dyad of new venture and venture capitalist as the unit of analysis. Nonetheless, we consider the unit of analysis to be on the syndicate level. They refer to the triad of two venture capitalists and a startup. They model the probability that the second venture capitalist joins, given the setting of the new venture. It is thus the dyad between the two venture capitalists that is in the focus of the analysis. A number of structural network-based measures are used. For instance, they consider direct ties, which are mutual investments in a funding round, and indirect ties if they have previously had the same co-investors in a measure of "geodesic length" (Sorenson and Stuart, 2008, p. 280). They also calculate the geographical distance in miles and the industry distance by comparing the share of investments in the given industry between two venture capitalists. They also calculate these distance measures between the investors and the target firm.

The same research approach using a triad of two investors and one investment target is applied by Meuleman et al. (2010). The authors show

that an investor is more likely to join a certain syndicate when it has previous ties to the lead investor, and the effect is weakened if the non-lead investor has a strong reputation. Interestingly, they do not find support for the hypothesis that "knowledge complementarity" (Meuleman et al., 2010, p. 1008) increases the likelihood of joining a certain syndicate. The most relevant independent variables are "relational embeddedness" (p. 1,006), a count of the direct ties between the two venture capitalists, knowledge complementarity, defined as the difference between the two venture capitalists' share of investments in the new venture's industry, and "reputation of potential partner" (p. 1,008), which is measured by the number of IPOs in a focal venture capitalist's portfolio. The authors use the buyout segment, which is outside of the venture-capital focus of this study. Nevertheless, we include it here on account of the relevance of the applied measures.

Syndicate formation also depends on signals sent by potential syndicate partners: In an analysis of the German venture capital market, Hopp and Lukas (2014) show that lead investors are more likely to choose other venture capitalists to join their syndicate if these investors have greater experience, have previously co-invested, and continuously show these signals over time. The effect is greatest in later funding rounds, whereas in early rounds, diversification is the main goal. The authors analyze the investors of a certain startup, which includes the investors of the current round and from previous rounds. They measure "industry experience" as the number of deals in the past year in the focal industry and analyze "direct relationships" (Hopp and Lukas, 2014, p. 647), considering whether a lead investor has invited a certain venture capitalist before and whether this invitation was reciprocated.

Kogut, Urso, and Walker (2007) conduct an in-depth longitudinal analysis of the development of network structures in the venture capital industry. Their analysis shows that investors who invest in similar industries tend to cluster together in groups, resulting in a so called "giant component" (Kogut, Urso, and Walker, 2007, p. 1,184), which is caused by the effect that central venture capitalists tend to syndicate with venture capitalists with whom they have previously co-invested, so they can trust them.

Evaluating the dyad between two investors, Zhelyazkov and Gulati (2016) analyze the effects of a syndicate break-up on future tie formations. They analyze the number of times the focal investor has withdrawn from a syndicate that his or her dyad partner continued to invest in as a share of the total number of syndicates of the two investors and the rate of total withdrawals from all syndicates, the "overall withdrawal rate" (Zhelyazkov and Gulati, 2016, p. 285). They measure "social overlap" (p. 285) as structural embeddedness based on the number of shared contacts in a standardized

Jaccard index measure. In a sample of US venture capital investments between 1985 and 2008, the researchers show that, after a withdrawal from a syndicate, an investor is less likely to invest with his or her prior co-investors again. The withdrawal also affects the overall reputation of a venture capitalist such that he or she becomes a less attractive syndication partner. On the one hand, this is caused by the publicly known fact that the venture capitalist withdrew from a syndicate. On the other hand, prior co-investors will distribute negative information through informal information flows in their network, further harming the reputation of the venture capitalist who withdrew. They control for industry-specialization distance based on a 10-industry categorization by calculating a distance measure introduced by Sorenson and Stuart (2008). This measure compares the share of investments made in the focal industry with that of the syndicate partner.

Structural embeddedness also affects a venture capitalist's internationalization: In an analysis of 1,010 US venture capital firms, Guler and Guillen (2010) find that their home country syndication networks affect the internationalization behavior of venture capitalists. Venture capitalists with high status can also carry this quality signal to foreign markets, whereas "brokerage advantage" (Guler and Guillen, 2010, p. 399), that is, information access and transfer, is local and cannot be transferred. If a local network partner is already present in a foreign market, brokerage increases the likelihood of internationalization. They measure status with the standard eigenvector centrality score and apply the brokerage measure developed by Burt (1992) based on the syndication networks. They also control for whether or not syndicate partners in a certain country exist.

Jääskeläinen, Maula, and Seppä (2006) use syndicate structures to measure a venture capitalist's involvement. The authors analyze a sample of 94 major US venture capital firms and demonstrate that a venture capitalist's involvement in their portfolio company improves their performance and affects the optimal portfolio size. Their analysis shows that a venture capitalist's "allocation of attention" (Jääskeläinen, Maula, and Seppä, 2006, p. 192) improves a new venture's performance and that, to a certain degree, increased portfolio size improves the venture capitalist's performance. There is an optimal portfolio size for which the benefits from involvement and increased diversification maximize performance. The optimal size increases in a syndicate especially if the venture capitalist is not the lead investor, as the required time per portfolio company is reduced. The authors measure the performance as the number of IPOs in their portfolios and the attention as the number of managing partners per portfolio company. They also take into account the "syndication frequency" and "syndication role" (p. 193), that is, whether or not the investor was the lead investor.

Structural network embeddedness can also have a direct impact on a venture capitalist's performance: Echols and Tsai (2005) analyze venture capitalists' performance in an 80-firm sample, depending on the niche focus and the company's "network embeddedness" (p. 226). The analysis shows that a product niche focus and a process niche focus have a positive performance impact when the venture capitalist is highly embedded in the network. The greater the share of IPOs in the same industry as compared to other venture capitalist's IPOs, the smaller the "product niche" (Echols and Tsai, 2005, p. 225) focus of the focal venture capitalist is. In terms of "process niche" (p. 226), the authors compare the shares of all investments depending on the development stage, to those of all other venture capitalists. Network embeddedness is measured by a network redundancy score developed by Burt (1992) divided by the network size to account for network differences.

Similarly, Hochberg, Ljungqvist, and Lu (2007) find that more structurally-embedded venture capitalists have better performance, as measured by the number of IPOs in their portfolios. The authors use five different centrality measures to analyze network embeddedness including three measures for "degree centrality", eigenvector centrality as a measure of "closeness", and "betweenness" (Hochberg, Ljungqvist, and Lu, 2007, p. 256, 257, 258). Specifically, they find that connections to the most embedded venture capitalists are more important than bringing other previously unconnected venture capitalists together. The authors also find support for their hypothesis that venture capitalists can provide better support to their portfolio companies, for instance, by coaching, such that central venture capitalist's portfolio companies receive more funding and have a greater likelihood of going public.

iii. Structural – Status

One of the very influential papers in the venture capital industry was written by Podolny (2001). He differentiates between uncertainty that a startup has in how to create the best company and the uncertainty an investor faces about the success of the startup. Regarding the former, the closing of "structural holes" (Podolny, 2001, p. 54), for example, by connecting previously unconnected investors through a syndicate, helps to mitigate this uncertainty but not the uncertainty, about the startup. The structural-holes aspect relates to the flow of resources, whereas "status" (p. 54) has an informative role. Uncertainty about the startup's quality is mitigated by the status of the venture capitalist. This implies that investors who can close structural holes should invest in markets with high uncertainty where individual startups can convert their inputs into the best output, whereas high-status investors should be active in markets with high uncertainty concerning the quality of the startup output. Podolny follows

Burt (1992) to calculate the degree of structural holes and uses the eigenvector centrality score for status to analyze the effect on the investment stage of a venture capitalist's portfolio based on a three stage categorization of the portfolio companies.

Dimov and Milanov (2010) show that the decision to syndicate an investment depends on a venture capitalist's prior experience, status, and reputation. They differentiate in their analysis between a venture capitalist's uncertainty regarding the monitoring and selection of new ventures and uncertainty regarding the quality of the syndicate partner. The authors show that the need for syndication is higher in novel industries, but it is harder to find syndication partners. While a good reputation cannot easily be carried over as a quality signal to attract syndicate partners, status can be used as the required quality signal. The authors measure the non-novelty of a given startup segment to a venture capitalist by the number of times the investor has previously invested in the focal industry. They use a nine-industry segmentation. The authors measure "VCF status" (Dimov and Milanov, 2010, p. 337) with the eigenvector centrality score and calculate reputation as a composite of the venture capitalist's age, the total number of investments, and the number of IPOs.

The first syndicate partners plays a special role: Milanov and Shepherd (2013) analyze the effect of the status of the first syndicate partner from a medium- and long-term perspective for a focal venture capitalist. Building on categorization literature, the authors argue that venture capitalists are categorized based on their social and human capital. This type of capital creates legitimacy for the firm up to the point that it has created its first alliances. The first alliance through the first venture capitalist syndicate co-investor will thus shape the focal venture capitalist's status. They propose that the future "VCF status" is increased when the "first partner's reputation" is strong, especially if the first partner had a dense network prior to the focal funding round, that is, "past cohesion" (Milanov and Shepherd, 2013, p. 735, 735, 736). The authors measure reputation as the number of IPOs relative to the number of companies in the investment portfolio. The study confirms the hypothesis in a sample of 272 new-entry venture capitalists.

Contrary to Podolny (2005), who argues that a venture capitalist's status is mainly a signal, Ma, Rhee, and Yang (2013) argue that status also affects the way members of a syndicate behave and cooperate based on the "power source match" (p. 719). Effective syndicates have power parity from a similar ownership share, power parity from similar status, or a dominant ownership advantage. However, low-status investors with high ownership shares seek high-status co-investors, but this status mismatch will limit the

effectiveness of the syndicate. The effect is weakened by investors' familiarity among each other and by the entrepreneurial performance. They test their hypotheses in a sample of 1,053 biotechnology ventures in the US, measure success as time and the rate of IPO or M&A exit in the portfolio of a venture capitalist, and calculate status with the eigenvector centrality score. The authors calculate the coefficient of the variation of status and ownership share for each pair of investors, "familiarity" (Ma, Rhee, and Yang, 2013, p. 717) based on the number of direct ties, and performance of the new venture based on the number of patents.

For syndication network settings such as the venture capital domain, Bothner, Kim, and Lee (2015) develop a twofold status framework with an additional measure of status, namely, the "complementary status" (p. 592), which is the status of the co-investors. They define "primary status" (Bothner, Kim, and Lee, 2015, p. 592) as the 'standard' status measure for the lead investor. Their results show that complementary status has a positive effect on venture capital firm survival, while it reduces the effect of primary status.

The effects of status and reputation have been analyzed together, and researchers have shown that clear differences between these concepts exist. Dimov, Shepherd, and Sutcliffe (2007) find that venture capitalists with strong finance expertise are less likely to invest in early-stage startups. The effect is stronger for high-"status" venture capitalists, but reversed for high-"reputation" (Dimov, Shepherd, and Sutcliffe, 2007, p. 491) venture capitalists. The authors measure the "finance capacity" (p. 490) in a venture capital team by the relevant work experience of the team members. Reputation is calculated based on the number of portfolio companies, invested capital, age, and media attention. Status is measured by the eigenvector centrality score.

In an analysis of more than 400 venture capital firms, Pollock et al. (2015) show that reputation and status affect each other, with reputation having a greater effect on status than the reverse. They also find that the past magnitude has a stronger effect on the future magnitude for reputation than for status, especially when firms get older. Status is measured by the eigenvector centrality, and the researchers use a modified version of the reputation index developed by Lee, Pollock, and Jin (2011). Among other things, the authors control for "structural holes" and "industry diversification" (Pollock et al., 2015, p. 13, 14).

iv. Reputation

Reputation is important for individual venture capitalists, affecting the investments decisions and strategies of these investors. Petkova et al. (2014) find that, with increasing "VC firm reputation" (p. 432), venture capitalists

are more likely to invest in an emerging industry sector that is interpreted as risky. At the same time, venture capitalists use a number of measures, for instance, increasing the number of syndicate members, investments in later stages, and lower commitment based on the share of clean technology new ventures in the portfolio, to reduce their overall risk in the case of better reputation, which causes pressure to perform well. The authors use a sample of venture capitalists, of which 4.8 percent invested in the clean energy sector. They consider an emerging sector risky due to the lack of knowledge about the prospects of the sector, in addition to the general startup-related risk. The venture capitalists thus have to make investment decisions in ambiguity. They explicitly differentiate reputation from other measures, like status and "sector legitimacy" (Petkova et al., 2014, p. 432). The latter is based on the mentioning of clean technology in newspapers and articles, government incentive programs, and the number of other venture capitalists that have invested in the clean technology sector. They use a composite reputation measure based on five factors, such as the number of historic investments, amount invested, and total funds raised, similarly to the index developed by Lee, Pollock, and Jin (2011).

Clercq, Sapienza, and Zaheer (2008) find in their analysis that a venture capitalist's involvement decreases with the "reputation of focal VCF" and with the "total reputation of other syndicate members" (p. 1,181, 1,182). However, involvement increases with a high ownership share compared to the other investors of the focal venture capitalist. The authors analyze involvement based on a survey among venture capitalists and measure reputation as the number of IPOs in the focal venture capitalist's portfolio.

Gu and Lu (2014) analyze whether or not a deal was syndicated and explain this using the focal venture capitalist's "reputation" (p. 743) as measured by the number of IPOs. The authors analyze venture capitalists from 16 countries, including the US, major European countries, Japan, and China (which they consider the local country). They find that a venture capitalist's reputation increases the probability of syndication up to a point of strong reputation, where the probability decreases. The authors thus find evidence of a U-shaped relationship. The reputation threshold for the decrease in syndication probability is lower for non-local venture capitalists and more mature institutions, which means that strong-reputation venture capitalists decide sooner against syndication.

As seen above, often the number of IPOs in a venture capitalist's portfolio is used to measure reputation. Matusik and Fitza (2012) do not call it reputation; instead, they use the "IPO success rate" (p. 420) as a performance indicator. They show that there is a U-shaped relationship between the industry diversification in the portfolio of a venture capitalist and the performance measured as the share of IPOs. The positive diversification

effect is stronger for early-stage-focused venture capitalists and is weaker when an investor typically invests in larger syndicates. The authors measure diversification by the share of investments in a given industry based on a standard industry classification scheme.

In sum, the studies described show that reputation is an important factor for venture capitalists. It also creates pressure to push new ventures to go public quickly, especially when the venture capitalists are young and thus have an undeveloped reputation (Gompers, 1996). Gompers (1996) finds that this goes so far that even new ventures are made publicly listed quicker than the portfolio companies of strongly reputable, as measured by "the age of the lead venture capital firm" (p. 139), investors, which leads to high underpricing in the IPOs. Lee and Wahal (2004) also confirm these results for a more recent dataset and even find a general underpricing of venture-capital-backed new ventures in IPOs compared to non-venture-capital-backed companies driven by venture capitalists' desire to gain reputation from IPOs. The authors use the size of the next fund raised by the focal venture capitalists as their dependent variable.

v. Distances and others

The investment behavior of venture capitalists depends on the degree of the internationalization of a new venture. LiPuma and Park (2014) consider a startup's high degree of internationalization a risk-lowering factor, as a diversification effect can be expected. In less internationalized startups, which have more than zero but less than ten percent international revenue, namely, "opportunistic international" (LiPuma and Park, 2014, p. 8) companies, they invest smaller amounts in typically smaller syndicates, and the time between funding rounds is smaller. The authors analyze the venture-capital funding rounds of 334 new ventures.

In a global venture capital investment sample, Liu and Maula (2016) show that venture capitalists mitigate uncertainty when investing in new ventures in foreign countries by partnering with local venture capitalists. The need to partner with local venture capitalists increases with the uncertainty surrounding the new venture and the uncertainty in the foreign country. However, high "country-level uncertainty" (Liu and Maula, 2016, p. 1,415) makes finding a suitable local partner difficult. If venture capitalists become more experienced, the need to partner locally is reduced. Their dependent variable is categorical, as each investor could invest locally or internationally, and, if the latter, then invest alone, with a local partner, or with a foreign partner. They measure "venture-level uncertainty" (p. 1,415) by the development stage with seed being most uncertain and the later stage being least uncertain. They use a standard index for country-level uncertainty that lists countries according to strength of and accordance with

each country's legal system. A venture capitalist's international experience regarding early-stage startups is based on the number of investments in early-stage foreign new ventures; a venture capitalist's international experience in dealing with country-level uncertainty is based on the number of investments in countries above a certain risk level, and international experience in the focal country is measured by the number of investments in that country. For all measures more recent investments are more heavily weighted.

2.4.3 C. Venture capitalist - startup dyad focus

i. Cultural and categorical

A cultural measure based on category labels was used by Pontikes (2012) in her analysis of US software companies. She shows that various audiences judge categorical ambiguity differently. While consumers, who cannot affect the firms and who use categories to help them navigate their consumption behavior, dismiss ambiguous categories, venture capitalists have a preference for ambiguity providing capital faster. She argues that venture capitalists consider categorical ambiguity a sign of the flexibility of the new venture. However, this only holds true for independent venture capitalists, while corporate venture capitalists are more similar to consumers, preferring unambiguous investments. In her analysis, the author distinguishes between being associated with multiple categories that make the firm ambiguous and being associated with an ambiguous label. She assumes that claiming category membership is part of creating an entrepreneurial identity. 457 categories are taken from press release statements from 4,835 software companies in an automated manner. She develops two measures of a label's ambiguity: firstly, "fuzziness" (Pontikes, 2012, p. 92), which is based on the number of overlapping categories, for example, how many companies are also labelled in other categories. Secondly, she calculates "leniency" (p. 93), which indicates how broad or defining a category is, based on the fuzziness and the number of distinct other categories. She also measures whether an organization spans categories by the number of categories it is associated with.

Wry, Lounsbury, and Jennings (2014) analyze whether venture capitalists discount or favor category-spanning startups when selecting their investments. In line with Pontikes (2012), they argue that venture capitalists prefer category-spanning, which is contrary to the discounting found in other audiences' perceptions. In a sample of 58 nanotechnology new ventures, Wry, Lounsbury, and Jennings (2014) demonstrate that, depending on the type of startup, science versus technology, and the way they cross boundaries, venture capitalists show a preference for category-spanning

new ventures, as indicated by receiving another round of financing. Specifically, they look at category-spanning on the patent level, that is, "hybridization of patenting", the top management level, namely, "technology expertise" versus "science expertise" (Wry, Lounsbury, and Jennings, 2014, p. 1,318) and at the collaboration level. For instance, science startups are more likely to get funding if they are boundary-spanning on the patent level, while technology startups are less likely to get funded if they are boundary-spanning on the patent level. The authors measure boundary-spanning on three levels: patents with the number of patents outside of their category, management with the number of top management people with a science or technology background, and collaboration between the patent investors. They validate their results in 13 interviews with venture capitalists to build a deeper understanding of the decision-making criteria and how science and technology are ideally mixed. Overall, the authors lay out very well multiple layers on which boundaries can be spanned, demonstrating great heterogeneity in the startup sphere.

Similarly, in an earlier study, Wry and Lounsbury (2013) analyze the effect of the boundary-spanning patent categories of startups on the success of receiving venture-capital funding. They find support for their hypothesis that the funding probability decreases if multiple patent categories are crossed. They classify patent categories into scientific-related and product-development-related patent categories. However, they do not find support for their hypothesis that the discounting for spanning categories is reduced as long as categories are within either the scientific or product development fields. They show that the categories evolve over time and that the early category-spanning of patent categories that later merge is actually rewarded by venture capitalists. They also show that some categories are more similar to each other than others and that crossing more similar categories has a lower negative effect on receiving another round of venture capital financing. They analyze 62 new ventures in the nano technology industry and use patents classified in about 400 categories. Their dependent variable is whether or not the firm received another round of financing. Their main independent variable, "category dispersion" (Wry and Lounsbury, 2013, p. 123), is based on a modified Herfindahl measure, calculated with the proportions of memberships in a company's patent categories. It ranges from 0 if the company is the only member in one category and approaches 1 in cases of high category dispersion. They measure "category similarity" (p. 124) on the basis of a patent co-citation analysis, assuming that a high number of co-references between patents is a sign of a high degree of category similarity. The authors acknowledge that the sole use of patent categories might not be the most relevant to venture capitalists. Market or product categories might be more relevant and representative of

the business model than pure patent categories.

In sum, the papers neglect heterogeneity among the investors, who can also have varying strategies, preferences, and tastes. They are treated as a relatively homogeneous group. In addition, the analysis of small samples in niche industries highly limits the generalization and transfer of the results to other industries.

ii. Structural – General

The competences and foci of the investors as measured by their structural embeddedness are crucial for a new venture's performance. Ter Wal et al. (2016) show that startups are most successful, measured as having received another funding round, when their syndicate investors either form an open-specialized or closed-diverse network. The researchers use the approach developed by Oh, Chung, and Labianca (2004) to measure the social capital of the syndicate based on the prior syndication networks of the members. "Network closure" (Ter Wal et al., 2016, p. 408) measures the local density of the syndicate network. Openness is based on the prior co-investments of the syndicate members and measures the degree of new information, for example, few prior co-investments lead to a lot of new information. They measure "bridging ties" (Ter Wal et al., 2016, p. 408), that is, those that create connections between previously unconnected venture capitalists, based on the approach developed by Burt (1992). Culturally, Ter Wal et al. (2016) use a 32-industry categorization to calculate a "knowledge similarity" (p. 408) or specialization measure. This measure compares the shares of investments in a certain industry category to that of other investors.

The determinants of initial tie formation also depend on the lifecycle of an organization. Hallen (2008) uses a sample of 92 new ventures in a niche industry to show that early formed ties, in his analysis of professional investors like venture capitalists, are based on personnel connections, for instance, of the founder, whereas more mature organizations form their initial ties based on the company's achievements. He measures a founder's "direct ties" by looking at the investors of firms that the founder had previously founded and measures "indirect ties" (Hallen, 2008, p. 698) by counting the number of times a certain investor has previously invested with another investor that the founder has received funding from. He uses the eigenvector centrality to measure a venture capitalist's status. He also includes measures based on the founder's education, the universities he or she attended, and the number of previously founded companies. Management capabilities are assessed based on previous management roles, the diversity of the founders, and the prior working relationships of the founders. He also collects the number of new products and awards received. The analysis of

only a small niche industry within a three-year time frame from 2000 to 2002, during which the dot-com bubble burst, makes the generalizability of the results weak.

A startup's ties serve different purposes: Shane and Cable (2002) differentiate the effect of "social ties" (p. 372) on the funding decision regarding access to information and social pressure from these ties. They find evidence in a survey among venture capitalists that access to information is more important than social pressure.

Broad network ties can also be a substitute for status: Wuebker, Hampl, and Wüstenhagen (2015) analyze whether personal network ties or status hierarchies are more important to venture capitalists. Their analysis is based on a conjoint analysis such that the authors develop and ask questions regarding certain settings to make social ties and status measurable. For instance, they use three well-known venture capitalists as their high-status investors and an unknown company as the low-status venture capitalist. Their focal variable is thus the "relative strength of strong ties" (Wuebker, Hampl, and Wüstenhagen, 2015, p. 174), which is the difference between personal ties and status. They find that, when facing market uncertainty, investors care more about personal ties, that is, the deal source, than about another venture capitalist's status, namely, the lead investor. When venture capitalists become more experienced, they demonstrate an increased preference for status. However, when venture capitalists become highly experienced, they return to favoring the personal network. The authors only test this relationship for the screening phase.

iii. Structural – Status

Ozmel and Guler (2015) analyze the effect of a new venture's "relative standing" (p. 2,047) in a venture capitalist's portfolio on the chances of an exit through sale or IPO. The authors consider the relative position of a new venture based on the relative size in the portfolio, and they control for, among other factors, the "status" (Ozmel and Guler, 2015, p. 2,048) of the venture capitalist with the eigenvector centrality value. They first model the likelihood of a match between a certain venture capitalist and the focal startup and then control for this likelihood in their analysis. They find support for their hypothesis that a better relative position in the portfolio increases the chances of a successful exit, and this effect is stronger for high-status venture capitalists and large portfolio sizes.

iv. Reputation

New ventures also base their decision on the venture capitalist's reputation. Through a survey, Hsu (2004) finds that new ventures are willing

to accept a 10-14% valuation discount in order to be funded by strongly reputable venture capitalists. This further supports the argument that new ventures not only seek financial capital, but are also interested in the experience a reputable venture capitalist has gained. The author's main measure of reputation is "high industry deal experience" (Hsu, 2004, p. 1,824), which is a binary variable equal to one if the focal venture capitalist has industry segment experience above the median.

While most researchers focus on reputation gained from experience, that is, the number of prior deals, and performance, Drover, Wood, and Fassin (2014) add "ethical reputation" (p. 729) to the discussion. The authors find evidence in a conjoint experiment that venture capitalist's ethical reputation affects a new venture's decision to select a certain investor. A bad reputation regarding unethical behavior might even be more important to an entrepreneur than a venture capitalist's positive performance.

v. Distances and others

Various distance measures are used to explain the matching of venture capitalists and startups: Sorenson and Stuart (2001) analyze the effect of "geographic distance" (p. 1,563) and "industry distance" (p. 1,565) between startups and venture capitalists on whether or not a certain investor invests in a given startup. The industry distance is based on the share of previous investments of the focal investor that were not in the same industry category as the startup. The two authors demonstrate that, due to better information access regarding new deals, investors are more likely to invest in new ventures with less geographical and less industry distance. The effect is weakened if the venture capitalist has strong ties to other investors. In that case, more distant investments become more likely.

Industry fit also affects new ventures' later-stage performance. Lungeanu and Zajac (2016) are concerned with ownership heterogeneity among venture capitalists. They show that, depending on the strategic needs in each step of a new venture's lifecycle, there is a certain type of investor that best meets these needs. They focus on the advisory role of the venture capitalist in addition to his or her monitoring role. The researchers analyze almost 2,000 firms that aimed to go public in a seven-year timeframe and the experience of the venture capitalists invested. They argue that stage fit, industry fit, and investment time-horizon fit all have a positive effect on a company's performance. They confirm their hypotheses with regard to the increased likelihood of successful exit for all three types of fits, including industry fit, and confirm their hypotheses for stage and timehorizon fit with regard to short-term post-IPO share performance. Thus, the fit between the venture capitalist and the invested firm is crucial

to firm performance. The authors measure fit by the count of prior investments in the same type of startup on the three dimensions: development "stage fit," "industry fit," and investment "time-horizon fit" (Lungeanu and Zajac, 2016, p. 939), weighing more recent investments more.

2.5 Summary and outlook

The literature review, based on the empirical variables and lenses used, gives us three main insights. Firstly, the startup and the venture capitalist/syndicate level have been thoroughly researched. Different angles have been used to analyze the effects of status and reputation on, for instance, the performance of startups or venture capitalists, and the effects from cultural and structural measures. However, the number of researchers focusing on the dyad level and applying network-based and related measures has been limited, and only recently has this field received greater attention.

Secondly, we find different foci in the variables used to explain the relationships. For instance, on the startup level, the effect of a venture capitalist's reputation and the ties that startups have to strategic partners and investors are at the core of the empirical work. Not surprisingly, on the venture capitalist/syndicate level, general and status-related structural measures have been the focus. The syndication networks in the venture capital industry are a great setting for analyzing the development and creation of ties and the trade-off between weak and strong ties. Explaining the relationship formation between startups and venture capitalists has been at the heart of the dyad level. Distance measures based on market categories or geographical distance have been used, as well as some structural measures based on existing ties.

Thirdly, a number of blind spots exist that open avenues for further research. While there is agreement on the measurement of status, reputation is still based on a number of different measures. An empirical test of the different measures used might move the discussion on a standard measure forward. In addition, we see that the dyad level still remains partly unexplained. Researchers acknowledge that there exists heterogeneity among startups in terms of their cultural embeddedness and that this heterogeneity is perceived differently by different audiences. However, in these studies, the venture capitalists are primarily treated as a homogeneous group. From our perspective, focusing on the creation of dyads between startups and venture capitalists, one should take into account the differences between venture capitalists' preferences and strategies for the cultural and categorical embeddedness of new ventures. Furthermore, status and reputational differences among venture capitalists is another important factor when analyzing which venture capitalists invest in which startups. Based

on the theoretical framework developed in chapter 3, we aim to shed light on the factors affecting the dyad, considering the heterogeneity among startups and venture capitalists based on a categorical fit in the empirical analysis of this dissertation.

Chapter 3

Theoretical framework and hypotheses

In the previous chapter, we see that only a limited number of studies have used cultural and categorical measures to explain how the dyad between the startup and the venture capitalist is formed. Furthermore, most studies have focused on heterogeneity either on the venture capitalist or the startup side. In this chapter, we develop a theoretical framework and derive testable hypotheses taking into account a new ventures embeddedness and the varying preferences from the investor's perspective.

We start with an introduction of the sociological concepts of legitimacy and distinctiveness, focusing on the new venture's need for and trade-off between legitimacy and distinctiveness. We then describe how an entrepreneurial identity, which is the basis for the determination of legitimacy and distinctiveness, is created from cultural and structural embeddedness. However, we do not make any predictions about whether more or less distinctive startups have greater success in receiving financial capital. In order to answer the first research question, we continue with a consideration of historical experience, portfolio diversification, the status of venture capitalists, and the effects these variables have on the preference for the degree of distinctiveness in a new venture's entrepreneurial identity. In the last section, we develop the hypotheses to answer the second research question,

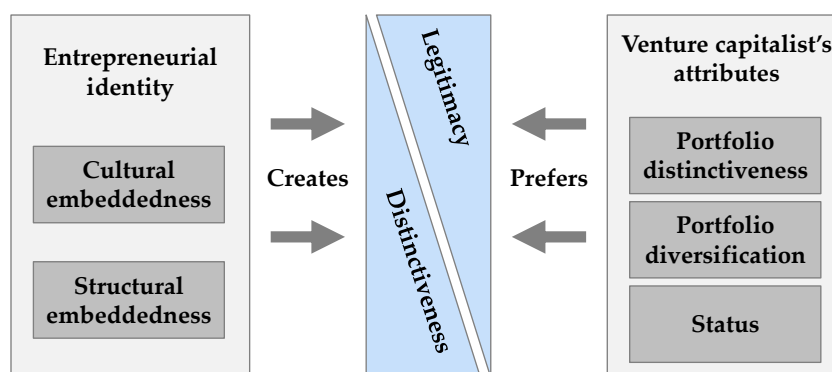


FIGURE 3.1: Theoretical building blocks

discussing the effect of the categorical distance between the focal startup and the venture capitalist's portfolio companies and what mitigates this effect. The building blocks of this chapter and their relationship are displayed in figure 3.1.

3.1 The importance of legitimacy and distinctiveness

A core argument of organizational institutionalism theory centers around an organizations' need for legitimacy (Scott, 2013), which is achieved by being associated with accepted beliefs, norms, or categories. Organizations need to be considered legitimate to gain resources, customers, or media attention. We argue that for organizations, particularly for new ventures, there is also a need to be distinctive and show novelty in order to be successful, for example, by making break-through innovations (Navis and Glynn, 2011). Legitimacy and distinctiveness follow from an organization's entrepreneurial identity, which is created through the organization's cultural and structural embeddedness. For venture capitalists, entrepreneurial identity and the resulting creation of legitimacy and distinctiveness is an integral factor in the funding decision.

3.1.1 Organizational legitimacy

Organizational legitimacy is a field that has long been studied, applied, and referred to, but that often lacks a clear definition. Researchers on legitimacy can be divided into strategic and institutional theorists (Suchman, 1995). The strategic theorists focus on the managerial aspects of legitimacy (e.g., Ashforth and Gibbs, 1990; Dowling and Pfeffer, 1975), which encompass managers' ability to influence and use legitimacy as part of their overall strategic management choices. This includes the use of symbols to convey messages to certain audiences and as a managerial resource (Suchman, 1995). On the other hand, institutional theorists are concerned with pressure on the organization derived from the operative beliefs, norms, and categories around them (e.g., DiMaggio and Powell, 1983; Scott and Meyer, 1983; DiMaggio and Powell, 1991). For them, it is most important that the existence of an organization is legitimized and that this is achieved by the organization's cultural setting, rather than by the organization itself.

At the core of both research traditions lies the notion that organizations gain legitimacy from congruence with existing beliefs, norms, and categories in the environment. Legitimacy is a perception of the audience, and, while an organization can be subjectively considered legitimate by its observers, it may actually divert from existing beliefs.

There is also a need for an organization to demonstrate legitimacy. For instance, an analysis of stock-price development has shown that companies

not covered by relevant industry analysts trade at a discount (Zuckerman, 1999). This means that these companies are not considered to be part of an industry that is covered by analysts and thus lack the required legitimacy. Generally, a desirable organization receives more resources as it becomes more trustworthy and meaningful with increasing legitimacy (Meyer and Rowan, 1977; Suchman, 1995) and legitimacy has been found to improve the resource acquisition and survival rate of companies (Rao, 1994)

In certain contexts, the complexity of legitimization increases. For instance, multi-national enterprises face legitimization complexity from the environment, from within the organization, and in the legitimization process due to their international operations and subunits (Kostova and Zaheer, 1999). Similarly, due to lack of information and often disruptive business models, the legitimization of startups is comparably complex. These companies suffer from the liability of newness (Stinchcombe, 1965), which describes a lack of proof of the business model and thus a higher failure rate. It follows, that legitimacy for new ventures is particularly important (Aldrich and Fiol, 1994).

In summary, legitimacy is beneficial to organizations in acquiring resources, becoming meaningful, and gaining trustworthiness. Legitimacy is gained by being embedded within the cultural setting, sharing norms, beliefs, and categories, and being structurally embedded. DiMaggio and Powell (1983) argue that, due to the need for legitimacy, organizations are very similar, leading to isomorphism among institutions. We argue that this is not generally true for new ventures, as they need to be innovative and potentially disruptive. Legitimacy can be acquired by a number of different means (Fisher, Kotha, and Lahiri, 2016). We postulate that the entrepreneurial identity affected by these means functions as an overarching integrator, on which various audiences can base their perception of legitimacy and distinctiveness.

3.1.2 Distinctiveness through boundary-spanning

While there is a clear need to be legitimate, organizations, particularly new ventures that seek venture capital funding, also have to demonstrate a certain degree of distinctiveness and novelty (Aldrich and Fiol, 1994). This distinctiveness shows how these organizations are different from competitors, for instance, in terms of their corporate culture or business model. It is crucial for the innovation process, for investors like venture capitalists, or for other audiences. The optimal degree of distinctiveness versus legitimacy depends on the audience (Pontikes, 2012).

Boundary-spanning is the sociological concept behind distinctiveness and can be considered an extension of the legitimacy concept. An organization typically has strict boundaries that define it (Aldrich and Herker,

1977) and that are the signals on which an audience bases its (unconscious) decision to assign legitimacy to an organization; therefore, legitimacy and boundaries play an informational role. Deviation is often called 'boundary-spanning' as it crosses existing, demarcated boundaries (Fennell and Alexander, 1987).

An important aspect is that boundary-spanning is required for the innovation process (Tushman, 1977), for instance, when boundaries are of a technological nature. For companies that focus on a certain technology and that continue to innovate incrementally within their technology boundaries (Rosenkopf and Nerkar, 2001), boundary-spanning is the process of combining different technologies. This may be achieved via the establishment of alliances or other forms of relationships, especially with technology companies (Stuart and Podolny, 1996; Nagarajan and Mitchell, 1998). While in-boundary technological developments are mainly small, incremental improvements (Rosenkopf and Nerkar, 2001), break-through innovation can only be achieved by spanning existing boundaries. Rosenkopf and Nerkar (2001) argue and demonstrate that, in the CD/DVD domain, in addition to technological boundary-spanning, organizational boundary-spanning can also lead to radical innovation. If the knowledge base of multiple organizations is combined, new technical re-combinations can be achieved. In order for new ventures to be innovative and develop new business models that can disrupt entire industries, they need to cross boundaries.

There is a tradeoff between the legitimacy that is gained from being associated with a clear and distinct category and the distinctiveness that is gained through crossing existing borders. Legitimacy is generally important for all firms, but it is especially significant for new ventures that have not yet gained legitimacy by sustained survival (Navis and Glynn, 2011). Venture capitalists are especially sensitive to how they can justify their investments to their limited partners based on the attributes of the startup. Contrary to this, Pontikes (2012) argues that venture capitalists, as opposed to other audiences such as consumers, prefer distinctiveness over legitimacy.

In sum, there is a need for new ventures to be simultaneously legitimate and distinctive, similar to individuals as argued in the optimal distinctiveness literature (compare Brewer, 1991; Zuckerman, 2015). Thus, there are differences in the degree of distinctiveness among new ventures. This heterogeneous group of new ventures also faces a heterogeneous group of venture capitalists, of which some prefer more legitimized and some more distinctive new ventures. We explore this further in section 3.3.

3.2 The creation of entrepreneurial identity

Entrepreneurial identity is important to create "legitimate distinctiveness" (Navis and Glynn, 2011, p.482) or in other words a coherent story on an organization. The entrepreneurial identity is created by the cultural and structural embeddedness¹, which we will outline in the following.

3.2.1 Introduction to entrepreneurial identity

Entrepreneurial identity has received little attention in the current state of research, with only a few exceptions (Navis and Glynn, 2011; Lounsbury and Glynn, 2001; Martens, Jennings, and Jennings, 2007; Tripsas, 2009). It is generally accepted that, in essence, an organization's identity is claiming certain attributes or categories, for instance, as a reference to its business model (Glynn, 2000; Ashforth and Mael, 1989). For instance, Navis and Glynn (2011, p. 480) describe entrepreneurial identity as "the constellation of claims around the founders, organization, and market opportunity of an entrepreneurial entity that gives meaning to questions of 'who we are' and 'what we do'." From our perspective, entrepreneurial identity also includes structural embeddedness, that is, the new venture's network, and the legitimacy and distinctiveness it gains from that network. Hence, entrepreneurial identity is an integral part of a new venture and is shaped by introducing distinctiveness and novelty on the one hand and conforming to shared beliefs on the other. The identity of a new venture is especially important as these young companies often have little more than an idea, business plan, or prototype, which increases the importance of the entrepreneurial identity for potential investors, employees, or customers (Elsbach and Kramer, 2003; Navis and Glynn, 2011).

Entrepreneurial identity follows from the combination of an organization's legitimating and boundary-spanning elements that formed the basis for the categorization. For a new venture to be successful, it needs a "legitimate distinctiveness" (Navis and Glynn, 2011, p. 482): A firm cannot randomly combine distinct elements to create a legitimate distinctive identity, but it needs to have a coherent story, identity, or business model in order to ensure its legitimate distinctiveness (Navis and Glynn, 2011). In a coherent story, all of the legitimizing and boundary-crossing elements on the personal, organizational, and market levels need to fit together. This argumentation is in line with other authors who indicate that any story around new ventures' business models need to be coherent in order for them to achieve the desired signaling effect (e.g., Magretta, 2002; Perkmann and

¹We follow the definition by Goldberg et al. (2016) of embeddedness as "one's degree of anchoring in a social context" (p. 1,216), which can exist in a cultural and a structural dimension.

Spicer, 2010) and that this makes it easier for new ventures to access funding, employees, and other resources (Lounsbury and Glynn, 2001; Martens, Jennings, and Jennings, 2007; Zott and Huy, 2007). For example, Zott and Huy (2007) show that new ventures that use symbolic management to create a legitimate identity are more successful at securing required resources.

While entrepreneurial identity is also a type of categorization, it is more than just an association of cultural categories. Other ways of creating entrepreneurial identity may be possible: the legitimacy that is gained from being funded by a high-status venture capitalist may affect entrepreneurial identity, for instance. Goldberg et al. (2016) argue and empirically demonstrate that legitimacy and distinctiveness can stem from cultural and structural embeddedness. Entrepreneurial identity integrates the cultural and structural embeddedness of an organization into a holistic concept. Categorization places organizations in categories without considering the whole picture, while, on the other hand, entrepreneurial identity binds all elements together into an ideally coherent story.

3.2.2 Cultural and categorical embeddedness

As outlined above, boundary-spanning and gaining legitimacy can be achieved through several means and layers. One major aspect is cultural embeddedness, which is also at the heart of this work. Most importantly, cultural embeddedness encompasses categorical embeddedness, that is, what categories a new venture is associated with. This so-called 'social categorization' is the process of dividing organizations into groups, by which each organization either forms part of a group (in-group) or does not (out-group), based on the organization's attributes (Hogg and Terry, 2000). Organizations are compared to group prototypes in order to determine whether they are in-groups. These group prototypes incorporate all of the attributes that are specific to a certain group. An organization's position is also dependent on the current environment and may evolve over time, which is why it is important to consider group membership at a certain point in time (Hogg and Terry, 2000).

Two parties can effectuate categorization: an organization itself (self-categorization) or others (Jenkins, 2000). While the management of an organization can try to influence its categorization, it is the perception of an audience that matters in the creation of legitimacy.

Categorization also has very practical, every-day applications as shown in consumer-psychology research. Consumers use categorization to navigate their consumption behavior (Pontikes, 2012; Loken, Barsalou, and Joiner, 2008): When they see a product, they immediately categorize it into one of their known categories and, if it is difficult to categorize, it is discounted. This has practical implications on the marketing, for example, the

positioning and advertising of products. In consumer-psychology research, categorization has been analyzed in a number of fields, such as brand categories (e.g., Barone, 2005), product categories (e.g., Meyers-Levy and Tybout, 1989), and cultural categories (Briley and Wyer, 2002).²

Placing organizations in categories helps audiences, as it serves as a summary of a firm's attributes, which is especially helpful in cases of uncertainty or ambiguity (Ashforth and Humphrey, 1997). Boundary-spanning concerns all types of boundaries, and a firm can easily have boundaries in many dimensions. Categories summarize part of these dimensions and are easier to treat for consumers, venture capitalists, or other audiences.

Many researchers such as Hannan, Goldberg, and Kovács (2016) as well as other sociologists (e.g., Peterson, 1992; Johnston and Baumann, 2007) have analyzed cultural embeddedness only in cultural settings that are understood in a narrow sense, for instance, in arts such as film-making or cuisine. In a broader sense, culture can be considered the shared beliefs within a society or group of people, and legitimacy is gained by an organization's or object's cultural embeddedness. This includes the perception of consumers, investors, and employees, all of whom form legitimacy.

Startups, for example, can distinguish themselves, that is, be distinctive, by boundary-spanning through defiance of established categories and combining distinct cultural attributes to create new ones (Hannan, Goldberg, and Kovács, 2016). If a single startup identifies itself with the categories of *biotechnology*, *telecommunication* and *software*, it is distinctive, as it combines multiple bounded categories to create a new category. Distinctiveness thus entails creating a new category, and, in our example, it may be a signal for a new business model that typically creates more risk, as it has not been proven. However, this higher risk is also rewarded by a potentially higher return.

Cultural sociologists consider boundary-spanning to be a consumption of multiple categories and view it as consumption behavior rather than as an object attribute. For instance, Hannan, Goldberg, and Kovács (2016) analyze movie goers' and restaurant visitors' consumption preferences. They introduce variety as a measure of the extent of consumption of different cultural types. A movie or restaurant itself cannot have a measure of variety: only consumers, investors, and other types of audiences that consume multiple objects can cross boundaries in this sense.

3.2.3 Structural embeddedness

Structural embeddedness is defined by the network to competitors, investors, or other partners, like research institutions or suppliers, that a new

²For a more in-depth review, please see Loken, Barsalou, and Joiner (2008)

venture has connections to. From this network, social capital, that is, access to resources within the network, is gained by the new venture (Nahapiet and Ghoshal, 1998). The resources from this network can include access to information, technology, and financial capital. The current network partners also affect the future network partners, for instance, an organization is more likely to form a direct tie to indirectly connected organizations than to unconnected partners (Gulati, 1995), and thus impact the identity and development of an organization. A recent meta analysis of 61 samples has shown a significant positive effect of social capital on small firms' performance (Stam, Arzlanian, and Elfring, 2014). One important discussion on structural embeddedness revolves around the question of whether greater access to broader information from less dense networks through the closure of structural holes (Burt, 2004), that is, a focal actors connections to many unconnected partners, is more beneficial to an organization than strong ties in dense networks (Burt, 1997), that is, a focal actor's connections to others that share many ties (Ter Wal et al., 2016).

According to Burt, Kilduff, and Tasselli (2013) as well as other authors, the term 'structural embeddedness' describes specifically the 'strong tie embeddedness', that is, an organization with a strong structural embeddedness has a highly dense network in which ties are well connected with each other and few structural holes exist. In a recent review, the authors show that, in the social-psychology context, structural embeddedness, that is, constraints in dense networks, reduce the advantages of social networks, especially regarding access to new information and innovativeness. However, they agree that structural embeddedness has a positive effect on reputation and the trust among network members. In sum, whether strong or weak ties are better for an organization depends on the setting and state, which either requires more diversity and information or deep insight (compare Burt, 1997; Ter Wal et al., 2016).

One important aspect in the discussion of structural embeddedness is the legitimacy gained from it such that it becomes part of a new venture's entrepreneurial identity. When other types of quality assessment do not exist, structural connections serve as quality signals (Podolny, 1993; Ozmel, Reuer, and Gulati, 2013). For instance, venture capitalists use the partner network of a new venture as a quality signal (Stuart, Hoang, and Hybels, 1999). Similarly, high-status venture capitalists can themselves be a positive quality sign for a new venture. For new ventures, these legitimizing signals are especially important (Rao, 1994). For instance, Higgins and Gulati (2003) show that when other quality signals do not exist, the personal ties of the new venture's management to other reputable organizations affect the IPO success.

For the creation of entrepreneurial identity, the legitimacy resulting

from structural embeddedness is most relevant. Aldrich and Fiol (1994) show that network partners can help overcome lack of legitimacy for new ventures. Specifically, these network ties can help in understanding the business and, secondly, convince other network partners to accept the company as an appropriate partner. Elfring and Hulsink (2003) develop a framework based on three case studies that showed that, especially for highly innovative new ventures, a combination of strong and weak ties can enhance legitimacy. The networks of young companies develop over time (Maurer and Ebers, 2006), depending on the development stage, such that they have a greater need for strong ties (dense connections), which evolves into a sparser network in later stages (Hite and Hesterly, 2001).

3.3 Venture capitalists and entrepreneurial identity

Based on the description of entrepreneurial identity and the factors influencing the legitimacy and distinctiveness of new ventures, we develop a number of hypotheses with regard to which venture capitalists invest in what type of startups. The mechanics at work in the venture capitalist - startup relationship along with the empirical approaches used, as described in chapter 2, provide the basis for our argumentation. We add some specific theory about the venture capitalist's preference for distinctiveness, portfolio diversification, and the impact of status to support our argumentation.

3.3.1 Distinctiveness and novelty in venture capitalists' portfolios

A new venture's affiliation with certain institutionalized categories in order to gain legitimacy varies according to the degree of distinctiveness. Above, we defined the startup-related attribute of distinctiveness as a measure of the degree of boundary-spanning in new ventures. Moreover, venture capitalists have a choice regarding investment in a certain degree of distinctiveness among their portfolio companies.

Categorical ambiguity – for example, when a new venture that is difficult to categorize, either because it falls into multiple categories or because the distinction between several categories is unclear – also creates distinctiveness. Pontikes (2012) argues that, depending on the audience, categorical ambiguity can either be a desired or an unwanted feature. She distinguishes between consumers, who prefer category association, and venture capitalists, who prefer categorical ambiguity due to their desire for distinctiveness. Consumers and other audiences who prefer sharply defined category association use categories to navigate and reach their goals, for instance, by looking for a certain type of company that can fulfill their need

for a specific service or product. As this type of audience cannot influence the companies, in their search, they discount companies with ambiguous business models versus companies without categorical ambiguity.

Other audiences, such as venture capitalists, prefer categorical ambiguity, as it is a sign of flexibility: for instance, the ability of a firm to adapt to changes in its environment such as customer demand (Pontikes, 2012). Once they have invested, venture capitalists have a strong influence on the company and can thus shape the company upon analysis. In addition, they spend more time analyzing and doing research on a company so that, although categorical ambiguity may exist at first sight, a closer look reveals that the venture capitalist might be able to assign a clear category to the firm. Once such venture capitalists have control or influence over a firm, they can shape the company in order to reduce ambiguity for other audiences as well. A recent empirical analysis of venture capital funding in the nanotechnology sector has also found empirical evidence that venture capitalists are more likely to fund new ventures with a hybrid identity of science and technology compared to firms with a pure science identity (Wry, Lounsbury, and Jennings, 2014). The effects were especially strong for certain categorical combinations together with certain header categories.

This raises the question of why there are still new ventures with low distinctiveness that are funded by venture capitalists if distinctiveness is generally desirable. Because, with higher distinctiveness, the risk of failure also increases, there are some venture capitalists that can afford to take such a risk while others cannot. Being distinctive is a risk factor that Zuckerman (2015) calls "valuation risk". If a new venture is highly distinctive, defying many established norms, there is a high risk that the company will not be accepted, for instance, that it will not find enough customers, will not reach its strategic or financial goals, or will even end up insolvent. Thus, entrepreneurial identity's perceptual effect, that is, how the new venture is perceived by various audiences, also affects the development of the business itself. Startups that cross boundaries or are generally difficult to classify may be devaluated by consumers. The business model of a startup with a strong boundary-crossing identity is more likely to be of greater novelty than that of more embedded startups. It can also be a sign of flexibility if a startup is highly non-conforming (Pontikes, 2012). For a venture capitalist, the degree of distinctiveness thus signals the degree of novelty, flexibility, and risk of a new venture. However, there might be a reward: The valuation or business risk associated with being distinctive might lead to increased performance (Zuckerman, 1999; Phillips, Turco, and Zuckerman, 2013; Sgourev and Althuizen, 2014).

We have defined a startup's distinctiveness as the opposite of legitimacy

in a continuous space. A startup is thus placed somewhere in the continuum between high legitimacy and high distinctiveness depending on a startup's cultural embeddedness. A highly distinctive startup is less embedded and more boundary-spanning than one with low distinctiveness. When venture capitalists face their investment decision, they follow a certain strategy with regard to their portfolio. This can be in order to invest in more established categories and business models, or to take the risk of investing in newer, more distinctive startups. Venture capitalists that have gained experience in selecting, monitoring, and possibly coaching distinctive startups are expected to continue to do so. Not only have venture capitalists typically communicated their investment strategy at the origination of the current fund, but they also reduce the valuation risk when they continue to invest in similarly distinctive new ventures. Consequently, the first hypothesis is as follows:

H1: Venture capitalists with experience in investing in distinctive startups have a greater probability of investing in distinctive startups in the future.

3.3.2 Diversification of venture capitalists' portfolios

Variety describes the consumption of multiple cultural categories by consuming several objects, although each object can exist within established cultural boundaries (Hannan, Goldberg, and Kovács, 2016). Transferring this to organizations, for instance, in the venture capital context, this is a measure of portfolio diversification. The greater the difference between the startup's identities in a venture capitalist's portfolio, the more diversified the portfolio is. A venture capitalist would thus have a high degree of diversification if he or she invested in multiple startups that have highly different attributes, such as investing in a *biotechnology*, a *telecommunications* and a *software* startup.

In addition to investment-specific risk, venture capitalists also face portfolio risk (Norton and Tenenbaum, 1993). One strategy by which to reduce individual investment risk in the portfolio is diversification (Sharpe, 1964). In an investment portfolio with many assets that are not fully correlated, each asset reduces idiosyncratic or firm-specific risk. With an increasing number of assets, the only relevant risk that remains is market risk, which is the risk of the volatility of the whole market and can be caused by an economic downturn, rather than by the failure of individual market participants. The main assumption of the Capital Asset Pricing Model is that investors are only rewarded with higher returns for systematic risk, as any investor could diversify away unsystematic risk (Sharpe, 1964). In venture capital, there is high idiosyncratic risk especially in the early investment

phases (Ruhnka and Young, 1991). Venture capitalists could thus diversify through investments in multiple assets within varying industries and regions with low correlation. Another way of diversifying is by investing in different stages, such as the seed stage, early stage, and late stage of venture capital (Norton and Tenenbaum, 1993). However, venture capitalists are typically asset managers who allocate their limited partners' financial resources into new ventures. These limited partners, for instance, pension funds, can diversify their portfolios themselves. Specialized venture capital firms with a focus on industry and region allow limited partners to give capital to the venture capitalists that best match their risk, return, and portfolio-constraints profile. The limited partners typically have a number of other asset classes, such as bonds, stocks, or later-stage private equity, which need to be taken into account when allocating capital to venture capital and selecting venture capitalists. Venture capital firms have a desire to diversify their portfolios in order to reduce the risk of bankruptcy: If the cost of the risk of bankruptcy is too high, it negatively affects the performance of the firm.³

Strategists argue that venture capital firms benefit from specialization (Chan, 1983; Bygrave, 1987). As described above, venture capitalists spend a lot of time with new ventures, advising and helping them to professionalize (Hellmann and Puri, 2002). In order to be as beneficial as possible to new ventures, venture capitalist's need domain-specific knowledge. Especially for young ventures in their early stages, the specific product, market, or technology knowledge is crucial in making the new ventures successful. A survey among venture capitalists with 98 responses shows that the need to specialize is greater than the need to diversify (Norton and Tenenbaum, 1993) and Gompers, Kovner, and Lerner (2009) support findings that specialized firms outperform diverse firms. Diverse firms are more successful if they are comprised of specialists rather than generalists. More recently, Matusik and Fitza (2012) analyzed the performance of venture capital firms over a period of 40 years, depending on their degree of diversification. They found a U-shaped relationship between firm performance and the degree of diversification moderated by co-investments and the investment stage. Lower levels of co-investments and earlier investment stages increase the U-shaped relationship. Based on these findings, the level of diversification has been shown to be an important strategic choice for venture capital firms: Either high specialization or high diversification are the most beneficial to fund performance. Typically, small portfolios benefit from a low

³The argumentation is similar to the arguments about optimal capital structure: Increasing financial leverage increases the risk of bankruptcy for an organization. As long as the tax savings from higher leverage are greater than the cost of bankruptcy, it is advisable to take more risk (Kraus and Litzenberger, 1973; Gruber and Warner, 1977).

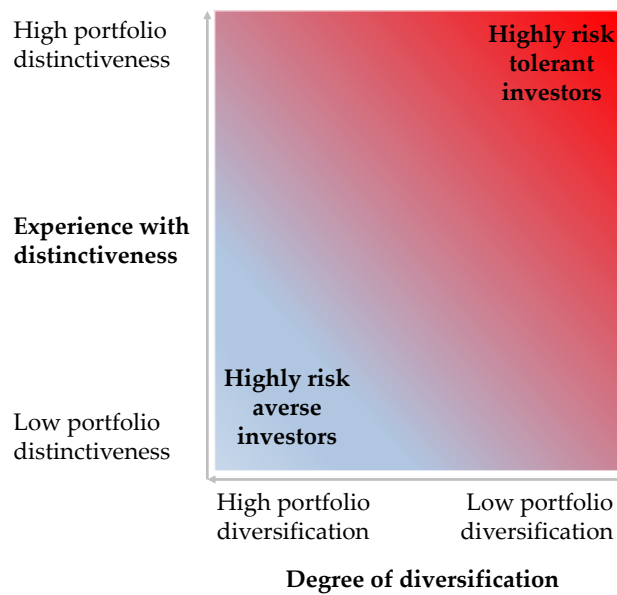


FIGURE 3.2: Summary portfolio attributes and risk tolerance

correlation between firms while large portfolios are typically more specialized (Fulghieri and Sevilir, 2009).

In another context it has been shown that movie goers and restaurant visitors who have a clear preference for objects in established categories display an increasing preference for consuming a great variety of categories (Hannan, Goldberg, and Kovács, 2016) and are considered open minded in doing so (Ollivier, 2008). Similarly, we expect that venture capitalists prefer highly diversified investment portfolios when investing in startups with low distinctiveness. This naturally makes sense: As venture capitalists with a preference for high diversification in their portfolios attempt to reduce risk by diversifying into multiple categories, we assume that they are risk-averse. If the legitimacy of an object reduces risk and distinctiveness refers to novel and unproven new ventures, such risk-averse venture capitalists will have a preference for a low degree of distinctiveness.

H2: Venture capitalists with diversified portfolios have a lower probability of investing in distinctive startups in the future.

3.3.3 The combined effect of average distinctiveness and diversification

We have now described and developed two venture capitalists' attributes based on their current portfolio. Firstly, the experience that is gained by investing in distinctive new ventures measured by the average distinctiveness of the portfolio companies. Secondly, there is portfolio

diversification, which is based on the differences in portfolio companies' identities. These two dimensions describing the venture capitalist's portfolio can be orthogonal to each other, as is displayed in figure 3.2. Based on its current portfolio, a venture capitalist can thus be anywhere in the matrix having either great or little experience in investing in distinctive new ventures, that is, high or low portfolio distinctiveness, and having either a highly or lowly diversified portfolio.

A highly diversified venture capital portfolio, which encompasses new ventures with very different identities, reduces overall portfolio risk, as described above, and is a sign of risk-averse venture capitalists. Consequently, investors with low portfolio diversification and high portfolio distinctiveness are most risk-tolerant, while portfolios with high diversification and low distinctiveness indicate risk-averse venture capitalists.

Our framework results in a four-type categorization based on an investor's portfolio. As diversification and distinctiveness do not simply add up (Hannan, Goldberg, and Kovács, 2016), and if hypothesis H1 and hypothesis H2 hold true, we would expect investors with a preference for high diversification and low distinctiveness to have the strongest preference for low distinctiveness.

3.3.4 Status benefits from structural embeddedness

Structural embeddedness describes the position of an object in a (typically social) network – the structure – and its advantages and disadvantages depending on the degree of embeddedness (e.g., Burt, 1997; Podolny and Baron, 1997). In addition to clearly demonstrated benefits from the flow of information from these networks and the closure of structural holes in syndicated venture capital investments (among others Podolny, 2001; Ter Wal et al., 2016), the structural position of a venture capitalist in the co-syndication network is also a signal of status and thus of the quality of a venture capitalist.

Generally, an organization's identity, including its status, depends on the network around it with respect to status and power differences (Hogg and Terry, 2000). Podolny (2001) proposes that, within a setting of "alter-centric uncertainty" (p. 37), a situation in which the quality of an actor is unclear, and specifically in the venture capitalist - startup relationship, networks can be considered prisms that convey information about the status of a venture capitalist. Furthermore, status has been shown to be a sign of relative quality compared to other producers, such as venture capitalists, in the market (Podolny, 1993; Podolny, Stuart, and Hannan, 1996).

Venture capitalists have to justify their investments to their limited partners, which can be achieved if the new venture's identity displays a high

degree of legitimacy. If the venture fails, the venture capitalist can still argue that it was, at least in parts, a proven or accepted concept and thus not a misallocation caused by insufficient due diligence. On the other hand, venture capitalists seek novel businesses for it to be possible to achieve abnormal returns on their investments.

The literature on the effect of status on preferences for either legitimizing conformity or boundary-spanning has argued for and empirically found opposing effects. Many researchers have found an inverted U-shape legitimacy preference, which entails middle-status actors' preference for legitimacy while high- and low-status actors do not share this preference (among others Hollander, 1958; Blau, 1960; Hollander, 1960). This argumentation is based on the assumption that an actor identifies with a role and a certain security that can be gained from it (Phillips and Zuckerman, 2001). High-status actors thus feel sufficiently secure in their role of having a boundary-spanning preference (Hollander, 1958; Hollander, 1960). Their risk is limited by the status and identity that they have already gained. For low-status actors, there is a 'nothing to lose' situation. As low-status actors are neglected no matter what their actions are, they do not need to conform to any expectations (Phillips and Zuckerman, 2001). In other words, a low-status actor can take the risk of choosing a boundary-spanning startup, because even if they fail, they do not lose any status. In sum, both high- and low-status actors have nothing to lose. In contrast, middle-status actors do not have the security of their established status identity, but are also not entirely unconsidered. Phillips and Zuckerman (2001) found that middle-status actors in the market for professional services and investment advice have the highest pressure to conform to expected standards, resulting in an inverted U-shape curve between status and legitimacy. In the venture capital domain, however, there is a matching or sorting component (Sorensen, 2007). In the relationship between venture capitalists and startups, not only do venture capitalists have a choice in which startups to invest; startups also select their investors to best match their needs, in addition to the required financial capital. This distinguishes the venture capitalist - startup relationship from most settings within the cultural domain, such as the relationship that a movie producer has with his consumers, in which the producer generally does not select the consumer but is happy for every paying customer. The best startups choose their investors and naturally select those with the best network and highest status as the high-status investors can increase the new ventures' legitimacy creating a positive quality signal (Podolny, 1993; Rao, 1994). This is especially relevant for those with low legitimacy and high distinctiveness making it hard for low-status venture capitalists to invest in this type of startups. Jensen and Roy (2008) show that organizations reduce the set of potential auditing firms based on the relative

status ranking of these firms and only then consider the actual experience. Analogously, we postulate that new ventures choose their venture capitalist based on the venture capitalist's status. Low-status investors are therefore not able to invest in highly distinctive startups for two reasons. Firstly, the most promising startups are often the ones with a high degree of distinctiveness due to a greater chance of radical innovation. Secondly, startups with the highest distinctiveness need legitimacy from high-status venture capitalists. Both factors make highly distinctive startups choose high- and not low-status venture capitalists. The inverted U-shape curve will thus not hold in the venture capital domain; rather, a linear relationship is expected.

We can summarize that, in some settings, the relationship between status and legitimacy follows an inverted U-shape (Phillips and Zuckerman, 2001). However, in the venture capital domain, we expect low-status investors to not be able to demonstrate their preferences. As startups also chose their investors, they will always pick the venture capitalist with the highest status, with the expectation that they will benefit from the network, knowledge, and legitimacy of a high-status venture capitalist. Low-status venture capitalists thus typically do not get the chance to invest in boundary-spanning, innovative startups. High-status investors can afford to make investments in high-risk startups with high distinctiveness, whereas middle-status investors are afraid to lose their status in case of failure. From this, the following hypothesis follows:

H3: The higher the status of a venture capitalist is, the greater the probability of investing in distinctive startups in the future.

3.4 Categorical distance and venture capitalists' attributes

In the previous section we developed hypotheses on the likelihood of the investments of venture capitalists based on the distinctiveness created by the entrepreneurial identity. Another broadly accepted argument concerning the formation of investment ties is that investments are more likely the shorter the distance between a new venture and a venture capitalist is (Sorenson and Stuart, 2001). As the second part of our analysis, we now develop hypotheses regarding the effect of categorical distance on the matching of startups and venture capitalists and, even more importantly, how the venture capitalist's experience of investing in distinctive new ventures, degree of portfolio diversification, and status affect sensitivity to categorical distance. The important link to our previous reasoning of a new venture's distinctiveness is how we measure categorical distance and distinctiveness:

We test our hypotheses by considering market categories, taking into account the distances between these categories. The methodological details are described in chapter 4, but it is important to understand how the concepts of categorical distance and distinctiveness are connected. From a conceptual perspective, one can see another similarity between the two measures: A startup's distinctiveness is a measure of the general novelty of the startup compared to all other startups, while categorical distance is a measure of the startup's novelty compared to the venture capitalist's portfolio companies.

We start this section by reiterating the main line of argumentation made by Sorenson and Stuart (2001) about why greater distance reduces the likelihood of investments. This is followed by the development of our hypotheses regarding the effect of a venture capitalist's attributes on sensitivities to categorical distance. We keep the description of the venture capitalist's attributes relatively short, as they have been discussed extensively in the previous section.

3.4.1 General effect of distance

Proximity is beneficial in the pre- and post-investment phases. In the pre-investment phase, venture capitalists first need to identify potential investment targets. This includes the identification of new ventures that match the desired attributes, for instance, with regard to distinctiveness. Secondly, these new ventures need to be seeking additional capital from venture capitalists at that moment to be considered a potential investment target. The more information the venture capitalist and the startup have, the better the chances of an investment. However, the flow of information is greater between closer actors than between more distant ones, which leads to a greater number of and more information about new ventures seeking capital in proximity than about distant new ventures. In addition, as the venture capital context is characterized by high uncertainty because the new ventures do not have proven business models and information is often sparse (Gompers and Lerner, 2001), venture capitalists tend to have more trust in closer partners. Secondly, after target identification, the pre-investment phase includes the evaluation of new ventures that are potential investment candidates. Due to informational asymmetries that lead to adverse selection and moral hazard problems (Amit, Brander, and Zott, 1998), venture capitalists need partners and information from partners that they can trust. They are more likely to find these in proximity rather than at a distance. Furthermore, the new ventures can also choose their venture capitalist. In line with the satisficing principle, which assumes that actors do not optimize until they find the optimal solution, but instead search until

they find a sufficient solution (Simon, 1955), a startup searches for a venture capitalist that fulfills the needs the organization currently has. They do not necessarily search among all potential venture capitalists. Rather, doing a cost-benefit analysis, they start with the venture capitalists in proximity to reduce the information and travel costs of searching for more geographically distant or otherwise distant venture capitalists. Consequently, venture capitalists gain easier informational access and receive information about new ventures seeking capital in proximity; new ventures would similarly start by searching for venture capitalists in their own neighborhood.

Knowledge spillovers happen, at least for certain industries, in geographically local cluster (Zucker, Darby, and Armstrong, 1994). These clusters are particularly important for knowledge spillovers to increase innovation (Audretsch and Feldman, 2004) and are most relevant to institutional ties (Bell and Zaheer, 2007). Singh (2005) also finds positive effects on the flow of information when geographical distances are short; however, when prior ties between actors exist, geographical distance becomes less important. In collaboration networks, social ties can even be more relevant to the decision of whether or not to collaborate than geographical distance (Autant-Bernard et al., 2007). The need for geographical proximity is also reduced when organizations share a similar background, as is shown for research institutions (Ponds, van Oort, and Frenken, 2007). In other words, there is higher "institutional proximity" (Autant-Bernard et al., 2007, p. 496) or smaller industry and cultural distance between these organizations.

In the post-investment phase, venture capitalists use close monitoring to overcome informational asymmetries (Kaplan and Strömberg, 2001). More than 80 hours per year on average are spent with each portfolio company (Gorman and Sahlman, 1989), in which discussions with the management regarding the achievement of milestones, the strategic direction of the company, and other topics take place. These monitoring and advising activities are the reasons for the effects venture capital has on startups, like greater innovation, stronger internationalization, and overall better performance (e.g., Sapienza, Manigart, and Vermeir, 1996; Lerner, 1995; Sapienza, 1992; Gorman and Sahlman, 1989). Greater distance reduces these benefits. For instance, if it takes more time to travel or if personal relationships are weak, because in-person meetings are replaced by video or telephone conferences, less time can be spent on value-adding activities, and meetings are less effective. The same is true for more industry-distant new ventures. If the venture capitalist does not have experience in the same industry, he or she is less familiar with it and cannot give the same quality advice, has a weaker network in that industry, for example, regarding collaborators, or at least needs more time to build up the required knowledge and understanding.

In sum, venture capitalists at a shorter distance gain access to more investment opportunities, are better at evaluating them, and can give better advice and spend more time with the new venture.

So far, we have used the general notion of distance, which we now define. While this is not intended to be a discussion about methodologies, we need to present a number of distance measures in order to clarify the different meanings between them. Distance can describe the geographical distance between the startup's headquarters and the venture capitalist's main or satellite office (Sorenson and Stuart, 2001). Generally, geographical proximity reduces the cost of monitoring and facilitates knowledge-sharing. Knowledge spillovers, in turn, have a strong impact on an organization's innovativeness, as long as the organization has the "absorptive capacity" (Cohen and Levinthal, 1990, p. 128) to do so. Industry distance is another distance measure and can be considered a familiarity measure, that is, how familiar the venture capitalist is with that industry segment as measured by the number of prior investments in that industry segment (Sorenson and Stuart, 2001). Furthermore, distance can represent cultural distance, which is a measure of how similar or dissimilar two cultures are (typically a comparison between two country's cultures) and has been broadly used in research as a factor explaining various relationships (Shenkar, 2001). Widely spread is the use of the four-dimensional model developed by Hofstede (1980) and an index based on these dimension created by Kogut and Singh (1988). Hofstede (1980) conducted a survey of IBM employees and found four independent factors that explained much of the response variation: 'power distance', 'uncertainty avoidance', 'individualism' versus 'collectivism', and 'masculinity' versus 'femininity'. The index is typically used in studies about foreign-country-entry behavior (Drogendijk and Slangen, 2006). An alternative cultural distance measure was introduced by Schwartz (1994), who criticized Hofstede's approach and dimensions. Schwartz also used a survey, but not from one individual company, and derived seven cultural dimensions. Drogendijk and Slangen (2006) have shown that both measures perform well in explaining international entry modes.

Cultural distance is also relevant because it impacts the way managers work together (Lincoln, Hanada, and Olson, 1981) and reduces the formation of ties between managers with great cultural distance (Manev and Stevenson, 2001). This can affect the collaboration behavior between the portfolio companies of a single venture capitalist or between a venture capitalist and a new venture. On the other hand, in cross-boarder acquisitions, cultural diversity can even enhance performance through learning from a new set of conventions for doing things (Ghoshal, 1987; Morosini, Shane, and Singh, 1998).

A concept related to cultural distance is institutional distance, which is also often used in international business research. It is developed as a broader advancement of the cultural distance concept, which is based on cultural dimensions developed by Hofstede (1980), and measures two countries' differences in terms of their institutional environments (Gaur and Lu, 2007). The institutional distance developed out of institutional theory, which assumes that organizations seek legitimacy by conforming to existing beliefs and cultural norms (Meyer and Rowan, 1977; DiMaggio and Powell, 1983) in their home countries. The distance measure has thus emerged as a measure of institutional difference based on these regulatory, normative, and cognitive elements and was introduced by Kostova and Zaheer (1999) (see also Xu and Shenkar, 2002). The way it emerged was complementary to cultural distance (van Tulder, 2010), but it has now become a broader concept. In a review of current research, Bae and Salomon (2010) have found five dimensions that institutional distances encompass, but most researchers have only analyzed a selection of these five: the cultural, the regulatory, the political, the economic, and the cognitive dimension. International firms try to assimilate to local peculiarities in a foreign country to create legitimacy and reduce the negative effects of being foreign (Salomon and Wu, 2012). The research of Salomon and Wu (2012) is particularly interesting, as they used four institutional distance dimensions in their analysis. The authors measure cultural distance with the index developed by Kogut and Singh (1988), economic distance by a ratio based on market capitalization, national gross domestic product and bank credits, regulatory distance by a survey-based database on banking regulation, and political distance by a standard index based on political volatility, developed by the World Bank.

This macro-level view, that is, considering cultural or institutional country differences, is not very useful in the venture capital domain as long as we are only analyzing a single country and is especially not useful for explaining which venture capitalists invest in which startups. However, the underlying notion of comparing cultural and some institutional elements is relevant, but we are interested in a micro-level consideration of the distance based on the cultural embeddedness. We introduce a measure of categorical distance that lies between institutional and industry distance. It measures how boundary-spanning an investment in a startup is compared to the existing portfolio companies of a focal venture capitalist. It is thus a measure of boundary-spanning between organizations (Hannan, Goldberg, and Kovács, 2016). Categorical distance has two distinct features: Firstly, as a measure of how boundary-spanning a startup is compared to the portfolio companies, it is also a measure of how much experience a venture capitalist already has with this type of startup. It is thus similar to industry

distance (e.g., Sorenson and Stuart, 2001), but more evolved in the way we calculate it (see chapter 4 for details). Secondly, through the use of market categories, it considers cultural embeddedness. A highly embedded startup associated with widely spread categories typically has a lower categorical distance than less embedded new ventures. This way, it is similar to institutional distance, but on an organizational instead of country level. Thereby, our measure based on market categories is different from industry classifications and distances used in other studies (e.g., Sorenson and Stuart, 2001; Sorenson and Stuart, 2008; Ter Wal et al., 2016).

Nevertheless, we expect that the general effect of categorical distance is similar to that of other distance measures. In line with prior findings that venture capitalists are more likely to invest in new ventures with low geographical and industry distance (Sorenson and Stuart, 2001), we expect that the same holds true for categorical distance.

H4: Categorical distance between startup and venture capitalist has a negative effect on the probability of an investment.

3.4.2 The effect of portfolio distinctiveness on distant investments

Venture capitalists, contrary to other audiences, have, to a certain degree, a preference for distinctive, boundary-spanning startups (Pontikes, 2012; Wry, Lounsbury, and Jennings, 2014). The more distinctive a new ventures is, the greater the risk and the higher the uncertainty of an investment become. We argue that venture capitalists make a choice regarding the desired degree of distinctiveness in a portfolio. The average portfolio distinctiveness is a measure of experience with managing the uncertainty of investing in distinctive new ventures. If a venture capitalist chooses to invest in these distinctive, boundary-spanning new ventures, which are harder to understand and potentially more disruptive, this demonstrates a willingness and possible ability to take more risk. More categorically distant new ventures represent high-risk investments because the venture capitalist lacks the required knowledge and knowledge access. It follows that a venture capitalist who is seeking more distinctive and disruptive new ventures and is willing to take the associated risk will also be willing to accept more distant new ventures.

H5: Venture capitalists with experience in investing in distinctive startups have a greater probability of investing in more categorically distant startups.

3.4.3 Portfolio diversification through distant new ventures

A venture capitalist's portfolio diversification is a way to reduce idiosyncratic risk in a portfolio. Diversification can take place in several dimensions, for instance, the industry, the geography, or the development stage of a new venture. Our definition of diversification is based on new ventures' entrepreneurial identity measured by market categories. Many authors have found evidence that the venture capitalist's specialization is more beneficial than diversification to the new venture's and the venture capitalist's performance (Chan, 1983; Bygrave, 1987; Gompers, Kovner, and Lerner, 2009). However, the degree of specialization versus diversification varies across venture capitalists. The risk-reducing effect of diversification might be a strategy pursued by risk-averse venture capitalists, or by venture capitalists who do not have a specific knowledge background that they could use and apply to a specialization strategy. For instance, a venture capitalist's investment strategy might be to (co-)invest in deals from their network, namely, a network-based investment strategy, which might be to invest across industries and thus to diversify. In order to diversify the investment portfolio, venture capitalists need to, by definition, invest in new ventures that are distant from the existing portfolio. A possible alternative argumentation is that, as we have put forth above, diversifying venture capitalists are risk-averse, and that these risk-averse investors would prefer less distant investments. However, we believe that the effect of the desire for diversification, that is, investment in categorically distant new ventures, is stronger than the risk aversion. Diversification can be achieved through spatially distant new ventures to diversify geographically or industry distant startups to diversify into various industry categories. Equivalently, categorically distant new ventures are more appealing to venture capitalists who generally try to diversify their portfolios.

H6: Venture capitalists with diversified portfolios have a higher probability of investing in more categorically distant startups.

3.4.4 Effect of status on distant investments

Venture capitalists gain status from their position in the syndication networks: The more central a venture capitalist is, the higher his or her status (Podolny, 1993). A central position in this network means that a venture capitalist has a large number of ties to other venture capitalists who also have several ties to other venture capitalists, which represents indirect ties to the focal venture capitalists. In this network, information is shared through these ties, for instance, regarding market trends, investment opportunities, or relevant experts. High-status venture capitalists thus have

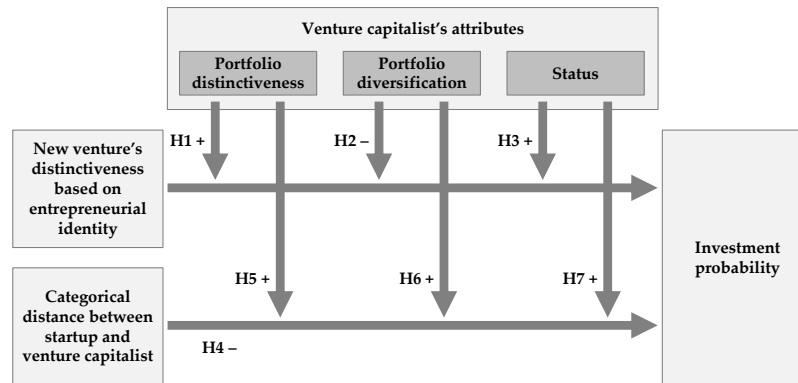


FIGURE 3.3: Conceptual model

access to a greater quantity of information, which reduces the need for proximity in other areas. Consequently, the need for geographical proximity is reduced (Sorenson and Stuart, 2001). Correspondingly, the same holds true for categorical distance: A venture capitalist with high status can use his or her indirect network, that is, the ties to the direct ties of the focal venture capitalist, to fill in knowledge gaps that he or she might have with a more distant new venture or to obtain information about a potential investment candidate. The need for proximity is thus reduced. Fully in line with the argumentation by Sorenson and Stuart (2001) regarding the effect of a venture capitalists centrality on geographical distance we expect that:

H7: The higher the status of a venture capitalist is, the greater the probability of investing in more categorically distant startups.

3.5 Summary and conceptual model

We have developed seven hypotheses to answer the overarching research questions of what type of investors invest in which startups and under what conditions. We use two venture capitalist's portfolio attributes and the status of the investor to predict probability changes in investing in distinctive and categorically more distant startups. The tradeoff between distinctiveness and legitimacy is summarized under the entrepreneurial identity created by a startup's cultural and structural embeddedness. The first line of reasoning anticipates a positive effect of more experience with distinctive new ventures and of a venture capitalist's higher status, and a negative effect of portfolio diversification on the probability of investments in distinctive new ventures. The second line of reasoning anticipates a generally negative effect of categorical distance between a venture capitalist and a startup on the investment probability, which is positively moderated

by all three attributes of a venture capitalist, namely, portfolio distinctiveness, portfolio diversification, and status.

Both parts of our argumentation are connected by the use of market categories to measure a startup's distinctiveness and categorical distance, as well as the average distinctiveness and diversification of a venture capitalist's portfolio as described in chapter 4. Distinctiveness and categorical distance also share a similar interpretation; both being measures of novelty, the former measures a startup's novelty compared to all other startups, whereas the latter measures novelty compared to the portfolio companies of a focal venture capitalist. The conceptual model in figure 3.3 provides a graphic display of the proposed relationships.

Chapter 4

Methodology: data selection and variable definition

4.1 Introduction to methodology: data selection and variable definition

In the following section, we outline how the data were collected and filtered in order to derive the base dataset. We subsequently describe in detail how the dependent, independent, and control variables were calculated. While the general definitions of our variables are mostly proven definitions that have been applied previously in similar or otherwise applicable research (e.g., Sorenson and Stuart, 2001; Ter Wal et al., 2016), we attempt to be more specific in our definitions by including a consideration of treating the exception. We start with a description of the data collection process and which filters were applied to derive the base dataset. We then briefly describe the dependent variable, followed by the five independent variables, of which three function as moderators to test our hypotheses. The controls defined in the last part of the chapter include variables on three levels: the startup/funding-round level, the venture capitalist/syndicate level and the venture capitalist - startup dyad level.

4.2 Data collection

Our main dataset is based on funding rounds that are included in the Crunchbase database¹. Crunchbase was founded in 2007 as a crowd-based platform through which to collect information on the funding rounds of startups, which are featured on the *TechCrunch* technology blog. The data have been used in prior research and are considered to offer a complete overview of funding activity in the US technology sector (Ter Wal et al., 2016; Alexy et al., 2012). The platform is continually further developed by adding more information and structure. While it is still crowd-based,

¹www.crunchbase.com

Crunchbase is one of the major informational sources for investors, startups, and founders, many of whom are contributors.

While we began our research in the beginning in 2015, we continuously updated our dataset with the latest available data until we introduced a version-freeze for our research, with version 3.0 of the spread-sheet format's downloadable version from February 12, 2016. However, in June 2016, a new startup categorization system was introduced on the platform. Not only were individual startups categorized, but they were also grouped into a higher-level categorization system. For our research, we used version 3.0 as the main dataset and, after applying the filters described below, added to the remaining startups the individual details and group categorization information from the database version 3.22 from June 22, 2016.

We used the downloadable spreadsheet format, which includes information from five perspectives: startup-related information, investor-related information, funding rounds, acquisitions, and initial public offerings (IPOs). We used the startup, investor, and funding-round information from the database for our research and applied a number of filters to the raw data download, firstly on the startup level, secondly on the funding-round level, and eventually on the venture capitalist - startup dyad level.

The total dataset is comprised of 121,257 funding rounds that were received by 93,793 startups. We removed all of the startups with missing founding dates or a funding date before the founding date leading to a negative company age² (29,236 startups), missing country information (4,808 startups), non-US startups (21,343 startups), no funding rounds (6,206 startups), no category information (590 startups),³ and a founding date before 2005 (7,069 startups). After the application of these company-level filters, a subset of 24,541 startups with 48,328 funding rounds remained. Naturally, Crunchbase's data quality and completeness has risen in recent years and is higher for US American startups due to the origination of the platform in 2007 and its founding location in San Francisco, CA, USA. A more detailed inspection showed that data were particularly sparse before 2005. That is the reasons why we excluded non-US startups and new ventures founded before 2005 from our main dataset.

On the funding-round level, we were only interested in funding rounds that were classified as venture capital and thus excluded all other types

²We excluded all startups without a founding date (29,148 startups), but excluded only the individual funding rounds that had a recorded funding date before the founding date (243 funding rounds, removing 88 startups). We included subsequent funding rounds if these were recorded after the founding date.

³This splits into 547 startups without category information in the dataset from February 2016, 41 startups that were no longer included in the June dataset (v. 3.22), from which we took the updated company categories, and two startups included in the June dataset but without category information.

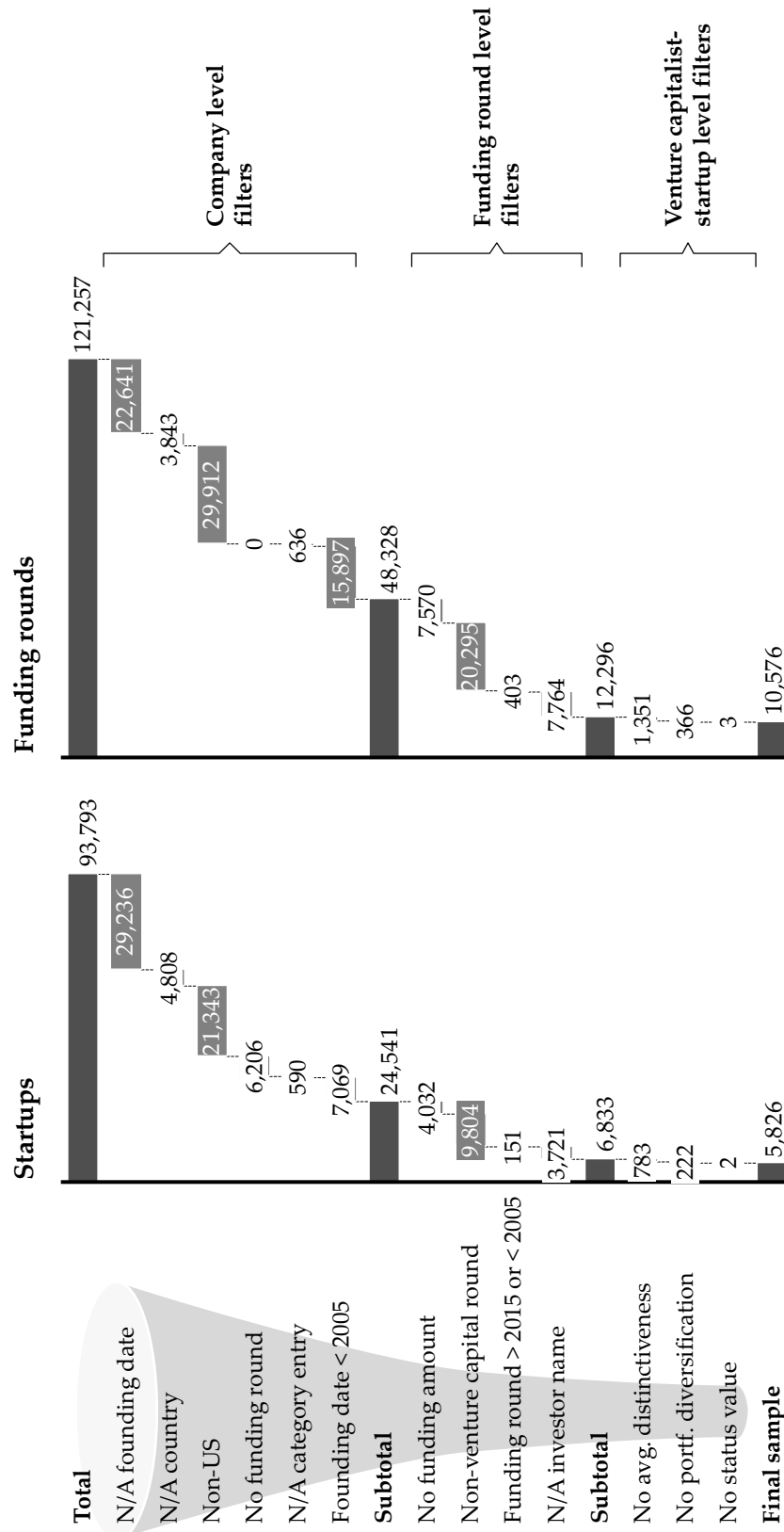


FIGURE 4.1: Filters to derive base dataset from the original Crunchbase file

(for example, debt financing, mezzanine, seed financing, and IPOs). We excluded funding rounds in which no funding amount was indicated (7,570 rounds), that were classified as non-venture capital (20,295 rounds), that took place after December 31, 2015 (341 rounds), funding rounds that took place before 2005 (62 rounds), and with omitted investor names (7,764 rounds). The remaining dataset included 12,296 funding rounds that were received by 6,833 startups.

The dataset after application of the company and funding-round-level filters resulted in 31,803 realized venture capitalist - startup combinations, as many funding rounds were syndicated. Due to the nature of our independent variables, we had to apply three additional filters to the venture capitalist - startup level. Firstly, to calculate the average distinctiveness of the startups in a venture capitalist's portfolio, he or she needs to have made at least one investment prior to the focal round. We thus excluded all venture capitalist - startup observations in which the focal venture capitalists had not made prior investments (8,086 observations, fully excluding an additional 783 startups and 1,351 funding rounds)⁴. Secondly, in order to calculate a venture capitalist's portfolio diversification, he or she needs to have made at least two prior investments in different startups. We therefore excluded an additional 2,638 venture capitalist - startup observations (222 startups and 366 funding rounds in total were dropped from the dataset). Thirdly, we excluded all venture capitalist - startup observations in which the focal venture capitalist was not connected to the core syndication network, to avoid distorting the status measure (115 observations, fully excluding two additional startups and three funding rounds; see section 4.4.4 for methodological details).

The final base dataset includes 29,000 realized venture capitalist - startup observations in 10,576 funding rounds by 5,826 startups. Figure 4.1 gives a graphic depiction of the funnel that was applied to derive this dataset.⁵

For the calculation of network-based measures such as distinctiveness, categorical distance, average portfolio distinctiveness, portfolio diversification, and status, we did not apply the date filters (that is, founding and

⁴It is important to note that, for instance, the 1,351 funding rounds dropped do not include all of the 8,086 observations dropped from the sample. If, in a syndicated funding round, only one investor had not made a prior investment, the remaining observations stayed in the sample, and the funding round was not included. The funding round was only completely dropped in the case that all the investors had not made prior investments or the focal investor was the sole venture capitalist.

⁵The exact order in the R script of the applied filters can differ slightly from the displayed and described order due to simplifications in the coding. However, the figures presented match the effect of the applied filters exactly, and the resulting sample is unaffected by the order. For instance, we excluded observation of venture capitalists who are not connected to the core network as the last step after adding the unrealized ties. This was necessary to check the robustness of the analysis to including these observations.

funding dates after 2005) in order to be able to calculate these retrospective measures for the startups in our dataset. The creation of different data sets for the calculation of measures and for the final analysis to test the hypotheses has been found useful by other researchers (e.g., Ter Wal et al., 2016).

4.3 Dependent variable

The dependent variable is the probability of an investment by the focal venture capitalist in a specific startup. To make this measurable, we also added to the base dataset of realized venture capitalist - startup investments unrealized investments, as described in section 5.1. It follows that the dependent variable in our regression model is a binary variable called *dummy actual data*. The variable equals 1 if an investment is realized and equals 0 if it is an unrealized investment tie. By definition, 50% of our observations are realized ties with the dependent variable being equal to 1, as each realized investment is matched with one unrealized investment tie.

4.4 Independent and moderating variables

We use five independent variables in our research to test our hypothesis. The startup's distinctiveness and the categorical distance between venture capitalist and startup are the primary independent variables. Three venture capitalist attributes, the average portfolio distinctiveness, the portfolio diversification and the status function as moderators of the primary independent variables. A detailed specification follows, including examples of these variables.

4.4.1 Startup's distinctiveness

We follow the lines of Kennedy (2008) and Navis and Glynn (2011) by measuring legitimacy from cultural embeddedness through the analysis of associated categories. As we are interested in the startup's distinctiveness, which is the antagonist of legitimacy, we are concerned with a startup's level of boundary-spanning. Our approach to measuring the distinctiveness or novelty mainly follows that of Hannan, Goldberg, and Kovács (2016), who measure the distinctiveness of films and restaurants by a measure called "atypicality" (p. 216). Startups are objects that are consumed by venture capitalists after rigorous due diligence. A key element of new ventures that interests venture capitalists is how the focal startup compares to other startups, specifically concerning the degree of novelty. In order to measure this novelty, we have developed a measure of distinctiveness that measures how much a startup resembles or creates new aspects

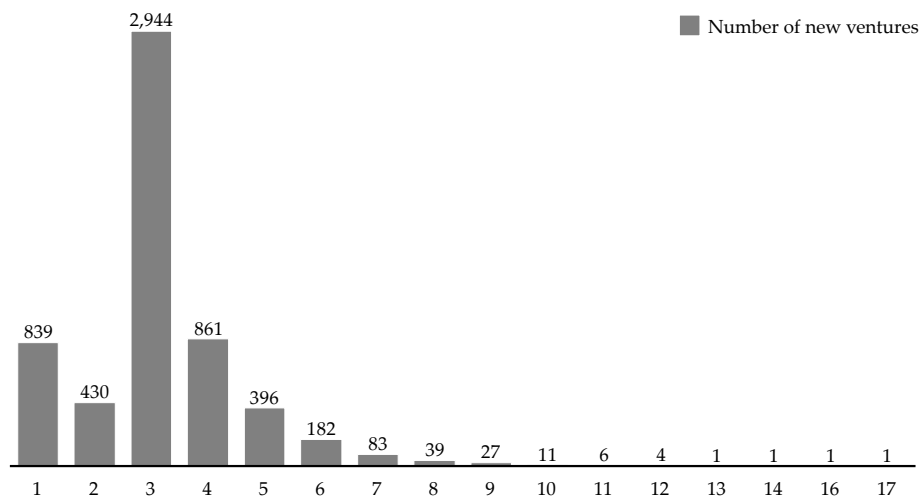


FIGURE 4.2: Frequency distribution of categories per new venture in the sample

and thus deviates from existing beliefs. We use the Crunchbase categories with which each startup is associated in order to measure this distinctiveness. For instance, *M.Geni* is an online fashion store that sells Italian leather shoes. It falls into four Crunchbase categories: *e-commerce*, *fashion*, *online shopping*, and *retail*. Based on this categorization, we need a measure that considers three aspects: Firstly, from an investor perspective, it makes a difference if investors invest in two startups that are both categorized as *e-commerce*, *fashion*, *online shopping*, and *retail*, or if they invest in one startup that is categorized as *e-commerce* and *online shopping* and one startup that falls under *fashion* and *retail*. It is fair to assume that the first two startups with the same categories are more similar than the second two. Only counting the individual categories that the investors have invested in would thus not be sufficient. Secondly, we need to measure how similar two categories are: *E-commerce* and *online shopping*, for example, are relatively similar categories, as one depends on the other. We measure the similarity of two categories based on their prior co-occurrence. Thirdly, as we aim to develop a measure of distinctiveness or novelty, we need to consider whether the categories themselves and whether the combination in which they occur are new. The distinctiveness measure that was developed by Hannan, Goldberg, and Kovács (2016) fulfills these criteria. The authors use it to explain the consumption behavior of restaurant visitors and movie-goers who consume and rate restaurants and movies that are categorized. Similarly to our goal, the researchers are interested in explaining how consumers vary in their preferences for boundary-spanning objects depending on their preferences for distinctiveness.

Distance – In our Crunchbase dataset, each startup has been assigned

to at least one and up to 17 categories by contributors such as startup founders, employees, or venture capitalists. The full distribution, with a mean of 3.2 categories per company for the 5,826 companies in our sample, is displayed in figure 4.2. We denote the set of categories as γ . In order to measure the similarity between two categories, we use the Jaccard (1901) similarity measure: a simple count of the number of startups that are associated with both categories, divided by the number of startups that are associated with either one or both of the categories. Formally, if k and l are two categories of the set of categories γ of startup i , $|k \cap l|$ is the number of startups that are associated with both categories and $|k \cup l|$ is the number of startups that are associated with either or both categories; then the Jaccard similarity J between the two categories k and l is as follows:

$$J(k, l) = \frac{|k \cap l|}{|k \cup l|} \quad (4.1)$$

This similarity measure has a maximum value of 1 in case of full co-occurrence, meaning that both categories always occur together, and 0 if they have never occurred together before. While this similarity measure based on co-occurrence is useful in comparing two categories with each other, we are interested in the distance between the categories and in the magnitude of said distance. Shepard (1987) set out to find a general law with which to make distances interpretable within a metric space. By analyzing a great number of other researchers' work on stimuli that were applied to humans, to animals, and in other domains, he found an upward concave relationship between the psychological distance and the measure of generalization. Generalization here is meant as the transferability of the response to a certain stimulus, which means that there is an exponential-decay relationship between the distance and the generalizability. For instance, if there is one observation with an attribute on an arbitrary scale with a certain outcome, the law helps to predict the likelihood that the same outcome will occur, depending on the distance between the attributes. The mathematical considerations confirm that, ideally (for example, when there is no delay in stimuli), the similarity has a negative exponential relationship with the psychological distance. It has also been shown that, theoretically, the so-called 'Shepard's Universal Law' will also hold true in more general settings (Tenenbaum and Griffiths, 2001; Chater and Vitányi, 2003). Mathematically, the similarity is thus defined as follows:

$$sim(k, l) = e^{-\lambda d(k, l)}, \quad \lambda > 0. \quad (4.2)$$

We thus transfer these findings as other researchers have (e.g., Hannan, Goldberg, and Kovács, 2016) to our setting by combining the categorical

Jaccard similarity measure (equation 4.1) with the relationship from Shepard's Universal Law in equation 4.2 so that the distance between two categories is defined as follows:

$$J(k, l) = \frac{|k \cap l|}{|k \cup l|} = e^{-\lambda d(k, l)}, \quad \lambda > 0. \quad (4.3)$$

Solving for the distance between two categories, we attain the following equation:

$$d(k, l) = -\frac{\ln(J(k, l))}{\lambda} \quad (4.4)$$

This distance measure is the basis for the startup attribute distinctiveness, the categorical distance, the average portfolio distinctiveness, and the portfolio diversification. The sizing coefficient λ is set to 0.5. We calculate the distances of each category pair for each month based on the assigned categories in the previous five years, excluding the focal month.⁶ For instance, in order to attain the category pair distances for startups that received funding in January 2009, we use the categories of all startups that were funded in the five years up to and including December 2008. We use the dataset for the calculation of network measures, which includes the startups that received funding before the year 2005 recorded in the Crunchbase database. While five years is an arbitrary number, we feel confident that any category combinations that have occurred more than five years prior to the focal month will have little meaning for the current analysis. Multiple researchers have also used five years in their calculation of network-based measures in venture capital settings (e.g., Sorenson and Stuart, 2001; Sorenson and Stuart, 2008; Zhelyazkov and Gulati, 2016). Furthermore and as previously described, we aim to be very explicit about our approach: If a category pair has never appeared before, the Jaccard similarity is equal to 0. As the $\ln(0)$ is not defined, we added 0.000000001 to the Jaccard similarity score. The maximum distance value for a set of categories is thus $-\ln(0.000000001) = 20.72$. The minimum value is naturally limited by the maximum value of the Jaccard similarity, which is 1, of which the natural logarithm is 0. The minimum distance thus also equals 0.

Distinctiveness – As previously outlined, we define the measure of distinctiveness of a startup as the degree of novelty or deviation from existing cultural beliefs that can be attributed to a startup. We consider the categories to be anchored in a conceptual space, which we analyze with a

⁶An organization's position is also dependent on the current environment and thus may evolve over time (Hogg and Terry, 2000). Therefore, we always measure a startup's distinctiveness at the time of a funding round. This way, we account for the fact that the environment, and, with it, the perceived categorization, changes over time. For instance, if certain category combinations have occurred before, they are less distinctive than when they occur for the first time.

geometric model so that each startup category is a subset in this domain (compare Hannan, Goldberg, and Kovács, 2016). In contrast to research in which each object is associated with only one category, the underlying assumption here is that, if an object is placed in multiple categories, it cannot be fully associated with any one of those categories. Rather, in a geometric space, the startup lies somewhere between all of the categories with which it is associated. This 'betweenness' makes the startup distinctive, which, in the context of the startup domain, can be considered to be a measure of newness.

Thus, the more categories with which a startup is associated, the more distinctive it is. In addition, the more diverse the set of categories as measured by the distance between them, the greater the distinctiveness. The latter is especially important because, if categories are very similar, for example, *e-commerce* and *online shopping*, they are more indicators of lower distinctiveness than others.

We base our measure of distinctiveness on the approach developed by Kovács and Hannan (2015). While we follow their definition closely, we make an adaptation to improve the distribution of the distinctiveness measure. For binary settings like ours, that is, a category is either assigned to an object or organization or not, the authors propose using the average pairwise distance between the startup's categories as input for the measure of distinctiveness. Using their approach with the distances obtained from our network of categories,⁷ the distinctiveness measure is highly skewed to 1 and has a relatively small standard deviation if we exclude the observations with only one category assigned, in which cases distinctiveness equals 0. Our adapted approach, which we outline below, has the same general properties, but leads to a more even distribution, capturing more variance in a startup's distinctiveness. The first property is that, with increasing distance from an increasing number of associated categories, the distinctiveness increases. The second property captures the effect of increasing distinctiveness when the distance between the categories increases.

Our approach is as follows. We sum up the distances between two categories for each startup multiplied by the weight of each category. We denote the set of categories of startup x as C_x and the weight of category k as $w(k, x)$. In our dataset, every startup is only associated with each category no more than once, so that the weight of each category is $1/|C_x|$ with $|C_x|$ being the number of categories with which the startup is associated. In other applications, when several category associations would be possible, this weight could be different for each category. For each startup, it holds true that $\sum_{k \in C_x} w(k, x) = 1$. Let k and l be two categories and $d(k, l)$ the

⁷Each research setting results in different distance levels, depending on the number of different categories used, the distribution of the number of categories assigned per object, and the degree of overlap.

distance between the two categories, then the sum of the distances for a startup $D(x)$ is defined as follows:

$$D(x) = \frac{\sum_{k \in C_x} \sum_{l \in C_x} w(k, x)w(l, x)d(k, l)}{2} \quad (4.5)$$

One might be surprised that we divide the sum of the weighted distances by 2, but this is necessary as we want each distance to only enter the total distance calculation once. The numerator in equation 4.5 results in a double summation of each weighted distance, making the correction necessary.

For all cases in which a startup is associated with more than one category, distinctiveness for startup x is defined as follows:

$$A(x) = 1 - \left(\frac{1}{1 + \frac{D(x)}{|C_x| - 1}} \right) \quad \text{if } |C_x| > 1. \quad (4.6)$$

For cases in which a startup is only associated with one category, distinctiveness is 0, as, in the geometric space, the startup is not in between categories but clearly assigned one category. The minimum distinctiveness value thus equals 0, while the maximum value approaches 1 as the average of the sum of the distances becomes larger. The larger the distances, the larger the denominator in the distinctiveness formula in equation 4.6 will be.

In what follows, we aim to briefly demonstrate our calculations by using the previous example of the startup *M.Geni*, which falls into four categories: *e-commerce*, *fashion*, *online shopping*, and *retail*. We create all unique pairs from these categories (six in the case of four categories). For each pair, such as *fashion* and *e-commerce*, we calculate the distance. Let us assume that 10 startups are categorized with *fashion* and 12 with *e-commerce*, and that 6 of these fall into both categories. We only consider startups that have been founded in the past five years. We then calculate the Jaccard similarity as $J(\text{fashion}, e - \text{commerce}) = 6 / (10 + 12 - 6) = 0.375$, which leads to a distance of $d(\text{fashion}, e - \text{commerce}) = -\ln(0.375) / 0.5 \approx 1.96$. In fact, the distance was ~ 4.03 . The other seven distances were calculated similarly, with values ranging from ~ 4.04 to ~ 7.68 . We multiply the weights of each category in a pair with the distance so that the values per category pair range from ~ 0.25 to ~ 0.48 . We then add the weighted distances in order to attain the total distance score of the startup, which, in our case, equals ~ 2.02 . *M.Geni* is associated with four categories, so that $|C_{M.Geni}| = 4$, and the distinctiveness is calculated as follows:

$$A(M.Geni) = 1 - \left(\frac{1}{1 + \frac{D(x)}{|C_x|-1}} \right) = 1 - \left(\frac{1}{1 + \frac{2.02}{4-1}} \right) \approx 0.40.$$

4.4.2 Average portfolio distinctiveness

The group of venture capitalists is heterogeneous, and they adopt different strategies for their portfolios. Regarding distinctiveness, some venture capitalists have a preference for investments in more distinctive startups while others invest in those less distinctive. We analyze historical investments that were made in the five years prior to the focal investment round. We denote the companies in the investment portfolio from the past five years from the venture capitalist y as P_y and take the average of the distinctiveness of the startups in their portfolio:

$$PA(y) = \frac{1}{P_y} \sum_{x \in P_y} A(x). \quad (4.7)$$

For the interpretation of the measure, we assume that an investment in a company has been voluntary. We thus assume that prior investment behavior also demonstrates a preference for the chosen investments: If an investor has invested in highly distinctive startups, we assume that he or she has a preference for distinctive startups.

4.4.3 Portfolio diversification

In addition to the distinctiveness of the startups in venture capitalists' portfolios, we are also interested in understanding whether they focus their investments on similar companies or instead diversify their investments. We measure the portfolio diversification by a measure similar to a measure Hannan, Goldberg, and Kovács (2016) call "variety" (p. 216) in their analysis of consumption behavior. For instance, an investor who invests in a *fashion* startup and an *e-commerce* startup is more diversified than one who invests in two startups that are both categorized as *fashion* and *e-commerce*. Again, we create a measure that takes into account not only whether an investor invests in different categories or category combinations, but also how different these investments are. We use the previously defined pairwise distances of categories to measure this difference. One aspect complicating this process is the fact that we are comparing sets of categories rather than only a pair, which requires an appropriate measure.

In the categorization literature, differences between category sets are typically measured by the Hausdorff distance, which was developed by Felix Hausdorff (Burago, Burago, and Ivanov, 2001). The general idea is to

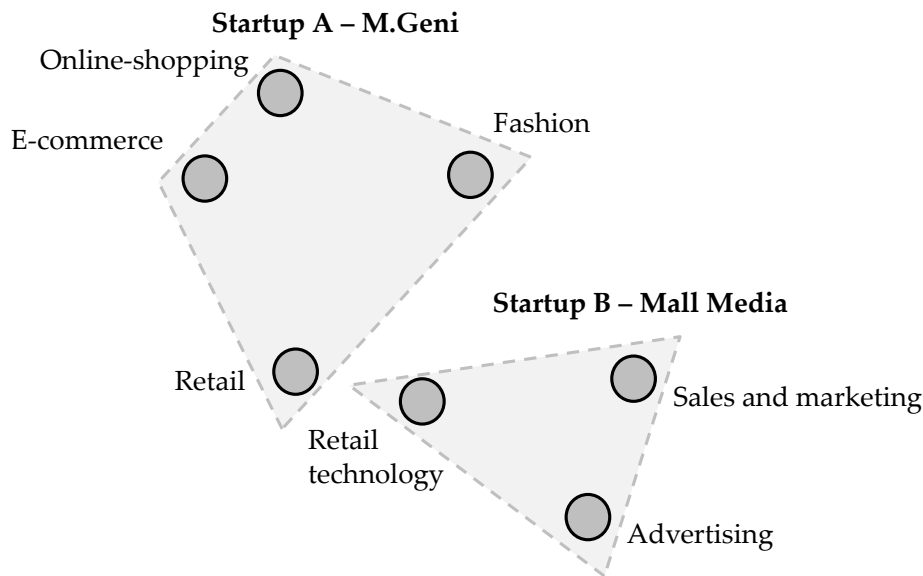


FIGURE 4.3: Comparing category sets of two startups in a geometric space.

compare two subsets of a metric space and to calculate their distance based on the longest distance of the pairwise distances from one set to the closest point of the second set. It is thus a good measure of resemblance (Huttenlocher, Klanderman, and Rucklidge, 1993) and has specifically been shown to be a good measure in cases of linguistic variables (Hung and Yang, 2004), which makes it applicable to our research.

As an example, we compare the categories of our previous example startup, *M.Geni*, to another new venture in our sample called *Mall Media*⁸: a provider of mall-marketing solutions such as coupons, which has been placed into the three categories of *retail technology*, *sales and marketing*, and *advertising*. As is visualized in figure 4.3,⁹ some categories, such as *retail technology* and *retail*, lie relatively closely to each other, whereas *e-commerce* and *advertising* lie further apart. The Hausdorff distance is then calculated as the maximum of the distances between each category of startup A and the closest category of startup B, and each category of startup B and the closest category of startup A. In the hypothetical example from figure 4.3, we would calculate the pairwise distance between all categories from *M.Geni* and *retail technology*, as the latter is the category closest to that of *Mall Media*. In addition, we would calculate the distances between all of the categories from *Mall Media* with the category *retail* from *M.Geni*. The maximum value of the calculated distances equals the Hausdorff distance. Formally, we denote a category of startup A as c_A and the set of categories of

⁸<http://mallmediainc.com/>

⁹The distances in the visualization are arbitrarily chosen. The figure is intended to help the reader in understanding the concept presented. It is not an exact representation of the actual distances.

startup A as $C_A = \{c_{A,1}, c_{A,2}, \dots, c_{A,n}\}$. The distance between a category c_A and a set of categories C_B is defined as $d(c_A, C_B) = \min_{c_B \in C_B} \|c_A - c_B\|$. The longest distance between all categories of set C_A and the closest category of set C_B is defined as follows:

$$d(C_A, C_B) = \max_{c_A \in C_A} d(c_A, C_B). \quad (4.8)$$

The Hausdorff distance based on the distance measure above is given by the following equation:

$$HD(A, B) = \max(d(C_A, C_B), d(C_B, C_A)). \quad (4.9)$$

The Hausdorff distance only takes into consideration the maximum of the calculated distance so that, for instance, the *fashion* category of *M.Geni* does not impact the calculation of the Hausdorff distance even though the existence of the category places *M.Geni* closer to *Mall Media*. We therefore use an extension of the Hausdorff distance not only to consider the furthest distance between the category pairs, but also to consider the whole set. This modification was introduced by Dubuisson and Jain (1994) and has been applied to socio-cultural domains (e.g., Hannan, Goldberg, and Kovács, 2016). The modification involves calculating the average of the pair-wise distances of the categories for one set between the closest category of the second set and vice versa. The maximum value of these two averages is the so-called 'modified Hausdorff distance'. Formally, the average of the distances between each category in a set and the closest category in the second set is given by the following equation:

$$dI(C_A, C_B) = \frac{1}{n_A} \sum_{c_A \in C_A} \min(d(c_A, C_B)). \quad (4.10)$$

The modified Hausdorff distance is then the maximum of the two sets and given by the following equation:

$$HDI(C_A, C_B) = \max(dI(C_A, C_B), dI(C_B, C_A)). \quad (4.11)$$

In order to calculate portfolio diversification for each venture capitalist, we take the average of the pairwise modified Hausdorff distances between each of the portfolio companies. For example, if an investor has made investments in five startups in the past five years, up to the month before the focal month, then we calculate the modified Hausdorff distance for each of the ${}_n C_r = \frac{n!}{r!(n-r)!} = \frac{5!}{2!(5-2)!} = 10$ unique startup category set pairs. We then take the average of these pairwise distances so that the portfolio diversification PV is given by the following equation:

$$PV(y) = \frac{1}{P^y} \sum_{x \in P_y} \sum_{z \in P_y} HDI(C_x, C_z). \quad (4.12)$$

4.4.4 The venture capitalist's status

Our third independent variable is the status of the venture capitalist. As previously explained, we are interested in the status that is assigned to venture capitalists based on the syndication networks of which they have been part. In order to measure status, we use the widely accepted eigenvector centrality that was first introduced by Bonacich (1987). The appeal of this measure is that it not only consider the direct network of a venture capitalist, but the status also depends on the centrality of its ties, whose centrality depends on their ties, and so forth. The status measure is widely accepted and has been used by many researchers as a status measure for venture capitalists (among others: Podolny, 2001; Sorenson and Stuart, 2001; Shipilov and Li, 2008; Bothner, Kim, and Lee, 2015).

In order to calculate the status of a venture capitalist at the point of an investment (the assigned status may change over time, depending on the new syndicate networks that are created), we consider the co-investment network from all investments in our dataset (including those before 2005) in the past five years prior to the focal month.¹⁰ For each month m , we create an adjacency matrix R_m based on the co-investments of the venture capitalists so that the rows and columns are investors and the matrix value is the number of times that investors have co-invested. Each cell R_{ijm} is the number of startup funding rounds in which the two investors i and j have jointly invested. CI is all of the co-investors in a focal funding round. We then use the eigenvector centrality score developed by Bonacich (1987) to measure status S for investor i at time m defined as follows:

$$S_{im}(\alpha, \beta) = \sum_{j \in CI} (\alpha + \beta S_{jm}) R_{ijm}. \quad (4.13)$$

As described previously, the status of investor i depends on the status of the other co-investors, such as S_{jm} , whose status also depends on the status of his or her co-investors and so forth. The β parameter sets the effect of the centrality scores of an investor's co-investors. If β is positive, the investor gains status if his or her co-investors have high centrality scores, whereas a negative β would result in a positive effect on status if co-investors have low centrality scores. In the case $\beta = 0$, the status would be independent of the centrality of the investor's co-investors and would only depend on the number of prior investments, which we consider in the experience variable

¹⁰We have limited our analysis to five years, as it is unlikely that, in the case of non-investment in the past five years, relationships have remained intact (Sorenson and Stuart, 2001). We hereby take into account that relationships erode over time.

below (Sorenson and Stuart, 2001). As we also consider the information flow and experience that is gained from the co-investors' central network positions as positive, we assign β a positive value. We follow other researchers (e.g., Podolny, 1993; Podolny, 2005; Bothner, Kim, and Lee, 2015) and set β to three-quarters of the reciprocal of the largest eigenvalue.

α is a scaling parameter that we use in order to make the status scores comparable over time (that is, independent of the network size) so that, for each period m , the mean centrality score equals 1 (Bonacich, 1987).

For the calculation of the status measure, we only included the venture capitalists who are connected to the core network at time m . A small share of investors is either not connected to a network at all or builds its own network. In order to avoid measuring high centrality scores in the non-core networks, we exclude any non-core parts in the calculation of the status score. 115 realized venture capitalist - startup observations and 24 venture capitalists were fully excluded due to this adjustment, leaving a total of 2,323 venture capitalists in the sample. All the venture capitalists who did not have any prior syndicated investments are assigned a status score of 0.

4.4.5 Categorical distance

In the theoretical framework derived in chapter 3, we developed a notion of categorical distance between two organizations: the venture capitalist and the startup. Our measure of categorical distance lies between the classical cultural distance based on the cultural dimensions developed by Hofstede (1980), the institutional distance (Kostova and Zaheer, 1999), and the industry distance, that is, how dissimilar the industries are that two organizations are active in (Sorenson and Stuart, 2001). We use the same market categories that we use to measure a new venture's distinctiveness, namely, the cultural embeddedness, and take into account how similar or dissimilar these categories are. That is why our measure of categorical distance is different from the industry distance measure developed by Sorenson and Stuart (2001), who only count a venture capitalist's number of prior investments in the same industry segment as the focal startup.

Specifically, as we do to model portfolio distinctiveness, we calculate the modified Hausdorff distance between the focal startup and each of the existing portfolio companies of the venture capitalist. The calculation involves taking the average of the distances between each set of one startup to the closest category of the second startup and vice versa. The maximum of the two averages is the modified Hausdorff distance between two startups, as we can see from equation 4.11, which we repeat here for reference purposes:

$$HDI(C_A, C_B) = \max(dI(C_A, C_B), dI(C_B, C_A)). \quad (4.14)$$

To calculate the categorical distance, we take the average of pairwise modified Hausdorff distances between the focal startup and the portfolio companies of the focal venture capitalist. Please note that, contrary to the calculation of portfolio diversification, we do not calculate the modified Hausdorff distances between each of the portfolio companies. The categorical distance between the portfolio companies P_y from the last five years of venture capitalist y and the focal startup B is thus given by the following equation:

$$PV(y) = \frac{1}{P_y} \sum_{x \in P_y} HDI(C_x, C_B). \quad (4.15)$$

4.5 Control variables

We control for a number of aspects based on previous research in order to isolate the effects of our independent and moderating variables. The control variables exist on three levels: the startup/funding-round level, the syndicate/venture capitalist level, and the venture capitalist - startup dyad level. This structure reflects the first dimension of the literature review framework from chapter 2 and figure 2.2. Table 4.1 gives an overview of the control variables used by level. In cases where we use logarithmic transformations, they are specified as $\log(1 + variable)$ (Hsu, 2006).

4.5.1 Startup and funding-round level controls

In this section we describe the startup and funding level controls, which are independent of the venture capitalist.¹¹ The funding-round level controls include variables that may change from one funding round to another for an individual startup.

Startup location – On the startup level, independent of the funding round, we only control for the region in which the startup is located. We use three US state dummies for California, Massachusetts, and New York,

¹¹One may notice that we do not directly control for the quality where some other researchers control for it, for example, by counting the number of patent applications a startup has made (e.g., Ter Wal et al., 2016). We do not do so, as we have a broad industry perspective and the number of patents varies from industry to industry and even within one industry. For instance, Mann and Sager (2007) show that even within the software industry, the number of patents varies strongly depending on the sector, and only about a quarter of all firms in their sample have issued patents at all. We thus decided to not control for the number of patents, as we do not expect this to be a good indicator for quality in our cross-industry analysis. Furthermore, we also only consider startups for the unrealized ties that received funding from a venture capitalist in that focal quarter. Consequently, they must at least be above a quality level threshold to qualify as a potential investment. Therefore, we strongly believe that further controlling for the quality of a new venture would not improve our results.

Startup/funding round level	Venture capitalist/syndicate level	Venture capitalist-startup dyad level
Startup location (3 dummies)	VC age	Geographical distance
Startup age	General experience	Industry distance
Investor count (26 dummies)	Industry experience	Prior investment (dummy)
Raised amount	Mean affiliation	Affiliate geographical distance
Round number (9 dummies)		Affiliate industry distance
Year (10 dummies)		

TABLE 4.1: Overview control variables by unit of analysis

in which we find a high concentration of startups¹² (Ter Wal et al., 2016). The dummy variables equal 1 if the headquarters of the focal startup are located in the respective state.

Startup age – We calculate the startup age at the time of the funding round in months and exclude cases in which the variable is negative due to inconsistency in the data (typically, funding does not occur in the venture capital stage if the startup has not already been funded). If the investment was made in the same month as the foundation, we set the age at funding to one month.

Investor count – We control for the number of investors in the focal funding round. The syndicate size is important because larger syndicates can distribute their risk (Wilson, 1968) and because more financial and intellectual resources are available in larger syndicates (Ter Wal et al., 2016; Lerner, 1994). Furthermore, larger syndicates might be able to invest in better new venture, as they can overcome competition (Bygrave, 1987; Casamatta and Haritchabalet, 2007). We use dummy variables to control for each level of the number of investors leading to 26 dummy variables.

Raised amount – We control for the amount raised in the focal funding round, as this may attract or exclude certain investors who are restricted in their funding amounts. We use the natural log of the USD value (Ter Wal et al., 2016).¹³

Round number – We count the number of venture-capital-labelled funding rounds in our dataset for each startup, which serves as an indicator for the startup’s lifecycle. The monitoring requirements and thus the need for geographical or industry proximity is highest in the early stages of a

¹²We also found a high concentration of venture capitalists in these areas. However, we do not need to directly control for the location of the venture capitalist, because we control for the region of the startup, and we have multiple controls for the distance between the venture capitalist and the startup, as well as for the geographical distance of the syndicate. Controlling for the venture capitalist’s location would thus not add any explanatory value.

¹³We do not control for the funding amount per investor, because, by controlling for the number of investors and the amount raised, we have already captured the direct effect. Nevertheless, it is important to note that the dollar amount each investor has to contribute might affect the decision to make an investment, as the financial risk increases with greater amounts.

startup (Sorenson and Stuart, 2001). In addition, differences in startup success expectation caused by the lifecycle phase in which a startup finds itself (Podolny, 2001). We try to capture this effect with nine dummy variables because, in our dataset, up to 10 venture-capital funding rounds were received.

Year dummies – Our dataset includes data from January 1, 2005 to December 31, 2015 so that we include 11 full years of funding information. We include a control variable for ten of the 11 years, 2005 to 2016, in order to account for cyclical effects.

4.5.2 Venture capitalist and syndicate level controls

In what follows, we describe the control variables that are attributes of the focal venture capitalist or its portfolio and that are independent of other investors and the focal startup. Furthermore, we include one syndicate-level control that analyzes the syndicate members' affiliation with each other. Syndicate distance measures for the focal startup are included under the venture capitalist - startup dyad controls.

VC age – Similar to the control for the startup age, we control for the venture capitalist age. A venture capitalist's investment patterns may change depending on their maturity. Our dataset does not include the founding dates of the venture capital firms. Sorenson and Stuart (2001) found in their dataset that on average the first investment after inception was made 366 days after the official foundation date. We thus first calculate a possible founding date 12 months prior to the first investment in our dataset. We then count the number of months between the calculated founding date and the focal founding round.

General experience – The experience of a venture capitalist may also affect the investment pattern. We count the number of prior investments that were made by the focal investor. Formally, the general experience of investor r is defined as follows, where x equals 1 when an investment was made:

$$ge_r = \sum_{p \in P} x_{rp} \quad (4.16)$$

Industry experience – We create a similar measure, counting only the prior investments in startups that have the same group category as the focal startup. We denote the group category by g and define the industry experience as follows:

$$ie_{rg} = \sum_{p \in P} x_{rpg} \quad (4.17)$$

Mean affiliation – The mean affiliation measures how frequently venture capitalists in the syndicate of the focal funding round have previously co-invested with each other. Investors may be willing to take more risk or accept more distant investments when they have already invested with the focal co-investors. We calculate the average number of startups in which the investor has co-invested with other syndicate members. We first calculate the number of co-investments ci of investor r with investor s in all previous funding rounds.

$$ci_{rs} = \sum_{p \in P} x_{rsp} \quad (4.18)$$

P is the vector of all funding rounds in the past five years, and x equals 1 if both investors have invested in p . The mean affiliation ma of investor r in funding round i is the average prior number of co-investments of investor r with its syndicate partners:

$$ma_{ri} = \frac{\sum_{v \in V_i} \sum_{p \in P} x_{rsp}}{|V_i|} \quad (4.19)$$

$|V_i|$ denotes the number of investors in the syndicate except for investor r . If no prior investments exist or if a funding round was given by only one investor, the mean affiliation is set to 0.

4.5.3 Venture capitalist - startup dyad level controls

In what follows, we describe the controls that depend on both the individual investor or group of investors and the individual startup. The geographical and industry distance have been found to significantly affect the decision to invest or not to invest in a given startup (Sorenson and Stuart, 2001). Syndicate distance measures are additionally included, as they might mitigate the effect of distance. We follow Sorenson and Stuart (2001) in their definition of affiliation measures.

Geographical distance – Greater geographical distance between a venture capitalist and a startup decreases the likelihood of an investment (Sorenson and Stuart, 2001). We measure distance as the kilometers between the center of towns of the startup's headquarters' and the venture capitalist's headquarters. Due to data limitations, we have no information about the venture capitalist's satellite offices, which may be closer to the startup than the venture capitalist's headquarters. However, we do not expect this to affect our analysis, as we do not see a systematic bias (Bernstein, Giroud, and Townsend, 2016). We take the natural log of that distance measure because, with greater distance, other methods of transportation are used: Cars and trains are replaced by airplanes, for instance. Thus, neither the time nor the required financial resources increase linearly (Sorenson and

Stuart, 2001). In addition, digital communication solutions, such as video conferences, have become less expensive, and their quality has improved. New collaboration tools and a widespread use of cloud-based solutions have also decreased the need for geographical proximity.

Technically, we downloaded the latitude and longitude position data from Google Maps¹⁴ for the town, state, and country combination in our dataset. We used spherical trigonometry¹⁵ to calculate the geographical distance. For about 3.2% of our observations, we were not able to calculate the geographical distance, mainly because we did not have the location data for the venture capitalist. We used mean imputation, that is, inserting the mean of the observed distances, to deal with the missing data. We created the dummy variable *No geographical distance dummy*, which equals 1 for observations where we inserted the mean and 0 otherwise. An overview of methods for imputation and reasons why we chose mean imputation can be found in section 5.2.3.

Industry distance – Similarly to geographical distance, the industry distance affects venture capitalists' decision to make an investment. If they are familiar with the industry and have done research or made investments within it, the likelihood of an investment increases. We measure industry distance by calculating the share of a venture capitalist's investments in a category different than that of the focal startup (Sorenson and Stuart, 2001). We are interested in a more general categorization than that which we used for the calculation of distinctiveness, and we thus use the group categories in our dataset that represent an aggregation of the detailed categories. We denote the group categories of all the investments in the past five years of investor r by G_r and the number of investments in the categories of focal startup i by G_{ri} . The industry distance id is then defined as follows:

$$id_{ri} = 1 - \frac{G_{ri}}{G_r} \quad (4.20)$$

The measure ranges from 0 if all prior investments were made in the same group category to 1 if none of the prior investments were made in any of the startup's group categories.

Prior investment dummy – We created a dummy variable that equals 1 if the focal venture capitalist has already invested in the focal startup in a previous funding round. Venture capitalists are more likely to finance a startup when they are already invested in it in order to demonstrate to other investors that they continue to believe in the new venture.

¹⁴We used the API that is implemented in the *ggmap* (Kahle and Wickham, 2013) R package in order to download the data.

¹⁵We apply the implementation in the *geosphere* (Robert J. Hijmans, 2016) R package and use the most exact method, assuming that the earth is an ellipsoid. We calculate the distance in meters with the default values of the World Geodetic System 1984 (WGS84) with a radius of 6,378,137 meters and an ellipsoid-flattening parameter of 1/298.257223563.

Affiliate geographical distance – We previously described how we control for the geographical distance between a venture capitalist and a new venture. However, when some of the investors in a syndicate are closer to the startup than the focal investor, the latter may trust his or her co-investors to overtake the monitoring and controlling functions (Sorenson and Stuart, 2001). We thus create a measure of the shortest geographical distance of all syndicate members to the startup and invert the measure, because, in cases where there are no other syndicate members, we would need to assign an arbitrary distance. We thus define the affiliate geographical distance agd as follows:

$$agd_{ri} = \frac{1}{\min(gd_{vi})} \quad (4.21)$$

The measure approaches 0 as the distance becomes larger and if no syndicate partners exist. We add 0.01 to the minimum distance to avoid difficulties in cases in which we have to divide by a minimum distance $\min(gd_{ri})$ equal to 0 when the venture capitalist and startup are headquartered in the same city.

Affiliate industry distance – We created a similar measure for the industry distance in order to attain the minimum industry distance in the syndicate. As the industry distance has a range from 0 to 1, there is no need to take the inverse. When there are no syndicate partners, we set the measure to 1 and thus define the affiliate industry distance aid as follows:

$$aid_{ri} = \min(id_{vi}) \quad (4.22)$$

Chapter 5

Methodology: empirical approach and analytical method

In this section, we outline our empirical strategy towards creating the dataset from our base dataset, which is described in section 4.2. This is followed by a more general presentation of the statistics that were applied in our models. In section 5.3, we describe the specific statistical models used to test our hypotheses. In terms of software, we use the open-source programming language R (R Core Team, 2016), including additional packages that add specific functions to the base functionality of the software environment. The detailed packages are listed in appendix B.

5.1 Empirical approach: the addition of unrealized ties

Our base dataset includes 29,000 observed, realized investment ties between a venture capitalist and a startup in a total of 10,576 funding rounds.¹ In our analysis, we aim to test the determinants of this tie formation, that is, how the probability of funding by a certain investor changes based on the startup's, the venture capitalist's and the ties' attributes. In tie-formation terms, we thus have a combination of node characteristics, such as the startup's distinctiveness, and dyadic characteristics, such as the distance between the venture capitalist and the startup (Stuart, 1998). Some researchers use the dyad as the unit of analysis (e.g., Podolny, 1994; Gulati, 1995), while others focus on the node as a unit of analysis (e.g., Powell, Koput, and Smith-Doerr, 1996), and some claim to analyze neither but instead the type of relationship (Powell et al., 2005).

In order to test our hypotheses, we use the venture capitalist - startup dyad as a unit of analysis, and we compare the realized ties from our base

¹Multiple ties between a venture capitalist and the same startup may exist if the investor invested in multiple funding rounds.

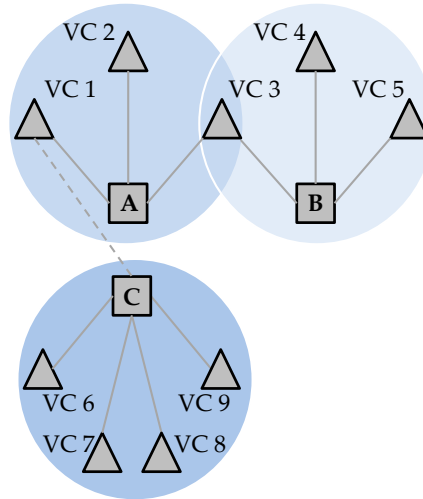


FIGURE 5.1: Realized and unrealized ties in an investment network.

dataset to a set of unrealized ties. This dyad-level approach has two important characteristics. Firstly, there is a mutual interest in tie formation, in our case, the exchange of social and financial capital for an equity stake in the new venture (Sorenson and Stuart, 2008). Secondly, based on certain characteristics, we predict in which type of startup a venture capitalist invests (Gulati, 1995).

As previously outlined, venture capitalists who have sufficient financial resources have a great number of startups in which they could invest. A large share of their daily work consists of scanning new ventures and pitching books. From this selection, and after rigorous due diligence, they enter into a financing round with selected startups who also consider the focal venture capitalist to be a useful investor. Naturally, we thus have a great number of non-materialized ties (not yet included in our dataset) and a far smaller number of realized ties. Some studies have analyzed all unrealized and realized dyads, such as that of Podolny (1994), who analyzed investment-grade and non-investment-grade investment banks and how lead banks choose their co-managers. As only 170 banks in the investment grade and 171 in the non-investment grade sample exist, it is likely that a lead bank would consider all or at least a great share of all possible banks as potential co-managers. The same is true for other researchers, such as Gulati (1995), who analyzed tie formation in certain industries, with up to 62 companies in his sample, or Stuart (1998), who analyzed up to 144 firms per observation period. The assumption underlying these studies is that there is a non-zero probability that ties between these firms are realized. However, it is very unlikely that a typical venture capitalist reviews all possible investments in a given quarter. In our sample, we have 10,576 funding rounds (a multiple of that number if we were also to include the

first-time and second-time investment of all venture capitalists, as well as the funding rounds and companies that we excluded due to missing data) in 44 quarters, rendering about 240 potential investment opportunities to consider per quarter. It would thus be false to assume a non-zero probability of all possible dyads in our sample.

In addition, considering all possible dyads creates a methodological problem of non-independence so that standard errors are underestimated (Sorenson and Stuart, 2001). Attributes (such as venture-capitalist-level attributes) that do not change with the unrealized ties may report lower standard errors. While, for a large dataset such as ours, it would also be highly computationally intensive to consider all possible dyads (Sorenson and Stuart, 2001), with the use of cloud-based server power, it is possible to do so nowadays. Computational intensity thus has not influenced our choice for the empirical approach, which is described in the following paragraph.

In order to make our approach clear, we created a schematic representation of three funding rounds in figure 5.1, including the syndication networks. Each circle represents one funding round of the startups A, B, and C. The triangles represent the venture capitalists and the continuous lines each represent a realized tie or a dyad between a venture capitalist and a startup. Only these realized ties are included in our base dataset. We have then created one unrealized tie for every realized tie, so that each investor is matched to a startup that has received funding in the same quarter, but in which the investor did not invest. We have randomly sampled the startup for the unrealized tie from the list of startups that received funding in that quarter, excluding the startups that the focal investor invested in during the quarter, for example, if he or she had made multiple investments. We used the sampling function in base R, which draws upon the random number generator for setting the seed. Taking a startup that received funding in that quarter ensures, at least theoretically, that the dyad could have been realized. By doing so, we also indirectly control for the quality of the startup, as it indicates that it is worth a venture capital investment. In figure 5.1, the dotted line between venture capitalist 1 and startup C represents an exemplary unrealized tie. Our resulting dataset for the regression now includes the dependent variable *dummy actual data*, which equals 1 in the case of a realized tie and 0 otherwise. Naturally, half of our observations are coded as 1.

This approach is superior to sampling a number of unrealized investment ties randomly, because most of the information for the model estimation is included in the realized investments (Imbens, 1992; Imbens and Lancaster, 1996; King and Zeng, 2001). The advantages of choosing only one unrealized tie over all other possible ties have been outlined above.

For each unrealized tie, some of the variables remain the same as in

Variable	Status in non-realized ties
<i>Independent variables</i>	
Startup's distinctiveness	Updated
Portfolio distinctiveness	Unchanged
Portfolio diversification	Unchanged
Venture capitalist's status	Unchanged
Categorical distance	Updated
<i>Startup/funding-round level controls</i>	
Startup location (3 dummies)	Updated
Age at funding	Updated
Investor count (26 dummies)	Updated
Raised amount	Updated
Round number	Updated
Year (10 dummies)	Unchanged
<i>Venture capitalist/syndicate level controls</i>	
VC age	Unchanged
General experience	Unchanged
Industry experience	Updated
Mean affiliation	Unchanged
<i>Venture capitalist - startup dyad level controls</i>	
Geographical distance	Updated
Industry distance	Updated
Prior investment (dummy)	Updated
Affiliate geographical distance	Updated
Affiliate industry distance	Updated

TABLE 5.1: Overview of the creation of variables in unrealized ties

the corresponding realized tie, whereas others are changed. For example, investor-level variables such as status remain the same, whereas the distinctiveness of the startup changes. Table 5.1 shows a complete overview of the variables and whether they are updated or remain the same as in the corresponding realized tie.

5.2 Analytical method: logistic regression and rare-events application

In this section, we broadly describe our statistical approach, including the main characteristics of logistic-regression models and the peculiarities in rare-events situations. For the basic statistical concepts, we have used the overview of regression modeling by Harrell (2015), which has been found to be an extensive standard work with a high quality reference section (Helmreich, 2016).

5.2.1 Logistic regression

The choice of a statistical model depends on the nature of the dependent variable. As described in section 4.3, our dependent variable is binary and takes the value of 1 if an investors has invested in the focal startup and 0 in all other cases. Our base statistical model is thus a logistic-regression model, which is suitable for binary dependent variables (Harrell, 2015).²

We denote the binary dependent variable by Y and build a model that predicts the value of Y based on a number of predictor variables X_k in vector $X = \{X_1, X_2, \dots, X_k\}$. Furthermore, we denote the vector of regression coefficients by $\beta = \{\beta_1, \beta_2, \dots, \beta_k\}$. As the logistic regression is a special case in the ordinary linear model, we begin with the following linear model:

$$E\{Y|X\} = X\beta \quad \text{with} \quad X\beta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (5.1)$$

Equation 5.1 can take values that fall outside of the range of our binary variable. We thus adapt the equation in order to predict the probability that Y takes the value of 1 given the values of the independent variables X . Based on the work of Cox (1958) and Walker and Duncan (1967), the probability is given by as follows:

$$E\{Y|X\} = \text{Prob}\{Y = 1|X\} = [1 + \exp(-X\beta)]^{-1} \quad (5.2)$$

The following equation thus shows the so-called logistic regression function:

$$P = [1 + \exp(-x)]^{-1} \quad (5.3)$$

The function in equation 5.3 now has a range from 0 to 1 and can take any value from our set of predictors X , as illustrated graphically in figure 5.2, which shows equation 5.3 with x ranging from -5 to $+5$.

Model assumptions – An important difference between the logistic-regression and ordinary multiple-regression models is that the former lacks a normality requirement for the dependent variable. This is true as long as the observations are independent of each other and simple random sampling is used (Harrell, 2015). In addition, the general assumption that all relevant co-variables are included needs to be true.

²Another widespread model by which to describe a binary dependent variable is the probit model (e.g., see application in Stuart, 1998). It defines the probability of Y being equal to 1 by the cumulative normal distribution and discriminant analysis (Harrell, 2015). While the shape of the resulting function is similar to that of the logistic regression, it has a number of disadvantages. According to Harrell (2015), the probit model is computationally more complicated, the regression coefficients have no natural interpretation, and more assumptions are made. We have thus chosen to use logistic regression.

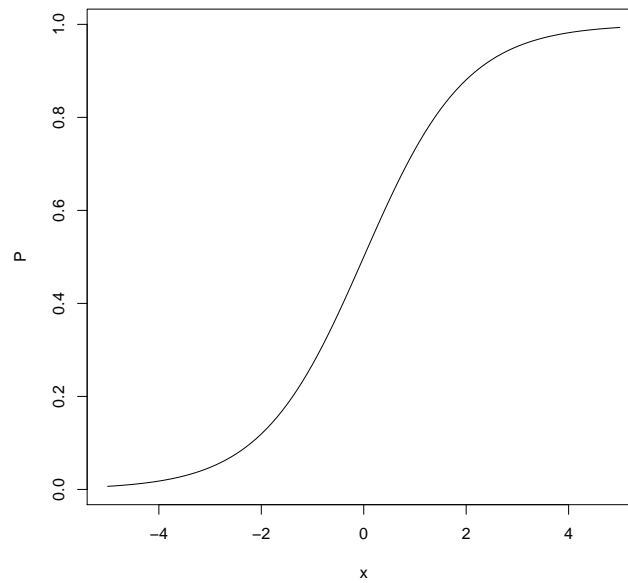


FIGURE 5.2: Plot of logistic function

Model coefficient interpretation – Solving equation 5.3 for x allows us to better understand the regression and to interpret the coefficients. After rearranging the equation, x can be considered to be the natural logarithm of the odds that $Y = 1$:

$$\begin{aligned}
 P &= \frac{1}{1 + \exp(-x)} \\
 \exp(-x) &= \frac{1}{P} - 1 \\
 \exp(x) &= \frac{P}{1 - P} \\
 x &= \log \left[\frac{P}{1 - P} \right]
 \end{aligned} \tag{5.4}$$

It is important to note that $P/(1 - P)$ are the odds and not the probabilities that $Y = 1$. The equation can thus take any value and is not bounded, as is the case with probabilities. Having solved for x in equation 5.4 allows us to build a linear model based on the log odds that $Y = 1$. In the case of no interaction effects, the so-called 'logit' of the logistics regression is defined as follows, with X_k being the factors in the model:

$$\begin{aligned}
 \text{logit}(Y = 1|X) &= \log \left[\frac{P}{1 - P} \right] \\
 &= X\beta \\
 &= \beta_0 + \beta_1 X_1 + \dots + \beta_j X_j + \dots + \beta_k X_k
 \end{aligned} \tag{5.5}$$

In order to interpret the coefficient β_j , we assume that all factors except for X_j in equation 5.5 are held constant. β_j is then the linear change of the log odds when X_j changes by one unit. We can rewrite this to attain the non-log change in the odds that $Y = 1$ by the following equation:

$$\Delta odds(Y = 1|X) = \exp(\beta_j X_j) \quad (5.6)$$

Generally, the odds can be defined as follows, where C is defined as all factors except for j , given by $C = \beta_0 + \beta_1 X_1 + \dots + \beta_{j-1} X_{j-1} + \beta_{j+1} X_{j+1} + \dots + \beta_k X_k$.

$$\begin{aligned} odds(Y = 1|X) &= \exp(X\beta) \\ &= \exp(\beta_j X_j) \exp(C) \end{aligned} \quad (5.7)$$

The use of logistic regression thus allows us to interpret the coefficients of the regression by transforming them into odds changes. In many cases, it is easier to interpret probabilities than odds. Returning to the specific case from equation 5.6, we can obtain the change of the outcome probability when X_j changes by one unit through rearrangement according to equation 5.4:

$$\begin{aligned} \Delta odds(Y = 1|X) &= \exp(\beta_j X_j) \\ \Delta \frac{P}{1-P} &= \exp(\beta_j X_j) \\ \Delta P &= \frac{\exp(\beta_j X_j)}{\exp(\beta_j X_j) + 1} \\ \Delta P &= \frac{1}{1 + \exp(-\beta_j X_j)} \end{aligned} \quad (5.8)$$

Interaction effects – Interaction effects describe situations in which the effect of one predictor is affected by the magnitude of a second predictor. For instance, the body height of a person is positively related to a person's weight. The effect of body height, however, is different for males and females, which the interaction effect would show. In our analysis, for example, we expect the effect of categorical distance to be mitigated by a venture capitalist's status. The latter is the second predictor, which affects the effect size of the categorical distance on the investment probability. The effects of each of the predictors depend on each other. In a simple model, we then denote the interaction term of two independent variables X_1 and X_2 by $X_3 = X_1 X_2$ so that

$$C(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 \quad (5.9)$$

The interpretation of the effect of a change in X_1 is more complicated than in cases of purely separate effects. Assuming that X_2 remains constant, a change of one unit in X_1 is the first derivation of equation 5.9 and has an effect that is described as follows:

$$C(Y|X_1 + 1, X_2) - C(Y|X_1, X_2) = \beta_1 + \beta_3 X_2 \quad (5.10)$$

It is important to note that adding an additional predictor to the model as an interaction term not only makes the additional predictor dependent on the original predictor, but also vice versa.

Parameter estimation – In logistic regression, parameters β are estimated based on the maximum-likelihood method. Only in special cases can one explicitly solve for the parameter β . Instead of this, a maximization method is used to iteratively find the best solution for β in order to maximize the log likelihood. One of the most widespread methods is the Newton-Raphson method (Harrell, 2015): a stepwise approximation until the change in β does not change by more than a small, previously defined amount. This amount is the threshold by which the statistical significance is not further affected on the χ^2 scale. A more detailed review of the method's development can be found in Ypma (1995), and more practically interested readers can refer to Harrell (2015).

Test statistics – In logistic regression, the Wald statistic is typically used to obtain the statistical significance that coefficients are different from zero, that is, there is an effect of the predictor on the dependent variable on a certain probability level. In order to test whether the predictors increase the power of the model, a test statistic may be used. Either the likelihood-ratio method or the Wald statistic is used, but the latter does not perform well for cases in which the sample is small and the parameter's true value is significantly different from the null value (Harrell, 2015). In the likelihood-ratio method, the likelihoods of two models' test statistics are compared and whether the improvement³ of the likelihood is statistically significant is then tested. Another test statistic is the Lagrange multiplier, which is the slope of the log likelihood normalized by the curvature of the log likelihood.

5.2.2 Peculiarities in rare-events settings

Examples of rare events are wars, earthquakes, or atomic-core melt-downs that only occur rarely when analyzing historical data. For instance, analyzing international relations between country pairs since 1945 only has a 0.34% defined occurrence of a country pair being at war (King and Zeng,

³Typically, as a predictor is added to the model, the likelihood is improved but that does not necessarily mean that it is a better fit.

2001). Similarly, considering all the potential investments of a venture capitalist, such as the number of new ventures that approach a venture capitalist or try to make a pitch appointment, the actual investment in the part of an investor is a rare event. However, as previously indicated, most of the information can be found in the event data and not in the non-event data.

Working with these rare events has two peculiarities. Data collection can be very intensive with a limited amount of insight gained, and, if the dependent variable is binary, as in logistics regression, the event probabilities may be downwardly biased (King and Zeng, 2001). Regarding the data collection, the suggestion is to collect as much data as possible on the realized events and then sample randomly a number of zeros, that is, unrealized events or ties. This strategy effectively focuses the researchers' resources on the most informative data. Using this strategy for the dataset, it is necessary to correct the event probabilities, which we do by applying the approach developed by King and Zeng (2001). This correction has been widely used in tie-formation research in venture capital settings (Sorenson and Stuart, 2001; Trapido, 2007; Rider, 2012; Zhelyazkov and Gulati, 2016) and other research areas (e.g., Jensen, 2003).

From equation 5.2, we know that the logistic regression renders the probability of $Y = 1$ given X by the following equation:

$$E\{Y|X\} = \text{Prob}\{Y = 1|X\} = [1 + \exp(-X\beta)]^{-1} \quad (5.11)$$

In order to correct for the effects described above, King and Zeng (2001) show that β is biased by oversampling rare events in the following equation, with $\xi = 0.5Q_{ii}[(1 + w_1)\text{Prob}\{Y = 1|X\} - w_1]$ in which Q_{ii} are the diagonal elements of $Q = X(X'WX)X'$ and $W = \text{diag}\{\text{Prob}\{Y = 1|X\}(1 - \text{Prob}\{Y = 1|X\})w_i\}$. w_1 is the share of realized ties in the sample relative to the share in the population and represents the weights.

$$\text{bias}(\beta) = \frac{X'W\xi}{X'WX} \quad (5.12)$$

The estimation of β is then done by using weighted least-square regression with X as the predictor and ξ as the response variable. The unbiased β is then given as follows:

$$\beta_{\text{unbiased}} = \beta - \text{bias}(\beta). \quad (5.13)$$

In our analysis, we used the corrected β to remove any bias, and we used the implementation in the R package *Zelig* (Choirat et al., 2016) to run our models with the correct β estimates.

5.2.3 Imputation for missing data

Imputation means dealing with missing data, because, in many cases, data is incomplete, for instance, when survey participants skip questions or databases are not well-maintained. The following brief overview of different methods for dealing with missing data is based on the comprehensive review by Schafer and Graham (2002). One of the simplest methods is the 'case deletion' method, from which all observations with missing data are excluded, but it can lead to biases when data is not missing totally randomly, that is, independent of the remaining data, or when the number of observations is too low after deletion of observations. The result would be a dataset that does not represent the complete population. One way to mitigate this problem is a weighting of the variables in the remaining dataset to be more representative of the complete population. Another method of dealing with missing data is replacing it with possible data, the so-called 'single imputation', to keep all observations, which increases the sample size and keeps the information from other variables in the observations. Different approaches for this replacement of missing values exist. For instance, replacement values can be randomly drawn from the observed values to keep the distribution similar. Another alternative is to use the mean of the observed values to keep the overall mean constant, but this approach changes the distribution by reducing the standard error. Both methods can accommodate cases in which the focal variable is conditional on another variable. However, single imputation generally reduces the precision of the models and may give significantly biased results. Schafer and Graham (2002) show that, in cases of low shares of missing data, for example, 3%, single imputation can be reasonable and is preferred over deleting all the observations with missing values.

More statistically robust results produce two alternative methods. 'Multiple imputation', developed by Rubin (1987), is similar to single imputation, but, instead of creating a single dataset with missing values inserted from a certain distribution, it creates multiple complete datasets. These multiple datasets are then individually analyzed and the results are consolidated, which increases the workload, but the computation is intellectually easy (Rubin, 1996). The second alternative method is using a maximum likelihood estimation, which assumes a normal distribution and requires large sample sizes (Schafer and Graham, 2002). Based on the observed values, a maximum likelihood estimate for the unobserved values is made, that is, finding the greatest probability that these were the observed values (Newman, 2003). Today, most statistical software provides functionality to apply both methods.

For most of the missing data in our *Crunchbase* dataset, we use the case-deletion method for two reasons. Firstly, we expect the missing values to

be missing randomly due to the insufficient maintenance of a database to which anyone can contribute. Furthermore, imputation is, by definition, imperfect and will never lead to a perfect result if the missing data is not missing randomly. We do not believe that, in our case, the imputation would have improved our results compared to deletion. In addition, we do not have missing data for our independent variables, as we have excluded the companies without a category entry. Imputation of categorical values with a high number of categorical values is particularly difficult, especially when the number of categories is unknown, for instance, startups are categorized in up to 17 categories per startup in our dataset. The other variables, for example, the founding date or funding amount, are only control variables and are thus of lower relevance. The size of our dataset also remains large enough to have great predictive power. Only for geographical distance we have used single imputation. The reason why we have missing data in the geographical distance is an issue concerning matching data from two different Excel sheets in the Crunchbase dataset. To avoid dropping the observations, we have used mean imputation for geographical distance and included a dummy variable equal to 1 when we inserted the mean geographical distance and 0 otherwise. In our dataset, only about 3.2% of the geographical distance values are missing, which is reasonably small to not overly introduce a bias for that variable.

5.3 Specification of statistical models

We use logistic regression with rare-events corrected coefficients, as described above, in order to test our hypotheses. Our baseline model 0 only includes the control variables. In model 1, we add the four independent variables of the first part of the analysis, that is, the startup's distinctiveness, portfolio distinctiveness, portfolio diversification, and status.

In the following three models, we use the interaction effects between our three venture-capital-based moderators and the startup's distinctiveness, which allows us to test our expected effects. In order to incorporate the effects of the variables on the intercept, we also need to include the main effects (e.g., see application in Sorenson and Stuart, 2001). Models 3, 4, and 5 thus each include the main effects and the interaction effect between the startup's distinctiveness and portfolio distinctiveness, portfolio diversification, and the venture capitalist's status. Model 5 is the full model for part one, including all main effects and the interaction effects of the independent variable with all three moderators.

The second set of models adds the categorical distances. Model 6 is the same as model 1, complemented by the categorical distance to test the hypothesis on the main effect of categorical distance. Models 7, 8, and 9

include each one of the venture capitalist's attribute variables and the interaction effects with categorical distance to test our hypothesis. The last model is the full model for the second part of the analysis, including all three interaction effects between categorical distance and portfolio distinctiveness, portfolio diversification, and status.

Chapter 6

Results

This chapter has three parts. Firstly, we describe our variables highlighting their peculiarities to develop a better understanding for the dataset, for example, the distributions or types of startups and venture capitalists. Secondly, we present the results of the regression models and what they mean for the test of our hypothesis. We discuss the results in detail in chapter 7. Thirdly, we use multi-way clustering by the startup and the venture capitalist to test the robustness of our results.

6.1 Descriptive statistics

We split the discussion of descriptive statistics into a presentation of the behavior of the five independent variables and a more detailed description of the dataset based on the control variables and additional information. Table 6.1 displays the main descriptive statistics, and table 6.2 presents the bivariate correlations for the sample. We have excluded the year dummies, the funding round dummies, and the investor count dummies from these tables to avoid overloading them. Instead, we have included the investor count and round number variables as discrete variables for informational purposes. A breakdown of our observations by year is displayed in figure 6.5. For all analyses, we have used 58,000 venture capitalist - startup dyad observations.

6.1.1 Overview of independent variables

The startup's distinctiveness ranges from 0 to 0.912 with a mean of 0.480 and due to the way we calculate this measure, it is not normally distributed. We assign a value of 0 to a startup's distinctiveness if the startup is associated with only one category as it is not boundary-spanning in that case. This results in the non-normal distribution. Figure 6.1 displays the histograms and density plots with the corresponding normal curves for each of the five independent variables, and graphical examination shows that neither of the variables is normally distributed.

Statistic	N	Mean	St. Dev.	Min	Max
Dependent variable					
Dummy actual data	58,000	0.500	0.500	0	1
Independent variables					
Startup's distinctiveness	58,000	0.480	0.178	0.000	0.912
Portfolio distinctiveness	58,000	0.479	0.081	0.000	0.912
Portfolio diversification	58,000	19.321	5.135	3.349	41.447
Venture capitalist's status	58,000	2.934	2.819	0.018	13.561
Categorical distance	58,000	20.563	6.868	3.609	41.447
Startup/funding round level controls					
State dummy CA	58,000	0.518	0.500	0	1
State dummy MA	58,000	0.098	0.298	0	1
State dummy NY	58,000	0.135	0.342	0	1
Startup age	58,000	41.174	25.737	0	129
Investor count	58,000	4.307	3.091	1	32
Raised amount (log)	58,000	15.954	1.248	8.517	20.906
Round number	58,000	1.987	1.239	1	10
Venture capitalist/syndicate level controls					
VC age	58,000	9.118	4.940	1.083	29.917
General experience	58,000	69.860	94.302	2	624
Industry experience	58,000	1.333	5.383	0	91
Mean affiliation	58,000	4.785	10.216	0.000	179.000
Venture capitalist-startup dyad level controls					
Geographical distance (log)	58,000	6.029	2.898	0.000	9.821
No geographical distance dummy	58,000	0.032	0.176	0	1
Industry distance	58,000	0.776	0.198	0.000	1.000
Prior investment dummy	58,000	0.175	0.380	0	1
Affiliate geographical distance	58,000	13.384	34.049	0.000	100.000
Affiliate industry distance	58,000	0.839	0.234	0.000	1.000

TABLE 6.1: Descriptive statistics

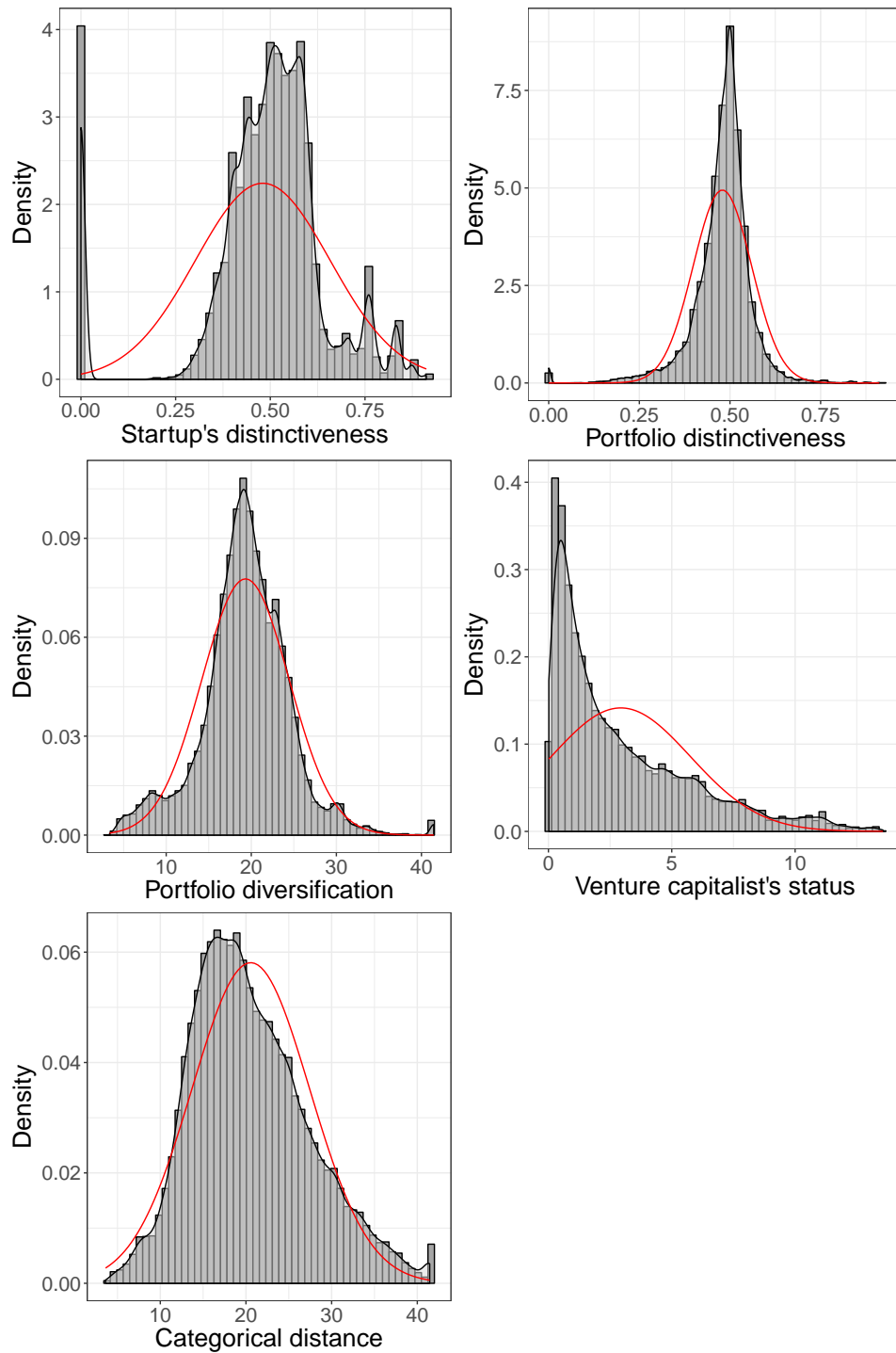


FIGURE 6.1: Histograms and density plots with respective normal curves

	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7	Var 8	Var 9	Var 10	Var 11	Var 12	Var 13	Var 14	Var 15	Var 16	Var 17	Var 18	Var 19	
1. Startup's distinctiveness	1																			
2. Portfolio distinctiveness	0.058	1																		
3. Portfolio diversification	0.022	0.123	1																	
4. Venture capitalist's status	0.023	0.132	0.141	1																
5. Categorical distance	0.026	0.033	0.347	0.034	1															
6. Startup age	0.211	0.025	-0.037	-0.029	0.002	1														
7. Investor count	0.085	0.004	-0.047	-0.001	-0.093	0.010	1													
8. Raised amount (log)	0.240	0.086	-0.026	0.081	-0.033	0.268	0.270	1												
9. Round number	0.211	0.061	-0.008	0.025	-0.048	0.530	0.104	0.402	1											
10. Geographical distance (log)	-0.003	-0.006	-0.029	-0.137	0.061	0.062	-0.059	-0.007	0.006	1										
11. No geographical distance dummy	0.002	-0.005	-0.082	-0.123	-0.017	0.037	0.006	0.022	0.024	0.118	1									
12. Industry distance	-0.075	0.035	0.241	0.113	0.513	-0.085	-0.070	-0.124	-0.117	0.046	-0.059	1								
13. Prior investment dummy	0.094	0.038	0.004	0.076	-0.106	0.197	0.084	0.189	0.374	-0.144	-0.012	-0.262	1							
14. Mean affiliation	0.050	0.043	-0.011	0.371	-0.105	0.024	0.148	0.165	0.127	-0.177	-0.047	-0.084	0.313	1						
15. Affiliate geographical distance	0.027	-0.008	-0.030	0.088	-0.103	-0.027	0.321	0.076	0.027	-0.145	-0.018	-0.071	0.151	0.259	1					
16. Affiliate industry distance	-0.072	-0.008	0.086	-0.110	0.251	-0.120	-0.328	-0.250	-0.209	0.134	0.004	0.465	-0.425	-0.359	-0.333	1				
17. VC age	0.038	0.150	0.132	0.517	0.031	0.108	-0.048	0.100	0.102	-0.067	-0.057	0.054	0.116	0.189	0.008	-0.086	1			
18. General experience	0.030	0.139	0.153	0.862	0.033	0.043	-0.021	0.096	0.062	-0.105	-0.101	0.087	0.102	0.335	0.053	-0.106	0.665	1		
19. Industry experience	-0.048	-0.031	-0.099	0.122	-0.140	0.096	0.017	0.129	0.081	-0.038	0.006	-0.336	0.200	0.129	0.039	-0.339	0.148	0.166	1	

TABLE 6.2: Pairwise correlation matrix for main variables

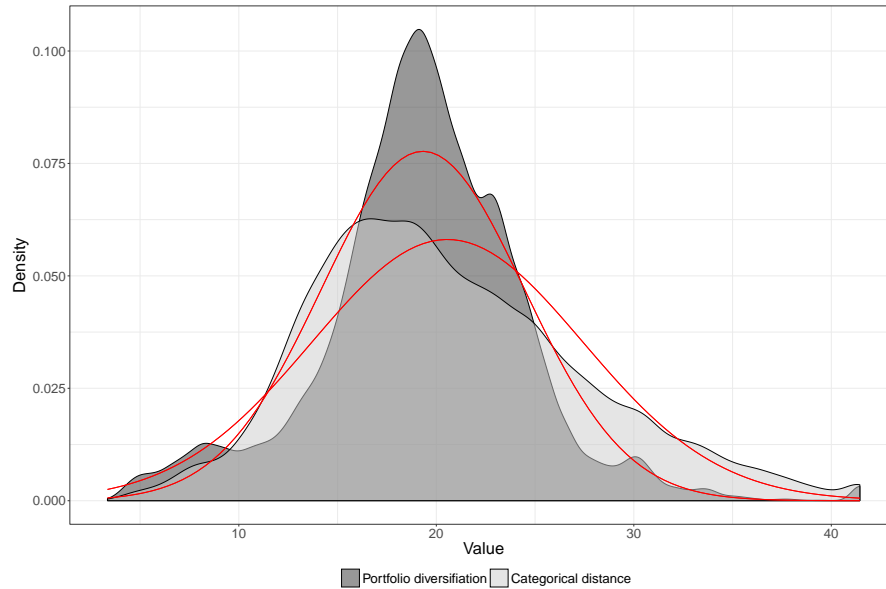


FIGURE 6.2: Comparison of portfolio diversification and categorical distance based on density plots and normal curves

The distribution and range of portfolio diversification and categorical distance are similar because we use the same calculation method for the two measures. We also acknowledge that there is a correlation between the two measures that is higher than any other correlation in our dataset (refer to table 6.2 for details). This, naturally, makes sense because, if a portfolio is highly diversified, that is, the categorical distance between the portfolio companies is large, then it is probable that the distance between the focal startup and the portfolio companies is also large. Only if the focal startup is more category-spanning than the portfolio companies can the categorical distance be lower than the diversification value. The comparison of the density charts in figure 6.2 shows that despite the similarity, the distribution of categorical distance is more widespread than that of portfolio diversification.

Furthermore, we are interested in understanding the relationship between the three attributes of the venture capitalist, that is, portfolio distinctiveness, portfolio diversification, and status. While we have not hypothesized about the relationships, it might still be helpful in the discussion of the results of the hypothesis tests to present these here. We have thus created a three-dimensional scatter plot in figure 6.3. The graphical inspection shows that, with increasing status, the deviations from the mean of portfolio distinctiveness and diversification are reduced. We have created four status-based quartiles, by splitting the status-ordered observations into four equally sized subsets, to analyze the distribution of portfolio

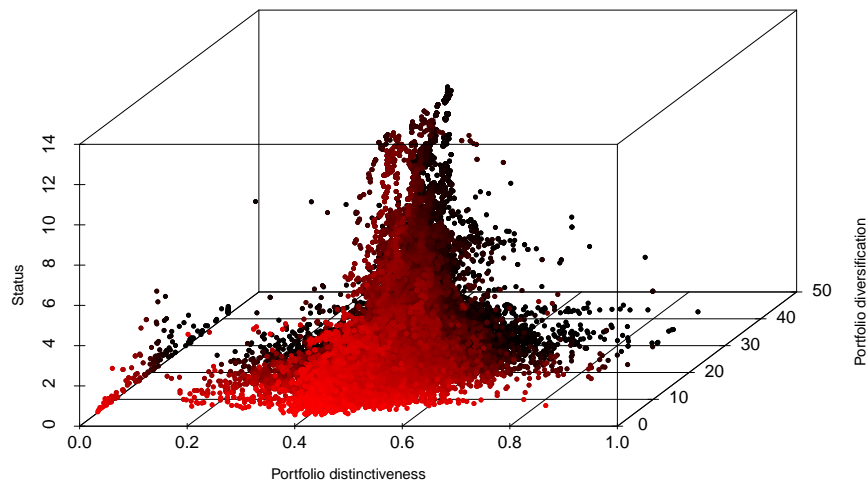


FIGURE 6.3: Three dimensional scatter plot of the three attributes of the venture capitalist

distinctiveness and diversification by status quartile in figure 6.4. The figure clearly shows that, in our sample, status reduces deviations from the mean for portfolio distinctiveness and portfolio diversification.

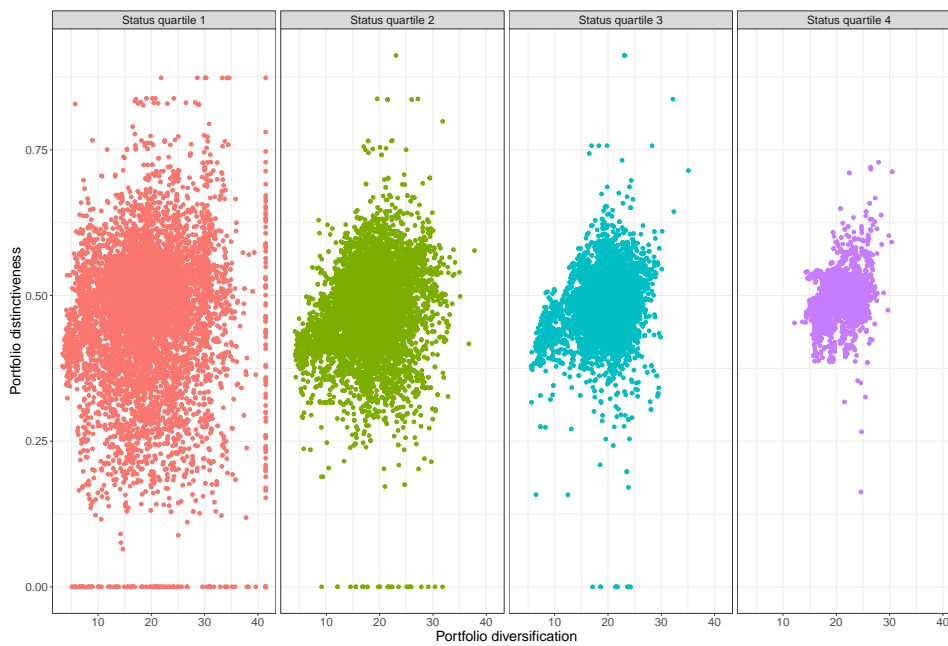


FIGURE 6.4: Analysis by status quartiles

6.1.2 The dataset

After the application of filters as described in section 4.2, the base dataset includes 10,576 funding rounds of 5,826 startups from 2,323 venture capitalists, that led to 29,000 venture capitalist - startup observations. Through the addition of unrealized ties, we doubled the sample size to 58,000 venture capitalist - startup observations. The startups were associated with one or more of 538 different detailed categories and 44 unique group categories. Table 6.3 summarizes this. The dataset is skewed to more recent years, as the Crunchbase platform and the related technology blog TechCrunch have increased their reach and attracted more contributors to the database during this time. We take this into account by controlling for the year of the funding round. Figure 6.5 splits the venture capitalist - startup dyads and the funding rounds by year.

Information	Number
Startups	5,826
Funding rounds	10,576
Investors	2,323
Unique categories	538
Unique group categories	44

TABLE 6.3: Overview core information final dataset

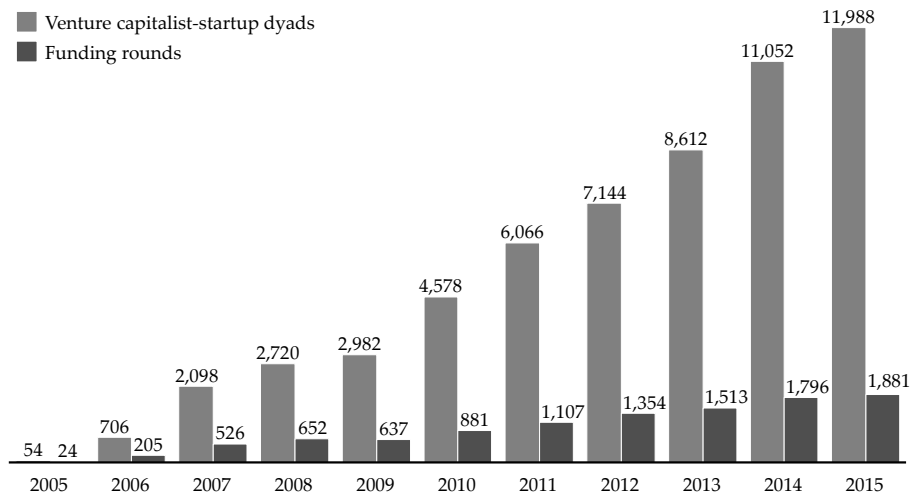


FIGURE 6.5: Frequency distribution by year

More than half (54%) of the new ventures in our sample enter it with only one funding round, and another quarter has two funding rounds in our sample. The startup might have previous funding rounds that are not labeled venture capital, for example, seed financing, but we excluded these

rounds from our sample. The pie chart in figure 6.6 displays the distribution of funding rounds per new venture.

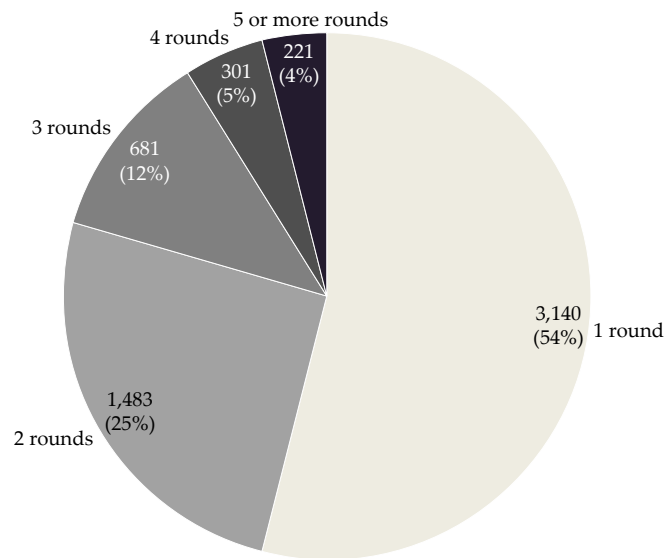


FIGURE 6.6: Frequency distribution of funding rounds per new venture

The geographical distribution of startups in figure 6.7 shows a concentration in the three states of California, New York, and Massachusetts, which is why we control for these states with three dummy variables. The figure excludes 14 startups for which we lacked information on their headquarters' location and two startups that have indicated headquarters on the islands of Hawaii. For graphical reasons the plot shows the point size relative to the log of the number of startups in the respective city.

6.2 Model results

In line with our hypotheses developed in chapter 2, we split our analysis into two parts. The first set of regressions helps to find evidence for the hypothesis regarding the startup's distinctiveness and the venture capitalist's varying preference for it. The second set of regressions displays the empirical results for the hypothesis tests regarding categorical distance. Tables 6.4 and 6.5 display the results of the rare-events logistic regressions. Three significance levels are indicated in the tables and used throughout this paper: weakly significant ($p < 0.05$), significant ($p < 0.01$), and strongly significant ($p < 0.001$). For an easier interpretation of the results, figure 6.8 includes the effect plots of the interaction terms. The figure plots the effects on a startup's distinctiveness and categorical distance for each of the venture capitalist's three attributes, which are set to two standard deviations above the mean and two standard deviations below the mean, as well as

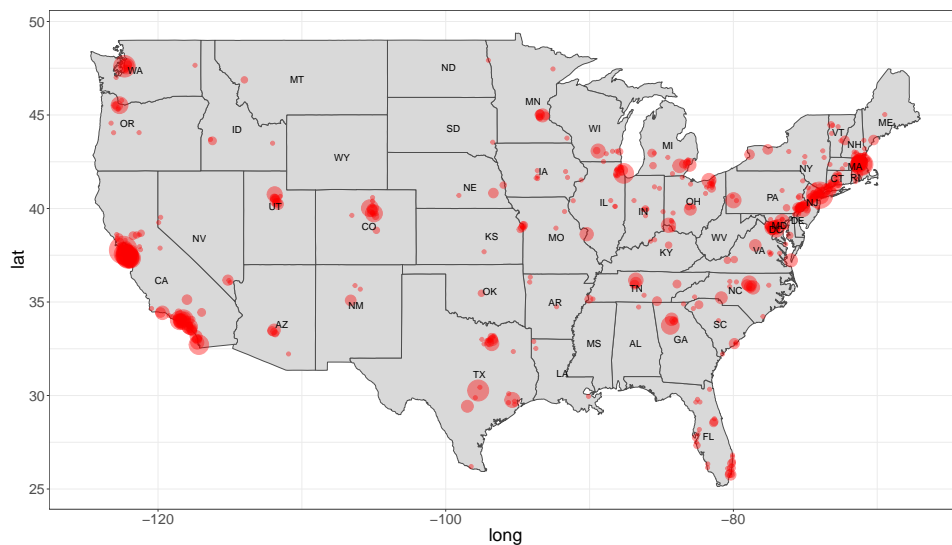


FIGURE 6.7: Geographic distribution of startups

the mean value. The bands around the plotted lines show the 95% confidence interval. We use logistic regression results without correction for rare events for the graphical display. A more detailed version of the regression outputs, including the coefficients for the dummy variables and a full model that includes all the variables and interaction effects of the two analyses can be found in appendix A.

6.2.1 Analysis part I: startup's distinctiveness

Model 0 only reports the estimated log odds for the control variables and serves as a baseline. The results for the major control variables are as expected. The greater the industry distance between a startup and an investor, the lower the probability of an investment is. The coefficient for geographical distance is also as expected, negative and strongly significant, reducing the probability of an investment. The same is true for our two measures of affiliate distance: The greater the industry and geographical distance of the closest syndicate partner, the lower the investment probability. It is important to note that the affiliate geographical distance is defined as the inverse of geographical distance, so that a positive value in the regression model indicates lower probability with larger distance. Interestingly, the effect on log odds is significantly larger for affiliate industry distance than for the direct industry distance, indicating that the syndicate distance is even more important than the direct distance. Furthermore, an

investor who has previously invested in the startup increases the probability of an investment with strong significance. We observe the largest coefficient for this control variable. Control variables that are only dependent on either the venture capitalist, for example, age or industry experience, or only dependent on the startup, for example, amount raised or geographical location, cannot be directly interpreted in our research design. As all of the startups in our sample received funding in that quarter, it would be wrong to interpret the coefficients as the log odds of receiving funding. We can only derive conclusions for variables that affect the venture capitalist - startup dyad. In other words, only the main effects of variables that simultaneously depend on the venture capitalist and the startup, such as any distance measure, can be meaningfully interpreted. That is also the reason why we have not included a hypothesis on the main effect of the startup's distinctiveness on the probability of receiving funding.

In model 1 we added the four independent variables from the first part of our analysis: the startup's distinctiveness, the portfolio distinctiveness, the portfolio diversification, and the venture capitalist's status. As outlined above, we cannot interpret the coefficients, as they are affected by the distribution of our sample but do not describe a dyad characteristic. With the following models, we test which investors prefer a startup's distinctiveness.

Model 2 incorporates the venture capitalist's historical average portfolio distinctiveness and the interaction effect with the startup's distinctiveness. As expected, we found strongly significant evidence for our hypothesis that venture capitalists who have previously invested in distinctive startups and gained experience in managing the associated risks have a higher probability of investing in a distinctive startup. Figure 6.8a clearly supports this hypothesis. In the plot, the solid line depicts the relationship of a startup's distinctiveness with the probability of investment in the case of portfolio distinctiveness two standard deviations below the mean portfolio distinctiveness. The line's negative slope is interpreted as a decreasing probability of investment when the a startup's distinctiveness increases for venture capitalists with low portfolio distinctiveness. The dotted line for two standard deviations above the mean of portfolio distinctiveness is upward sloping, which means that, for venture capitalists with high portfolio distinctiveness, the probability of investment increases with an increase in a startup's distinctiveness.

Model 3 adds to model 1 the portfolio diversification of the focal venture capitalist and its interaction effect with the startup's distinctiveness. We have expected a negative effect on the probability of investing in a distinctive startup, but we cannot confirm this expectation. Rather, we observe a weakly significant and positive coefficient of the interaction effect indicating that investors who try to diversify their portfolios in order to reduce risk

prefer distinctive new ventures. The graph in figure 6.8b supports this finding with an upward-sloping dotted line for highly diversified investors and a downward-sloping solid line for venture capitalists with lowly diversified portfolios.

In model 4, we have added the status measure based on venture capitalists' historical syndicate networks and the respective interaction effect. We find weakly significant support for the hypothesis that with the increasing centrality of a venture capitalist, the appetite for distinctive startups increases. Again, the slopes in figure 6.8c show graphical support.

In model 5 we introduce all the interaction effects for our analysis of startup's distinctiveness, the portfolio characteristics, and the venture capitalist's status. The direction of the interaction effects' coefficients remains the same but the previously weakly significant results for the interaction with portfolio diversification and status are no longer significant at the 5% level.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	4.99*** (0.35)	4.91*** (0.36)	5.68*** (0.38)	4.53*** (0.37)	5.59*** (0.35)	5.62*** (0.40)
State dummy CA	-0.26*** (0.03)	-0.25*** (0.03)	-0.26*** (0.03)	-0.26*** (0.03)	-0.25*** (0.03)	-0.25*** (0.03)
State dummy MA	-0.23*** (0.04)	-0.22*** (0.04)	-0.23*** (0.04)	-0.22*** (0.04)	-0.23*** (0.04)	-0.22*** (0.04)
State dummy NY	-0.15*** (0.04)	-0.13*** (0.04)	-0.15*** (0.04)	-0.14*** (0.04)	-0.14*** (0.04)	-0.13*** (0.04)
Startup age	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Raised amount (log)	0.06*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)
Geographical distance (log)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)
Geographical distance no data dummy	0.36*** (0.06)	0.35*** (0.06)	0.36*** (0.06)	0.40*** (0.06)	0.31*** (0.06)	0.35*** (0.06)
Industry distance	-1.52*** (0.07)	-1.61*** (0.07)	-1.53*** (0.07)	-1.68*** (0.07)	-1.46*** (0.07)	-1.60*** (0.07)
Prior investment dummy	5.84*** (0.17)	5.80*** (0.17)	5.83*** (0.17)	5.83*** (0.17)	5.81*** (0.17)	5.79*** (0.17)
Mean affiliation	0.04*** (0.00)	0.05*** (0.00)	0.04*** (0.00)	0.04*** (0.00)	0.05*** (0.00)	0.05*** (0.00)
Affiliate geographical distance	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Affiliate industry distance	-3.73*** (0.08)	-3.78*** (0.09)	-3.73*** (0.08)	-3.70*** (0.08)	-3.81*** (0.09)	-3.78*** (0.09)
VC age	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
General experience	-0.00***	-0.00***	-0.00***	-0.00***	-0.00**	-0.00***

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Industry experience	0.01***	0.02***	0.01***	0.01***	0.01***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Startup's distinctiveness		-0.16*	-1.51***	-0.64**	-0.28***	-1.78***
		(0.06)	(0.32)	(0.23)	(0.09)	(0.38)
Portfolio distinctiveness		-0.14	-1.45***			-1.26***
		(0.13)	(0.33)			(0.34)
Portfolio diversification		0.03***		0.02**		0.02**
		(0.00)		(0.01)		(0.01)
VC status		-0.14***			-0.17***	-0.16***
		(0.01)			(0.01)	(0.01)
Startup's distinctiveness x Portfolio distinctiveness			2.87***			2.42***
			(0.66)			(0.67)
Startup's distinctiveness x Portfolio diversification				0.02*		0.02
				(0.01)		(0.01)
Startup's distinctiveness x VC status					0.06*	0.05
					(0.02)	(0.02)
AIC	49,559	49,149	49,540	49,401	49,284	49,131
BIC	50,106	49,731	50,113	49,975	49,858	49,741
Log Likelihood	-24,719	-24,509	-24,706	-24,637	-24,578	-24,498
Deviance	49,437	49,019	49,412	49,273	49,156	48,9956
Num. obs.	58,000	58,000	58,000	58,000	58,000	58,000

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

* The table shows the results of the logistic regression models with correction of coefficients for rare events. Standard errors are shown in parentheses. Year dummies, round number dummies, and investor count dummies are included in all models. The unit of analysis is the venture capitalist - startup funding round dyad and the sample includes 58,000 observations from the period 2005 - 2015. The observations are half realized investments and half unrealized investments. The dependent variable *dummy actual data* is binary and equals 1 when the observation is a realized investment.

TABLE 6.4: Logistics regression output models part I: startup's distinctiveness*

6.2.2 Analysis part II: categorical distance

The regression output for the second part of the analysis of categorical distance is displayed in table 6.5. We have included model 0 again in the table as a reference for the new models. Similarly to model 1, we have included all main effects of our independent variables in model 7. We confirm with a strongly significant negative coefficient that categorical distance has a negative effect on the probability of receiving funding from the focal venture capitalist. This effect remains strongly significant in all models and has the highest coefficient in the full model 11.

Model 8 adds the interaction effect of portfolio distinctiveness on categorical distance and its main effect to model 7. As expected, we find a positive coefficient, that is, an increase in the log odds of receiving an investment when the venture capitalist has higher portfolio distinctiveness. However, the effect is not significant at the 5% level. The interaction effect plot in figure 6.8d also shows highly overlapping confidence intervals, even between values for portfolio distinctiveness, which are four standard deviations apart from each other.

Model 9 includes the main and interaction effects of portfolio diversification and shows strongly significant support for the hypothesis that venture capitalists with higher portfolio diversification have a greater probability of investing in settings with large categorical distance. While the coefficient of the interaction term of portfolio diversification with categorical distance is small (<0.00), the inclusion of the interaction effect and the main effect of portfolio diversification further decreases the negative coefficient (that is, a stronger effect) of the categorical distance's main effect. The plot in figure 6.8e shows the strong differences between the highly diversified venture capitalists with an upward-sloping dotted line and the strongly downward-sloping dotted line for venture capitalists with lowly diversified portfolios.

We can also confirm our hypothesis that high-status investors are more likely to invest in startups with larger categorical distance. Model 10 contains the main and interaction effects of venture capitalist's status and displays a strongly significant coefficient for the interaction term. Again, figure 6.8f shows that high-status venture capitalists have an increasing preference for categorically distant startups, that is, an upward-sloping dotted line, while low-status venture capitalists have a strongly decreasing probability of investing in categorically distant startups, that is, a downward-sloping solid line.

Model 11 contains all main and interaction effect regarding categorical distance allowing us to test the robustness of our results. We find that the results are unchanged with the coefficients showing the same degree of significance and sign.

6.3 Robustness test

Cluster sampling describes the issue of having groups in a sample in which members share common unobserved characteristics, causing correlation among standard errors (Wooldridge, 2002). These biased standard errors can lead to an over- or understatement of the significance levels, that is, validity of results, for the regression coefficients (Petersen, 2009). For instance, if, in an empirical analysis, respondents come from a number of companies, the respondents who work for the same company might share some unobserved characteristics, leading to cross-sectional correlation. According to Wooldridge (2002), this is not a problem as long as the number of clusters is large compared to the observations per cluster and a correction is being made. It is assumed that the clusters themselves are independent of each other. A number of approaches, such as including dummy variables for each cluster, adjusting the standard errors only, and adjusting the standard errors and the coefficients, to correct for this type of non-independence exist and have been applied in other studies (Petersen, 2009).

	Model 0	Model 6	Model 7	Model 8	Model 9	Model 10
(Intercept)	4.99*** (0.35)	4.96*** (0.36)	5.47*** (0.39)	5.32*** (0.37)	6.02*** (0.36)	5.93*** (0.40)
State dummy CA	-0.26*** (0.03)	-0.26*** (0.03)	-0.27*** (0.03)	-0.27*** (0.03)	-0.26*** (0.03)	-0.26*** (0.03)
State dummy MA	-0.23*** (0.04)	-0.20*** (0.04)	-0.21*** (0.04)	-0.21*** (0.04)	-0.22*** (0.04)	-0.21*** (0.04)
State dummy NY	-0.15*** (0.04)	-0.15*** (0.04)	-0.16*** (0.04)	-0.15*** (0.04)	-0.16*** (0.04)	-0.14*** (0.04)
Startup age	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Raised amount (log)	0.06*** (0.01)	0.08*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)
Geographical distance (log)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)	-0.17*** (0.00)
Geographical distance no data dummy	0.36*** (0.06)	0.37*** (0.06)	0.37*** (0.06)	0.41*** (0.06)	0.32*** (0.06)	0.36*** (0.06)
Industry distance	-1.52*** (0.07)	-1.09*** (0.08)	-1.13*** (0.08)	-1.02*** (0.08)	-1.01*** (0.08)	-0.92*** (0.08)
Prior investment dummy	5.84*** (0.17)	5.82*** (0.17)	5.86*** (0.17)	5.84*** (0.17)	5.83*** (0.17)	5.81*** (0.17)
Mean affiliation	0.04*** (0.00)	0.05*** (0.00)	0.04*** (0.00)	0.04*** (0.00)	0.05*** (0.00)	0.05*** (0.00)
Affiliate geographical distance	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Affiliate industry distance	-3.73*** (0.08)	-3.73*** (0.09)	-3.71*** (0.08)	-3.69*** (0.08)	-3.86*** (0.09)	-3.83*** (0.09)
VC age	0.00 (0.00)	-0.00 (0.00)	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
General experience	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00** (0.00)	-0.00*** (0.00)

	Model 0	Model 6	Model 7	Model 8	Model 9	Model 10
Industry experience	0.01*** (0.00)	0.02*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.02*** (0.00)	0.01*** (0.00)
Categorical distance		-0.03*** (0.00)	-0.03*** (0.01)	-0.09*** (0.01)	-0.04*** (0.00)	-0.09*** (0.01)
Startup's distinctiveness		-0.11 (0.06)	-0.11 (0.06)	-0.09 (0.06)	-0.10 (0.06)	-0.08 (0.06)
Portfolio distinctiveness		-0.17 (0.13)	-0.67 (0.39)			0.09 (0.39)
Portfolio diversification		0.03*** (0.00)		-0.03*** (0.01)		-0.02** (0.01)
VC status		-0.14*** (0.01)			-0.31*** (0.02)	-0.31*** (0.02)
Categorical distance x Portfolio distinctiveness			0.02 (0.02)			-0.01 (0.02)
Categorical distance x Portfolio diversification				0.00*** (0.00)		0.00*** (0.00)
Categorical distance x VC status					0.01*** (0.00)	0.01*** (0.00)
AIC	49,559	48,951	49,450	49,094	49,047	48,711
BIC	50,106	49,543	50,033	49,677	49,630	49,330
Log Likelihood	-24,719	-24,409	-24,660	-24,482	-24,458	-24,287
Deviance	49,437	48,819	49,320	48,964	48,917	48,573
Num. obs.	58,000	58,000	58,000	58,000	58,000	58,000

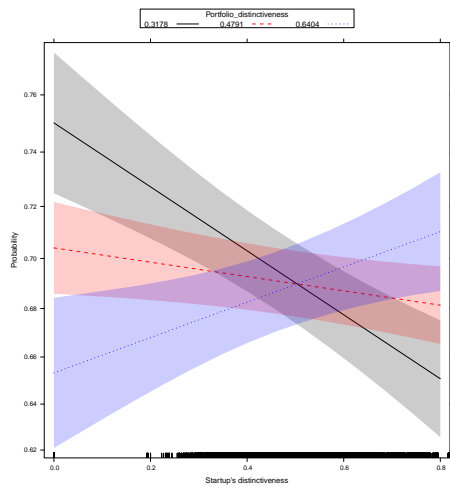
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

* The table shows the results of the logistic regression models with correction of coefficients for rare events. Standard errors are shown in parentheses. Year dummies, round number dummies, and investor count dummies are included in all models. The unit of analysis is the venture capitalist - startup funding round dyad and the sample includes 58,000 observations from the period 2005 - 2015. The observations are half realized investments and half unrealized investments. The dependent variable *dummy actual data* is binary and equals 1 when the observation is a realized investment.

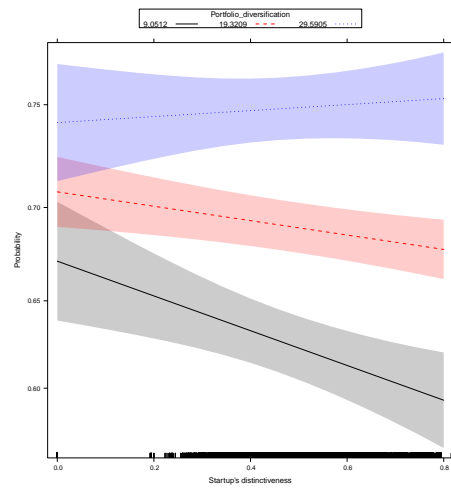
TABLE 6.5: Logistics regression output models part II: categorical distance*

In our dataset, each observation is a part of multiple clusters, which might lead to an unobserved correlation and thus non-independence of observations (Ozmel, Reuer, and Gulati, 2013). Firstly, a venture capitalist-startup observation is part of the startup cluster containing all the other observations of the focal funding round and other funding rounds of the focal startup. Secondly, the observation belongs to the venture capitalist cluster, including all observations in which the focal venture capitalist has invested. Thirdly, the funding round cluster includes all observations, that is, venture capitalist - startup combinations, in the focal funding round. Theoretically, within each of these clusters, unobserved characteristics might lead to correlation. We have thus conducted a robustness test by including multi-way clustering of standard errors for startups and for venture capitalists. We did not cluster the funding rounds, as they are clustered in a higher order through startup clustering. We used the approach for the multi-way clustering of residuals to estimate the robust standard errors developed by Cameron, Gelbach, and Miller (2012) because it allows multi-way clustering and is applicable to logistic regressions. We applied the implementation in the R package *multiwaycov* (Graham, Arai, and Hagstroemer, 2016). Cameron, Gelbach, and Miller (2012) adjust the standard errors to correct for potential in-cluster correlation.

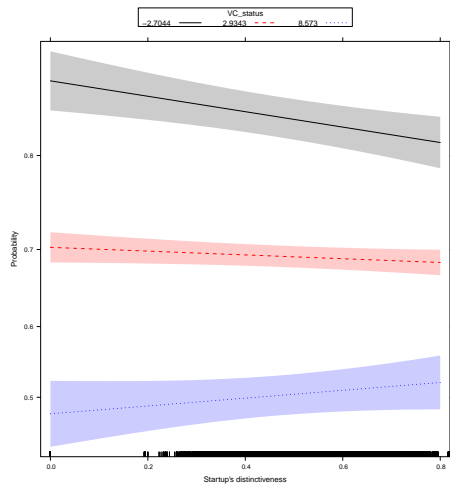
The results of the regressions, including the adjusted standard errors, can be found in table 6.6 and table 6.7. We have excluded the rare-events correction of coefficients described in section 5.2.2, which is why coefficients may differ from those in tables 6.4 and 6.5. The robustness check shows that, even with the correction of standard errors, the results remain the same qualitatively, that is, in the direction of coefficients, and quantitatively, that is, in the significance levels.



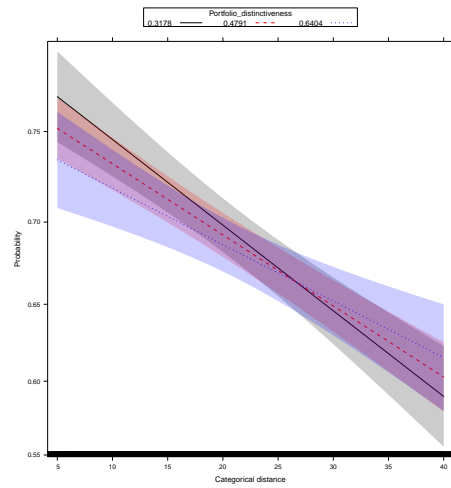
(A) Startup's distinctiveness with portfolio distinctiveness



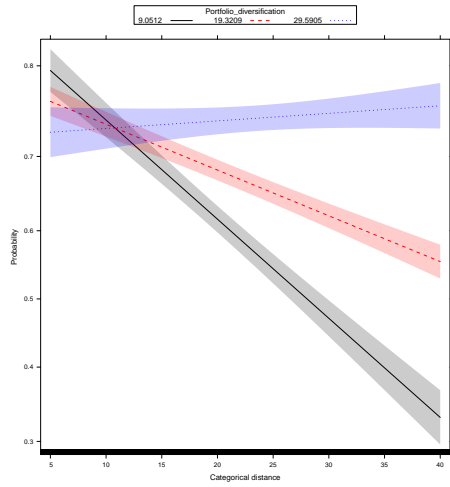
(B) Startup's distinctiveness with portfolio diversification



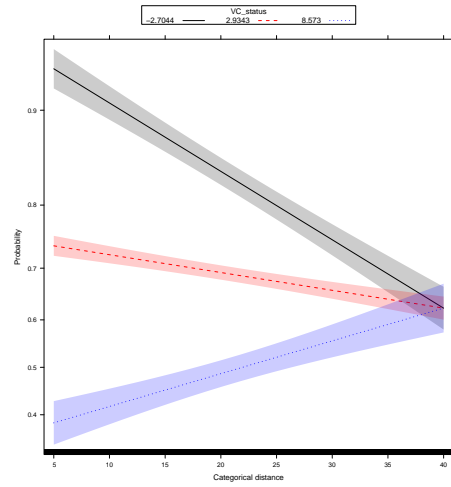
(C) Startup's distinctiveness with venture capitalist's status



(D) Categorical distance with portfolio distinctiveness



(E) Categorical distance with portfolio diversification



(F) Categorical distance with venture capitalist's status

FIGURE 6.8: Graphical analysis of interactions

	Model 0.1	Model 1.1	Model 2.1	Model 3.1	Model 4.1	Model 5.1
(Intercept)	5.00*** (0.47)	4.92*** (0.47)	5.69*** (0.48)	4.53*** (0.46)	5.59*** (0.46)	5.62*** (0.48)
State dummy CA	-0.26*** (0.04)	-0.25*** (0.04)	-0.26*** (0.04)	-0.26*** (0.04)	-0.25*** (0.04)	-0.25*** (0.04)
State dummy MA	-0.23*** (0.06)	-0.22*** (0.06)	-0.23*** (0.06)	-0.22*** (0.06)	-0.23*** (0.06)	-0.22*** (0.06)
State dummy NY	-0.15** (0.06)	-0.13* (0.06)	-0.15** (0.06)	-0.14* (0.06)	-0.14* (0.06)	-0.13* (0.06)
Startup age	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Raised amount (log)	0.06* (0.03)	0.07** (0.03)	0.07** (0.03)	0.07** (0.02)	0.07** (0.03)	0.07** (0.03)
Geographical distance (log)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)
Geographical distance no data dummy	0.36*** (0.06)	0.35*** (0.05)	0.36*** (0.06)	0.40*** (0.06)	0.31*** (0.06)	0.35*** (0.06)
Industry distance	-1.52*** (0.10)	-1.61*** (0.11)	-1.53*** (0.10)	-1.68*** (0.11)	-1.46*** (0.11)	-1.60*** (0.11)
Prior investment dummy	5.86*** (0.19)	5.83*** (0.19)	5.86*** (0.19)	5.86*** (0.19)	5.83*** (0.19)	5.82*** (0.19)
Mean affiliation	0.04*** (0.01)	0.05*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.05*** (0.01)	0.05*** (0.01)
Affiliate geographical distance	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Affiliate industry distance	-3.74*** (0.16)	-3.79*** (0.16)	-3.74*** (0.16)	-3.71*** (0.16)	-3.82*** (0.16)	-3.78*** (0.16)
VC age	0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)
General experience	-0.00***	-0.00	-0.00***	-0.00***	-0.00	-0.00

	Model 0.1	Model 1.1	Model 2.1	Model 3.1	Model 4.1	Model 5.1
Industry experience	(0.00) 0.01** (0.01)	(0.00) 0.02** (0.01)	(0.00) 0.01* (0.01)	(0.00) 0.01** (0.01)	(0.00) 0.01** (0.01)	(0.00) 0.02** (0.01)
Startup's distinctiveness		-0.16* (0.07)	-1.51*** (0.33)	-0.64** (0.24)	-0.28** (0.09)	-1.78*** (0.39)
Portfolio distinctiveness		-0.14 (0.11)	-1.45*** (0.34)			-1.26*** (0.32)
Portfolio diversification		0.03*** (0.00)		0.02** (0.01)		0.02** (0.01)
VC status		-0.14*** (0.02)			-0.17*** (0.02)	-0.16*** (0.02)
Startup's distinctiveness x Portfolio distinctiveness			2.87*** (0.69)			2.42*** (0.67)
Startup's distinctiveness x Portfolio diversification				0.02* (0.01)		0.02 (0.01)
Startup's distinctiveness x VC status					0.06* (0.03)	0.05 (0.03)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

* The table shows the results of the logistic regression models. Standard errors are shown in parentheses and clustered by startup and venture capitalist. Year dummies, round number dummies, and investor count dummies are included in all models. The unit of analysis is the venture capitalist - startup funding round dyad and the sample includes 58,000 observations from the period 2005 - 2015. The observations are half realized investments and half unrealized investments. The dependent variable *dummy actual data* is binary and equals 1 when the observation is a realized investment.

TABLE 6.6: Regression output models with two-way clustering of standard errors by startup and venture capitalist part I: startup's distinctiveness*

	Model 0.1	Model 6.1	Model 7.1	Model 8.1	Model 9.1	Model 10.1
(Intercept)	5.00*** (0.47)	4.96*** (0.48)	5.48*** (0.47)	5.33*** (0.46)	6.02*** (0.46)	5.93*** (0.47)
State dummy CA	-0.26*** (0.04)	-0.26*** (0.04)	-0.27*** (0.04)	-0.27*** (0.04)	-0.26*** (0.04)	-0.26*** (0.04)
State dummy MA	-0.23*** (0.06)	-0.20** (0.06)	-0.21*** (0.06)	-0.21*** (0.06)	-0.22*** (0.06)	-0.22*** (0.06)
State dummy NY	-0.15** (0.06)	-0.15** (0.06)	-0.16** (0.06)	-0.15** (0.06)	-0.16** (0.06)	-0.14* (0.06)
Startup age	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Raised amount (log)	0.06* (0.03)	0.08** (0.03)	0.07** (0.03)	0.07** (0.03)	0.07** (0.03)	0.07** (0.03)
Geographical distance (log)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)	-0.17*** (0.01)
Geographical distance no data dummy	0.36*** (0.06)	0.37*** (0.05)	0.37*** (0.06)	0.41*** (0.05)	0.32*** (0.06)	0.36*** (0.05)
Industry distance	-1.52*** (0.10)	-1.09*** (0.11)	-1.13*** (0.11)	-1.02*** (0.11)	-1.01*** (0.11)	-0.92*** (0.11)
Prior investment dummy	5.86*** (0.19)	5.85*** (0.19)	5.89*** (0.19)	5.87*** (0.19)	5.85*** (0.19)	5.83*** (0.19)
Mean affiliation	0.04*** (0.01)	0.05*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.05*** (0.01)	0.05*** (0.01)
Affiliate geographical distance	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Affiliate industry distance	-3.74*** (0.16)	-3.74*** (0.16)	-3.71*** (0.16)	-3.69*** (0.16)	-3.87*** (0.16)	-3.84*** (0.16)
VC age	0.00 (0.01)	-0.00 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
General experience	-0.00***	-0.00	-0.00***	-0.00***	-0.00	-0.00

	Model 0.1	Model 6.1	Model 7.1	Model 8.1	Model 9.1	Model 10.1
Industry experience	(0.00) 0.01** (0.01)	(0.00) 0.02** (0.01)	(0.00) 0.01** (0.01)	(0.00) 0.01* (0.01)	(0.00) 0.02** (0.01)	(0.00) 0.01* (0.01)
Categorical distance		-0.03*** (0.00)	-0.03*** (0.01)	-0.09*** (0.01)	-0.04*** (0.00)	-0.09*** (0.01)
Startup's distinctiveness		-0.11 (0.07)	-0.11 (0.06)	-0.09 (0.07)	-0.10 (0.06)	-0.08 (0.07)
Portfolio distinctiveness		-0.17 (0.11)	-0.67 (0.42)			0.09 (0.39)
Portfolio diversification		0.04*** (0.00)		-0.03*** (0.01)		-0.02** (0.01)
VC status		-0.14*** (0.02)			-0.31*** (0.03)	-0.31*** (0.03)
Categorical distance x Portfolio distinctiveness			0.03 (0.02)			-0.01 (0.02)
Categorical distance x Portfolio diversification				0.00*** (0.00)		0.00*** (0.00)
Categorical distance x VC status					0.01*** (0.00)	0.01*** (0.00)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

* The table shows the results of the logistic regression models. Standard errors are shown in parentheses and clustered by startup and venture capitalist. Year dummies, round number dummies, and investor count dummies are included in all models. The unit of analysis is the venture capitalist - startup funding round dyad and the sample includes 58,000 observations from the period 2005 - 2015. The observations are half realized investments and half unrealized investments. The dependent variable *dummy actual data* is binary and equals 1 when the observation is a realized investment.

TABLE 6.7: Regression output models with two-way clustering of standard errors by startup and venture capitalist part 2: categorical distance*

Chapter 7

Discussion

The overall aim of this dissertation is to further investigate the matching process of venture capitalists and startups in order to explain which venture capitalist - startup combinations are most likely. We use market categories for our analysis and apply the latest sociological research findings from the categorization literature on the venture capital domain. Specifically, using market categories and taking the distance between the categories into account, as well as the number of categories, is a new approach to analysis of the venture capital industry. In a sample of US startups that received funding during the period from 2005 to 2015, we find evidence that which venture capitalists invest in what startups depends on venture capitalists' experience regarding investments in distinctive new ventures, their portfolio diversification, and their status.

The chapter debates the results of the empirical tests regarding research question 1 on startup's distinctiveness, which is followed by the discussion regarding research question 2 on categorical distance, and the chapter closes with a brief consideration of how these results change for different types of venture capitalists, namely, independent and corporate venture capitalists.

7.1 The effects on the preference for new venture's distinctiveness

Organizations need to be legitimate in order to be accepted by various audiences, like customers, employees, or investors. Simultaneously, organizations also need to demonstrate novelty or distinctiveness (Aldrich and Fiol, 1994), and the optimal degree of distinctiveness versus legitimacy depends on the audience (Pontikes, 2012). We have defined a measure of distinctiveness based on the market categories that new ventures are associated with and the categories' prior co-occurrences. We discuss our hypotheses regarding the effects of a venture capitalists' average portfolio distinctiveness and portfolio diversification as well as status on the preference for distinctive new ventures and the resulting effect on the probability

of investing in them. Subsequently, we discuss the overarching contributions by our measure of a new venture's distinctiveness and the heterogeneity among venture capitalists. Most notably, while classical categorization theorists argue that boundary-spanning organizations are hard to understand and are discounted (Zuckerman, 1999; Zuckerman et al., 2003; Hsu, 2006), we are able to show that certain venture capitalists prefer distinctive, boundary-spanning new ventures.

7.1.1 Portfolio distinctiveness

We have postulated above that venture capitalists make portfolio choices with regard to the degree of distinctiveness and thus novelty in their portfolios. They also accumulate experience from investing in startups with a certain degree of distinctiveness. The greater the experience in investing in distinctive startups, the more likely the venture capitalists are to invest in distinctive startups in the future. We have found strong support for this hypothesis. This finding is not surprising for three reasons. Firstly, venture capitalists typically have an investment strategy or focus that they continue to pursue, so historical investments should predict future investment decisions. Secondly, venture capitalists have communicated their investment strategies to their limited partners, who themselves choose the venture capitalists to manage their capital based on a defined investment strategy. Thus, the focal venture capitalist cannot deliberately change the investment focus of their fund regarding the degree of startup's distinctiveness. Thirdly, we expect venture capitalists to improve when they gain more experience. Investing in a similar degree of distinctiveness allows the venture capitalists to gain experience managing these types of new ventures and the associated risks.

It follows that the current degree of distinctiveness of a venture capitalist's portfolio is an indicator for the future investment decision of the venture capitalist. Limited partners, employees, or startups can use this information when assessing a venture capitalist and then choose one that suits their needs or preferences. For instance, startups should approach investors who have previously invested in startups with a degree of novelty and distinctiveness similar to their own to increase the chances of being funded with the required capital.

7.1.2 Portfolio diversification

The results of our analysis with regards to the effect of portfolio diversification on the preference for distinctiveness are most surprising. Venture capitalists face individual investment risk, that is, the company-specific risk

of investing in an individual startup, and portfolio risk (Norton and Tenenbaum, 1993). The investment-specific risk can be diversified away (Sharpe, 1964), either by the venture capitalist him or herself to reduce the risk of low performance, or by the limited partner who can invest in multiple venture capitalists with high or low degrees of diversification. While diversification or specialization is possible in multiple dimensions, such as geographical, industry, or development-stage diversification, we measure diversification based on market categories. From a performance perspective, specialization has been found to improve venture capitalists' performance (Gompers, Kovner, and Lerner, 2009), while, more recently, a U-shaped relationship between diversification and performance has been found (Matusik and Fitza, 2012). It is generally accepted that diversification is a risk-reduction strategy (Sharpe, 1964), and we thus argue above that diversifying investors are risk-averse. More distinctive startups provide higher risk and, on average, rewards from their greater novelty. We thus hypothesize that risk-averse investors, as indicated by the degree of diversification, have a lower probability of investing in distinctive new ventures. Our empirical analysis, however, has not found support for this hypothesis. We observe quite the opposite: Instead of a negative effect of portfolio diversification on the probability of investment in distinctive startups, we have found a weakly significant positive effect. This means that venture capitalists with higher degrees of diversification in their portfolios are more likely to invest in distinctive startups than less diversified venture capitalists. Multiple explanations for this finding exist. One possible line of reasoning is that venture capitalists who diversify their portfolios, and thus reduce the risk associated with said portfolios, have more risk capacity to invest in more distinctive and riskier new ventures. For instance, venture capitalists with low levels of diversification and a high degree of distinctiveness have a higher risk of under-performance than venture capitalists with highly diversified portfolios and the same level of distinctiveness.

Another explanation could be that there is no relationship: The venture capitalist's decision about the degree of portfolio diversification is independent of the decision regarding the level of a new venture's distinctiveness in their portfolio. The investors might also choose to spread their investments in terms of the levels of distinctiveness, that is, they invest in new ventures with high and low degrees of distinctiveness. As in model 3, the relevant effect is only weakly significant, and, in the full model 5, the effect is no longer significant because, for example, parts of the effect are captured by the other predictors due to correlation. Thus, we consider it possible that there is no causal relationship between diversification and the preference for distinctiveness. The third explanation for the unexpected results is that our assumption about measuring a venture capitalist's risk aversion

by the degree of market-category-based portfolio diversification does not hold. It is possible that venture capitalists choose their level of diversification consciously or unconsciously based on their perception of an ideal diversification level, which varies even in empirical studies (compare Gompers, Kovner, and Lerner, 2009; Matusik and Fitza, 2012), independently of their level of risk aversion. Our argument about the relationship between risk aversion and the preference for distinctiveness would still be true, but the measure of risk aversion would be incorrect.

Our surprising finding shows that the venture capital context is different from other contexts. For instance, Hannan, Goldberg, and Kovács (2016) analyze the consumption behavior of consumers in the context of moviegoers and restaurant visitors. They use film genres and restaurant cuisine types as their categories and use ratings to measure the consumers' preferences. Their empirical analysis shows that there is a positive relationship between consumers' preferences for established, less boundary-spanning categories, that is, low levels of distinctiveness, and variety, that is, diversification in their consumption behavior. In line with other authors who demonstrate that venture capitalists' behavior and preferences are different from those of other audiences (e.g., Wry, Lounsbury, and Jennings, 2014; Pontikes, 2012), our results show that venture capitalists behave differently than consumers. Based on our empirical results, there is a positive relationship between portfolio diversification and the preference for distinctive new ventures.

7.1.3 A venture capitalist's status

A venture capitalist's status, as commonly agreed upon in literature, is based on the structural network contacts a focal venture capitalist has. It is quantified based on the centrality in the co-syndication network and dependent on the syndication networks of the syndication partners. Often, status is also considered a quality signal, as it is a relative ordering of actors in a network (Podolny, 1993; Podolny, Stuart, and Hannan, 1996). We find support for our hypothesis that high-status venture capitalists are more likely to invest in distinctive startups than venture capitalists with lower status levels. Two elements in our argumentation lead us to this conclusion.

Firstly, scholars have previously found a U-shaped relationship between status and the preference for boundary-spanning distinctive objects (Blau, 1960; Hollander, 1958; Hollander, 1960), that is, low- and high-status actors prefer boundary-spanning objects, whereas middle-status actors prefer legitimacy. In the venture capital context, venture capitalists have to justify their investments to their limited partners, who provide capital to the venture capitalists and expect a certain return on this investment. Similarly, to the previously found U-shaped relationship, we expect middle-status

venture capitalists to avoid distinctive new ventures, as they are afraid of losing their non-manifested status in the case of failure. Their limited partners might discard these venture capitalists if they invest in less legitimized new ventures with higher potential to fail.¹ On the contrary, high-status investors have enough status so as not to lose their status level in the case of failure. They can thus afford to invest in distinctive, riskier, and less legitimized new ventures. Based on the findings that these new ventures are more innovative and potentially more disruptive (Rosenkopf and Nerkar, 2001), we expect high-status venture capitalists to prefer distinctive new ventures. On the other end of the status spectrum are the low-status investors, who can also potentially invest in distinctive new ventures, because they are in a nothing-to-lose situation. They do not have a significant status level to lose, but they can gain by successfully investing in distinctive new ventures. However, we expect that these venture capitalists are not able to invest in these distinctive new ventures as outlined in the following.

Secondly, not only do venture capitalists choose their new venture investments, but these new ventures also choose their venture capitalists to best fulfill their required need for capital, knowledge, network contacts, or signals. This is one of the reasons why the venture capitalist - startup relationship is different from other typical consumer settings, such as moviegoers or restaurant visitors, in which typically only one side chooses from a great variety of product or service offerings. Consequently, we argue that highly distinctive new ventures choose the investors with the best networks, which provide them with the highest quality signal (compare Ozmel, Reuer, and Gulati, 2013), and help them to become more innovative (Pahnke, Katila, and Eisenhardt, 2015). As high-status venture capitalists provide this (Podolny, 1993; Rao, 1994; Pahnke, Katila, and Eisenhardt, 2015), distinctive new ventures favor high-status and disregard low-status venture capitalists. In support of this, new ventures choose highly reputable venture capitalists even if it leads to a discount in their own valuation (Hsu, 2004). It follows that the U-shaped relationship does not hold in the venture capital context, but a linear relationship exists.

As our empirical results support this argumentation, the implications are manifold and new. Low-status venture capitalists, for example, new venture capitalists who have raised their first fund, need to find other ways of making themselves attractive to startups as investors. For instance, this can be an experienced venture capital team with a great network in the investor and startup community, links to corporations, or the provision of specialized expertise. Furthermore, venture capitalists should seek to

¹Even if the startups have the same probability of failure, it is harder to justify investments in less legitimized startups than in highly legitimate new ventures.

quickly build up high levels of status to get access to more distinctive new ventures. They can achieve this through syndication with other high-status investors, even though it may be hard to attract high-status venture capitalists for syndication because low-status venture capitalists lack the required quality signal to do so (Dimov and Milanov, 2010). The line of reasoning that venture capitalists need to build up status quickly is based on the assumption that startups choose their investors based on status. We acknowledge that the outlined strategy for venture capitalists may just be one of many strategies a venture capitalist can pursue. For instance, we argue that distinctive, boundary-spanning startups can potentially lead to more disruptive innovations, as they combine new elements together, and the startups belong to the first companies to do so. While less distinctive startups may instead develop incremental innovations or simply be better at executing established business models with a lower degree of innovation, these new ventures can also be very successful. Some venture capitalists might thus pursue a strategy to focus on these type of startups and it might not be a necessary or desirable strategy for these venture capitalists to build up status quickly.

Another way of thinking about this is that more distinctive startups are able to attract higher status investors, who, in turn, not only have a denser network than their low-status peers, but also provide a positive quality signal for new ventures. Similarly to the Matthew effect (Merton, 1968), which states that even small advantages in the beginning can lead to large advantages or even self-fulfilling prophecies at a later point in time, one could argue that distinctive new ventures that are able to attract high-status venture capitalists in the beginning of their development, that is, in early funding rounds, are able to attract even more high-status investors and thus resources at later stages. This, in turn, can lead to faster growth and better performance. The effect might even lead to a manifestation of small quality differences between startups, as these differences are magnified by the high-status venture capitalist's attention and signals (Gould, 2002). It would thus be another avenue for further research to analyze the effect of early-stage distinctiveness on a new venture's future performance, as we discuss in section 8.3.

7.1.4 Overarching discussion of new venture's distinctiveness

It follows that the desire for distinctiveness varies not only between audience groups like employees, consumers, or investors, but also within these groups. Our analysis does not show whether more distinctive startups are more likely to get funding. However, it does show that differences in the levels of distinctiveness between the individual startups exist and that certain venture capitalists are more likely to invest in a certain degree

of distinctiveness. This is a major contribution to the categorization literature and extends the argumentation of Pontikes (2012) and Wry, Lounsbury, and Jennings (2014), who also propose that venture capitalists evaluate boundary-spanning organizations differently than other audiences. Other categorization researchers have demonstrated theoretically and empirically that the objects and organizations with a lack of legitimacy from being boundary-spanning, that is, spanning multiple categories, are less successful. For instance, Zuckerman (1999) shows that analysts discount companies that span categories and Dobrev, Kim, and Hannan (2001) find lower survival rates for companies that are boundary-spanning. Generally, organizations thus have pressure to conform to existing beliefs and categories to create legitimacy (DiMaggio and Powell, 1983; Meyer and Rowan, 1977). Hsu, Hannan, and Koçak (2009) identify two lines of argumentation as to why boundary-spanning organizations are discounted: Firstly, organizational membership in multiple categories signals – as, in most cases, the true attributes and characteristics cannot be observed – that this organization cannot have the same "ability" and "experience" (Hsu, Hannan, and Koçak, 2009, p. 151) in both or multiple categories as a single category member. The products are thus perceived to be of lower quality. Secondly, the authors identify the reasoning that category-spanning reduces the ability to be attractive to multiple audiences. In that argumentation, the organization cannot fulfill the expectations of all the different audiences for multiple categories at the same time. The organization is thus only attractive to one audience or more likely to be less attractive to any audience than a single category member. In sum, being associated with multiple categories has been found to negatively affect evaluations.

We build our analysis on the extension of this concept by Kovács and Hannan (2015), who consider the conceptual space in which categories are grounded, and add the distance between categories in their analysis. Applying their object-oriented approach to organizations in the venture capital domain is new and allows for a more differentiated consideration than that performed by Pontikes (2012) and Wry, Lounsbury, and Jennings (2014). Wry, Lounsbury, and Jennings (2014) only consider the categories of science and technology, but do not further distinguish between a more fine-grained categorization. Moreover, the authors have only tested their hypotheses in a sample from the nanotechnology industry, while we have been able to derive more general conclusions, as our sample is not limited to any single industry. Pontikes (2012) uses measures of category-spanning and category ambiguity, with the latter being a measure of how defining a category is; for example, a category that is more broadly used in combination with many other categories has less defining power than a category that is less broadly

used. She analyzes the co-occurrence network of categories from press releases to define her measure. However, she does not incorporate the categorical distance into her analysis. Our distance-based empirical analysis increases the robustness of the findings and has a number of highly relevant implications with regard to a new venture's distinctiveness: Firstly, distinctiveness, as measured by boundary-spanning categories, is not generally bad for new ventures, but might help them to receive venture capital funding. Secondly, it confirms prior findings that venture capitalists are different from other audiences in their evaluation of organizations. Thirdly, it extends prior research in the way that we demonstrate that there is great heterogeneity among venture capitalists. Pontikes (2012) argues that venture capitalists are different from other audiences, but her analysis is limited, as she does not differentiate enough between different types of venture capitalists.² Not all venture capitalists prefer distinctive new ventures in the same way, but the preference depends on the experience with these types of investments, the portfolio strategy regarding diversification, and the status derived from the venture capitalist's structural embeddedness. The valuation risk associated with investing in distinctive new ventures (Zuckerman, 2015) is thus not accepted to the same degree by all venture capitalists.

Having understood that venture capitalists behave differently than other audiences who have been found to discount category-spanning organizations (e.g., Zuckerman, 1999; Pontikes, 2012), we now discuss why this is the case. Firstly, in the pre-investment due diligence, venture capitalists conduct ample research on the new ventures that they are considering for investment. For instance, they get to know the management or founding team, try to understand the business model, and analyze the associated risks. Thus, they do not fully rely on categorization as a navigation tool, they spend more time on company analysis than a typical consumer, and they thus find it less difficult to interpret boundary-spanning organizations. Secondly, once invested, as equity or convertible debt-owners, venture capitalists can influence the strategy and direction of the new venture through board membership and informal contact with the management team. This is contrary to, for instance, the consumers, who cannot influence or control the organization (Pontikes, 2012). Thirdly, distinctive or boundary-spanning new ventures can be attractive, as the boundary-spanning attribute is a measure of novelty. Novel business models have the greatest potential for abnormal returns, for example, from breakthrough innovations (Rosenkopf and Nerkar, 2001); therefore, venture capitalists are interested in distinctive new ventures, which they try to understand and evaluate in their due diligence.

²She only argues that corporate venture capitalists are more similar to consumers than to independent venture capitalists, which we discuss in section 7.3.

Furthermore, we make a contribution to the literature on entrepreneurial identity. Navis and Glynn (2011) develop a conceptual framework to explain how the entrepreneurial identity from claiming certain labels can bridge the trade-off between legitimacy and distinctiveness. Our analysis supports their line of reasoning for two arguments: Firstly, we show that there is heterogeneity among new ventures despite the need for legitimacy, which has been found to lead to isomorphism among organizations. Secondly, we demonstrate that there is a tradeoff between being legitimate and being distinctive. If different audiences have different preferences, new ventures find themselves in an inherent trade-off to satisfy different expectations. Navis and Glynn (2011) argue that entrepreneurial identities need to tell coherent stories (see also Magretta, 2002) to create legitimate distinctiveness. Our measure of distinctiveness is the first step towards further exploring how legitimately distinctive a startup is. However, we acknowledge that our approach does not allow us to further distinguish between positively and negatively distinctive startups. When future researchers develop a method to measure the legitimacy of distinctiveness, we might even find stronger support for our proposed relationships. The non-existence of such a measure also supports our choice of not making any predictions as to whether distinctive new ventures are more successful, because we would need to be able to distinguish between legitimately distinctive and non-legitimately distinctive startups. Furthermore, Navis and Glynn (2011) differentiate between new and established markets. In established markets, startups are compared to prototypes of the existing market, whereas, in new markets, the startups become the prototype that others are compared to. Our measure of market-category-based distinctiveness integrates that concept: We compare a focal startup's categories to the previously used categories and measure the distance in case they have been used before. In the case that a category has not been used before, for instance, a new market category, we assign the greatest distance to that category. The next time a startup is associated with that category, it is compared to the startup that has previously been associated with that category and the other categories the startup has been associated with. The first startup with a particular category, thus, becomes the prototype for said category. Our distinctiveness measure thus also reproduces these elements of the authors' framework.

The trade-off between legitimacy and distinctiveness remains complicated for new ventures. Based on our findings and prior contributions (Wry, Lounsbury, and Jennings, 2014; Navis and Glynn, 2011; Pontikes, 2012), we recommend not generally perceiving category-spanning entrepreneurial identities in a negative way, but rather approaching them neutrally. If done right, distinctiveness can be beneficial, and it seems more

important to find the right match, that is, the best fitting venture capitalist, than to focus on acquiring a certain degree of legitimacy or distinctiveness.

7.2 The effect of categorical distance

The second part of our analysis concerns the effect of categorical distance on the probability of a match between a venture capitalist and a startup and how the sensitivity to categorical distance changes with three attributes of the venture capitalist. We use a measure of categorical distance based on the same associated market categories that we used to measure the distinctiveness of the new venture, which links the categorical distance to the distinctiveness attribute.

7.2.1 Overarching discussion of categorical distance

It is widely accepted that distance reduces the likelihood of tie formation. For instance, Sorenson and Stuart (2001) show that geographically and industry distant venture capitalists have a lower probability of realizing an investment than their closer peers. We find strong support for our hypothesis that the categorical distance also reduces the probability of an investment. There are three reasons for this effect. Firstly, access to information about investment opportunities is greater in categorically closer new ventures. Venture capitalists' expertise, experience, and cognitive behavior restrict access to new information about these more distant new ventures. Thus, venture capitalists do not know about potential investment targets. Secondly, the evaluation of potential investment candidates takes more time, as less prior knowledge exists, and data access is more difficult. Due to a lack of knowledge, the assessment is also harder for the venture capitalist, which can lead to a valuation discount and thus reduce the probability of investment. Thirdly, greater categorical distance also reduces the ability of venture capitalists to add value, as more time is required to become acquainted with the new venture, if at all possible. This also leads to valuation discounts and reduces the likelihood that new ventures choose more distant venture capitalists.

Our findings also make a methodological contribution regarding distance measures. Our measure of categorical distance is at the intersection of cultural, institutional, and industry distance. The use of market categories is similar to the use of industry categories by, for instance, Sorenson and Stuart (2001), but we do not use a simple count of previous investments in the same industry segment as they do. The market categories we use are also the basis for the degree of a new venture's distinctiveness. Startups can be distinctive, that is, boundary-spanning, when multiple categories

are assigned to a new venture and when the distance between each of these categories is large. Equivalently, we consider the distance between the focal startup's categories to the categories of each of the portfolio companies of the venture capitalist. If the focal new venture is highly culturally embedded, so that the categories are less boundary-spanning, and if the portfolio companies of the focal venture capitalist are also highly culturally embedded, then the categorical distance between these two is relatively small. We have thus developed a new distance measure that lies somewhere between the classical cultural distance of countries (e.g., Hofstede, 1980; Kogut and Singh, 1988), the institutional distance of countries (e.g., Kostova and Zaheer, 1999; Xu and Shenkar, 2002), and the organizational measure of industry distance (e.g., Sorenson and Stuart, 2001). The difference to the cultural and institutional distances is that we do not compare countries with each other, but instead we compare the categorical embeddedness of organizations. Going forward, our measure of categorical distance can be used to explain other relationships. For instance, the effect of and the sensitivities to categorical distance in the formation of syndicates among venture capitalists could be explored. Most likely, the categorical distance also decreases the chances of syndication, but status, prior investments, or indirect ties, for example, could reduce this effect. Another potential application is the creation of collaboration networks between organizations or the effect of the structural embeddedness of new ventures in their partner networks on sensitivities to categorical distance.

Another element to consider when discussing categorical distance is the use as a novelty measure. Categorical distance measures a startup's novelty compared to the existing portfolio companies of the venture capitalist. This way, the measure is similar to our measure of distinctiveness, which focuses on the novelty of an organization compared to all other organizations. In combination, the two measures can be used to investigate organizational novelty in these two dimensions in other contexts outside of the venture capital domain.

7.2.2 Portfolio distinctiveness

The average portfolio distinctiveness is a measure of how much experience a venture capitalist has with managing the uncertainty of investing in a distinctive new venture. The effect of the moderator portfolio distinctiveness on categorical distance is positive but not significant. We thus cannot support the hypothesis that venture capitalists with greater experience regarding investments in distinctive startups invest in more distant new ventures. An explanation might be that the decision to invest in distinctive new ventures is independent of the decision on categorically distant new ventures. The greater risk appetite as displayed by historical investment in

distinctive new ventures is satisfied by the investment in distinctive new ventures, but it does not increase the chances of investing in more categorically distant new ventures. It is possible that venture capitalists' preferences for categorically distant new ventures do not change with the average degree of portfolio distinctiveness. Instead, these may be two independent decisions, that is, venture capitalists prefer categorically closer startups no matter what their portfolio strategy is regarding distinctiveness.

7.2.3 Portfolio diversification

The empirical analysis supports our hypothesis that venture capitalists with diversified portfolios have a greater risk capacity to invest in more categorically distant new ventures. The risk of investing in individual startups, the idiosyncratic risk, is diversified away by diversification (Sharpe, 1964), which leaves the venture capitalist with only portfolio risk. If he or she then invests in categorically distant new ventures, it is another dimension of further diversifying the portfolio. The support for this hypothesis is also in line with the unexpected finding that well diversified venture capitalists are more likely to invest in more distinctive new ventures, as they have greater risk capacity compared to less diversified peers (see discussion of hypothesis 2 in section 7.1.2). In that light, our line of reasoning for using portfolio diversification as a measure of risk-taking behavior seems off; we cannot confirm that one leads to the other.

Consequently, for new ventures, it might be worth trying to approach more distant venture capitalists if these investors have a well-diversified portfolio. For instance, if a startup is interested in a venture capitalist with specific attributes, for example, specific network contacts, large categorical distance is less of an issue, when the existing portfolio is diversified.

7.2.4 A venture capitalist's status

A venture capitalist's status measures the centrality of the venture capitalist in his or her co-investment network. The data provides convincing evidence for our claim that a venture capitalist's status mitigates the negative effect of categorical distance. Venture capitalists can use the direct and indirect ties of their co-investment network to get more information, resources, or contacts that help the focal venture capitalist to overcome the difficulties of investing in and coaching more categorically distant new ventures. Our results extend prior findings that the need for geographical and industry distance is reduced with the increasing status of venture capitalists (Sorenson and Stuart, 2001) by the categorical distance dimension.

Having high-status investors can serve as a quality signal for new ventures (Podolny, 1993; Podolny, Stuart, and Hannan, 1996), which makes

it attractive for new ventures to have high-status venture capitalists. Our findings provide evidence that categorical distance is less of a problem for finding high-status investors than for low-status venture capitalists. Thus, if a startup has a strong need to gain legitimacy from being funded by high-status venture capitalists, it should not be set off by categorical distance.

7.3 The effect of corporate venture capital

Our empirical analysis aims to differentiate between different types of venture capitalists in terms of their portfolio strategies regarding distinctiveness and diversification, as well as their status. We have explicitly not explored differences in the types of venture capitalists regarding their backgrounds, that is, whether a venture capitalist belongs to a corporate organization or is independent (Shrader and Simon, 1997), because we wanted to use attributes that could change over time to explain the matching between startups and venture capitalists. Nevertheless, in this section, we discuss differences between corporate and independent venture capitalists that we have found in our sample, because the effects and preferences of corporate venture capitalists may differ from those of other types of venture capitalists. For instance, startups funded by corporate venture capitalists are more innovative than their competitors funded by independent venture capitalists (Park and Steensma, 2013; Alvarez–Garrido and Dushnitsky, 2016), compensation schemes differ in corporate venture capital firms, which affects the investment behavior and performance (Dushnitsky and Shapira, 2010), and corporate investors have different investment strategies than independent investors (Shrader and Simon, 1997).

In our dataset, 3,382 venture capitalist - startup observations (5.8% of all observations) made by 82 unique investors are classified as *corporate venture capital*. This is in line with other researchers' datasets that include 6% corporate venture capital investments (Gompers and Lerner, 2000). 41,610 observations are classified as *venture capital*, and the remaining observations were either categorized differently, for instance, as an *individual* or *family investment office*, or not matched, that is, we did not have information regarding the type of investor. The quality of the type of investor entry in the *Crunchbase* database is relatively weak, because the *venture capital* type includes not only independent venture capitalists, but also other investor types. The following additional analysis, the results of which can be found in appendix C, is thus only indicative and used to back our discussion of the effect of corporate venture capital.

Corporate venture capitalists often focus their investments in certain industries related to their core business (Dushnitsky, 2012; Shrader and Simon, 1997); they are often strategic investors who try to benefit from the

innovations in their startup investments (Dushnitsky and Lenox, 2005; Benson and Ziedonis, 2009) and use their industry resources (Sykes, 1990), for example, knowledge, networks, and factories, to assist the startups in their development. We thus expect corporate venture capitalists to have, on average, lower portfolio diversification. We have confirmed this expectation with a t-test comparing the means of the observations classified as *corporate venture capital* and those classified as *venture capital*. Due to variance differences between the two samples, we conducted a Welch two sample t-test with the following results: $t(4070.31) = 26.18, p - value = 2.2e - 16 < .001$. With a portfolio diversification mean for the corporate venture capitalists of 17.4 versus a mean of 19.5 for the non-corporate venture capitalist sample, we see higher specialization levels for the corporate investors.

Due to an informational advantage, corporate venture capitalists are better at evaluating startups than independent venture capitalists when the focal startup is related to the core business of the corporate investor (Gompers and Lerner, 2000). Corporate venture capitalists are thus typically more specialized, which allows them to invest in more distinctive new ventures as long as the startups are complementary to the investor's core business, because the corporate investor has better access to information and more expertise. To test this, we have used the same empirical approach that we used before for our main analysis and have added two interaction effects between a dummy variable, which equals 1 in the case of a corporate venture capitalist and 0 otherwise, and the independent variables measuring the startup's distinctiveness and the categorical distance. We have included only the observations in which the investor is either classified as *corporate venture capital* or *venture capital*. The model results in table C.1 in appendix C support our expectation with a positive and statistically significant coefficient for the interaction with a startup's distinctiveness and a negative and strongly statistically significant coefficient for the interaction with categorical distance. In this sub-sample, corporate venture capitalists are thus more likely to invest in distinctive startups, but they are less likely to invest in categorically distant ones. The latter is in line with prior findings that corporate venture capitalists typically specialize their investments such that those investments are related to their core business (Gompers and Lerner, 2000; Dushnitsky, 2012) and invest in less distant startups. However, the results are contrary to the research by Pontikes (2012), who found that corporate venture capitalists prefer less ambiguous startups. Her reasoning for the empirical finding is that corporate venture capitalists try to preserve their existing business, which is, by definition, clearly categorized and not category-spanning. She interpreted prior findings by Tushman and

Anderson (1986) and Sorensen and Stuart (2000), who show that large corporations are more likely to develop incremental instead of disruptive innovations, as an argument for corporate venture capitalists to invest in less distinctive startups. However, she neglected the fact that large corporations' existing businesses are often threatened by startups today: For instance, *Uber* is a threat to the taxi industry, *Amazon* to brick-and-mortar book stores, *AirBnB* to hotels, and *Tesla* to established premium car manufacturers. A great number of other less-well-known niche startups exist that may threaten the existing business of large corporations. In our opinion, the opposing findings of our empirical analysis and that of Pontikes may be a result of a change in strategy with regard to dealing with these types of startups. Many large corporations, particularly those with a venture capital unit, may invest in disruptive startups in order to learn from and collaborate with them rather than to neglect them; without these investments, they would not be able to make these types of disruptive innovations. However, corporate venture capitalists only prefer startups that are categorically closer to their existing portfolio companies. We acknowledge that our empirical analysis differs in the way we measured distinctiveness and the types of categories used for the calculation of the variables. Our analysis of corporate venture capitalists' behavior can thus be a piece of evidence that makes it necessary for future researchers to analyze the proposed change in the strategies of corporate venture capitalists over time in more detail.

Our findings also have implications for distinctive new ventures: If corporate venture capitalists with a core business related to the startup's business model exist, these investors are more likely to provide funding than independent venture capitalists. Therefore, a startup's choice for or against a corporate venture capitalists instead of an independent venture capitalist should take into account the different preferences of these two types of investors.

Chapter 8

Conclusion

Venture capital has been one of the most important funding sources for technology and non-technology new ventures in their growth phases for the last decades. The venture capital domain has also been widely researched and used as a setting to test new conceptual propositions. However, a gap remains in explaining the factors that affect which venture capitalists invest in which startups, despite this question being essential for researchers and practitioners alike. For researchers, it is crucial to understand the dimensions that affect the matching between venture capitalists and startups, as it can have major consequences on the way the two organizations work together, what effect venture capital funding has on the new venture, and, ultimately, on the success of the new venture and the venture capitalist. As the title of this dissertation suggests, we have aimed to find the perfect match between venture capitalists and startups by explaining the role of categorical fit. Multiple dimensions that play important roles in the venture capitalist - startup matching process exist. Recently, the role of category labels for the matching process has come to researchers' attention (Pontikes, 2012). Building on prior research, we set out to answer two main research questions: Firstly, what affects a venture capitalist's choice regarding a startup's novelty as measured by its distinctiveness? Secondly, how does categorical distance affect the matching of startups and venture capitalists and what mitigates this distance's effect? In this dissertation, we have also explored how venture capitalist's past portfolio strategies and their status gained from their co-syndication networks affect their preferences in the matching process with startups within each of the research questions.

After a brief synthesis of the empirical findings, we discuss the theoretical implications and contributions of the results and our discussions. This is followed by a description of the implications for practitioners. We then outline the limitations of this dissertation and develop areas for further research before concluding this chapter.

8.1 Synthesis of empirical findings

The first part of the empirical analysis of 10,576 funding rounds by 5,826 startups demonstrated support for the hypotheses that different venture capitalist demonstrate varying preference for distinctive and novel startups. Specifically, the more experience managing the risk of distinctive new ventures the venture capitalist has, the more highly diversified the portfolio is, and the higher his or her status, the more likely he or she is to invest in distinctive new ventures. These distinctive new ventures are associated with higher risk, but also potentially higher returns, due to their disruptive potential.

The second part of our analysis is concerned with the effect of categorical distance on the matching behavior. We have found strong support for the expected negative effect of categorical distance between venture capitalists and startups on the probability of an investment. This effect is mitigated when venture capitalists have a diversified portfolio or high status. We did not find support for an effect of experience with distinctive new ventures on the categorical distance.

8.2 Theoretical contributions and implications

The theoretical framework and empirical analysis in this dissertation contribute to existing research with five theoretical implications.

Firstly, it is generally accepted that organizations require the legitimacy that comes from being embedded in their environment with existing beliefs, norms, and categories (Scott and Meyer, 1983) and this need can lead to isomorphism (DiMaggio and Powell, 1983), that is, highly similar organizations. It has been shown that less legitimized organizations are discounted by various audiences (Zuckerman, 1999; Pontikes, 2012). Contrary to these general beliefs, our findings suggest that the need for legitimacy and the discount varies by audience. Boundary-spanning organizations are generally considered less legitimized or more distinctive. While, on the one hand, it is particularly important for new ventures to gain legitimacy (Aldrich and Fiol, 1994; Navis and Glynn, 2011), as they have fewer alternative ways of proving themselves, it might pay off for them to be boundary-spanning in order to be more innovative and taken into consideration by certain audiences. It follows that in future research, the organizational need for legitimacy should at least be questioned with regard to the setting, audience, and type of organization.

Secondly, we are contributing by supporting previous researchers who have argued that venture capitalists are different from other audiences (Pontikes, 2012; Wry, Lounsbury, and Jennings, 2014; Navis and Glynn,

2011). Our contributions to this are twofold: On the one hand, we provide support based on our large-scale, cross-industry analysis, whereas prior studies have focused on individual industries or niches. Our results thus increase the robustness of the findings. On the other hand, we show that not every venture capitalist has the same preferences. Rather, there is great heterogeneity among them, which has implications for their preferences: Depending on their portfolio strategies regarding average distinctiveness and diversification, or their co-investment network, that is, derived status, they may prefer distinctive new ventures, or the effect of categorical distance can be mitigated.

Thirdly, we are introducing a new measure of distance between two organizations based on market categories: categorical distance. Various distance measures, like geographical distance, cultural distance, and industry distance, are widespread and broadly used. Our measure of categorical distance lies somewhere between industry distance and cultural distance, as it compares the cultural embeddedness of organizations based on market categories. We have controlled for geographical and industry distance in our empirical analysis and categorical distance still has a strongly significant negative effect on investment probability. We thus encourage other researchers to use this measure in addition to other distance measures between organizations and to explore the effect of categorical distance on alternative organizational aspects.

Fourth, we are answering the research call made by Hannan, Goldberg, and Kovács (2016) to test their two-type framework of cultural preferences in the startup - venture capital setting. Our analysis demonstrates that the two dimensions of consumption behavior, atypicality and variety, can also be found in the venture capital setting, but the interpretation is different. On the one hand, some venture capitalists prefer distinctive or atypical new ventures. On the other hand, contrary to the consumers in the analysis of Hannan, Goldberg, and Kovács (2016), venture capitalists with diversified portfolios, the equivalent to consumers with a preference for variety, prefer distinctive new ventures, where the authors have found that these types of consumers prefer non-boundary-spanning objects.

Fifth, the use of a very recent, cross-industry dataset has made our findings highly relevant. Many of the influential papers most relevant to our analysis on the venture capital industry have analyzed data that is more than 10 to 15 years old (e.g., Sorenson and Stuart, 2001; Pontikes, 2012; Wry, Lounsbury, and Jennings, 2014). The startup domain is known for its fast-paced and ever-changing environment, which re-invents itself quickly. Not only do the types of new ventures and business models change, but also the technological and informational instruments available. This becomes very obvious when we consider communication tools: In the last 15 years,

the quality of high-speed Internet has significantly improved while the cost has significantly dropped, enabling, among other things, video conferencing, cloud-based storage, and new collaboration tools. This has led to a decreased or at least altered need for geographical proximity. Similarly, due to an easier access to information through new technologies and specialized data providers, the work of venture capitalists changes. We thus strongly believe, that research also needs to keep up with the rapidly changing real world by using more recent data to test certain relationships. As outlined above, during the development of this dissertation, we updated our underlying sample multiple times to integrate new information. As a side finding, our recent dataset, which also so far has barely been analyzed (for exceptions see Alexy et al., 2012; Ter Wal et al., 2016), provides evidence that some of the older findings also hold true today. For instance, as expected, all our distance measures, for example, geographical and industry distance, still have a negative impact on investment probability, as found by Sorenson and Stuart (2001) in a much older dataset.

Furthermore, we have paid great attention to making our research easily reproducible. Not only have we tried to be very explicit about our dealing with exceptions, and about which software-based implementations were used, but we are also able to reproduce the results of our research based on an R programming code at any time. While including the complete programming code in this dissertation's appendix would go beyond the page-scope of this dissertation and be irrelevant for most readers, we are able to provide the complete code upon request.

8.3 Practical implications

Explaining the behavior of venture capitalists is particularly useful for startups and respective entrepreneurs when searching for a venture capitalist. It is slightly less relevant for venture capitalists, as explaining how venture capitalists' preferences differ adds limited value for them. Our work creates three main insights for practitioners.

We set out with the question of what the perfect match is. We have added to the existing body of knowledge further dimensions that are critical in the venture capitalist - startup matching process. For a start, entrepreneurs need to be aware that the categories, which startups are associated with, are indicators of novelty and distinctiveness, which venture capitalists utilize. Similarly, categorical distance, that is, how different the market categories of portfolios companies are when compared to the focal startup's categories, matters. Avoiding repetition of the empirical results, we stress that entrepreneurs should approach venture capitalists who have the greatest chance of investing in them based on the venture capitalist's

portfolio distinctiveness, portfolio diversification, and status. This reduces the time spent searching for a suitable venture capitalist and leaves more time for other value-adding activities.

The second insight for entrepreneurs is that they can make themselves more or less attractive to or even attract certain audiences with their associated categories. Which categories the new venture is to be associated with should be a conscious decision. If, for instance, a new venture is searching for a high-status investor, it might decide to stress its distinctiveness and novelty by using boundary-spanning categories, attributes, or labels. Various means can be used to affect audience perceptions, for instance, through informational resources, such as Crunchbase, pitch documents, advertisements, news articles, or identity statements. The active use of relevant market categories can thus be a signal to various audiences and should be part of a new venture's overall strategy.

While we did not test whether startups also choose their venture capitalists, if one believes they do, our findings also have implications for venture capitalists: an investor needs to gain status quickly if he or she wants to have access to distinctive or more categorically distant startups. For instance, participation in syndicates of founding rounds with high-status investors can increase the own status by being connected to other well connected venture capitalists. This eases the access to information about potential investment opportunities, for example, in distinctive new ventures, or helps them to evaluate more distinctive new ventures. Thus, our results support the need for high status for venture capitalists.

8.4 Limitations and future research

We see a number of areas for future research. Some of them are based on the shortcomings of our empirical analysis, while others represent natural continuations of our research.

The crucial element of our empirical analysis is the categorization of startups. Each new venture is placed in one or more categories, which are the basis for four of the independent variables. The categorization is done by contributors to the Crunchbase platform. This includes Crunchbase employees, venture capitalists, startups themselves, and other individuals. Not knowing the author of the categorization can lead to a number of problems. Firstly, given that we are not able to differentiate between the originators of the assigned categories, we are not aware of the intentions of a given contributor. A startup's employee, who would be familiar with how the startup views itself, might differ in his or her categorization from the categorization of a customer, who only knows the company from

advertisements or friends' experiences. An investor, for instance, a venture capitalist, who wants to spread information about his investment, may categorize the same startup differently. Secondly, not only might different intentions be a problem, but varying quality in the data entry might also be an issue. For instance, someone who enters the information for multiple companies might be more consistent in his or her categorization than someone who only rarely contributes information to the database. Some categories are more general, whereas others are more specific, and the usage might depend on the contributor. While we were not able to control for the quality of the data entry, we feel confident that these errors are uncorrelated to each other such that no groups exist. Thirdly, the categorization scheme changed over time, due to changes in the database. Our controls for the year of funding should capture this effect, but we cannot fully rule out the possibility that changes to the database affected the categorization.

Nevertheless, future research could overcome the aforementioned problems caused by not knowing the author of the categorization; a different categorization scheme could be used. For instance, one could use the new venture's homepage and extract relevant categories to obtain self-assigned categories. In our view, most interesting for future research would be the analysis of the categories that others have assigned to a new venture, but with quality control. One approach would be to extract the categories from newspaper articles, as these have quality control from the newspaper publisher and the author of the article. The categories assigned by others are most relevant for external audiences like venture capitalists or consumers. This approach would differ from the one used by Pontikes (2012), who uses the identity statements of the companies and thus considers only self-assigned categories.

We have developed a measure of novelty, distinctiveness, for new ventures based on the associated market categories and have found different preferences among venture capitalists. Due to the research design, we are not able to make any predictions about the general desirability of distinctiveness for new ventures. Our empirical analysis only shows, for instance, that the likelihood of receiving funding from a high-status investor is higher for more distinctive new ventures. The positive quality signal could lead to superior performance of these startups. Future research would have to explore whether more distinctive new ventures are more successful, for example, in terms of their receiving of funding, their valuation, their innovation power, or their growth. Another angle to consider in future research would be how a new venture's distinctiveness evolves over time, that is, with the lifecycle, and how the venture capitalist's preferences depend on the phase the new venture is currently in. In addition, the desirability of distinctiveness and venture capitalists' preferences might depend on the country. Our

dataset was limited to startups from the US. While the US is the largest market for venture capital, we acknowledge that there might be different perceptions present in other countries. A comparison of differences across countries would thus be a natural extension of our study.

Future studies could continue to explore the perceptions of venture capitalists by gathering primary data about how category labels affect their decisions. For example, interviews or experiments could further increase our understanding of the mechanics that have led to the empirical results of this dissertation. So far, our measure of distinctiveness and categorical distance, despite being strongly embedded in the literature, remains a theoretical construct.

One of our theoretical contributions is the development of a categorical distance measure. Future studies would have to investigate the role of categorical distance on other parts of the venture capitalist - startup relationship. For instance, how is the relationship affected by categorical distance, and how good is the knowledge transfer? Furthermore, researchers could examine the effects of categorical distance between organizations outside of the venture capital setting.

The supplementary analysis of differences between corporate and independent venture capitalists resulted in findings contrary to an earlier study by Pontikes (2012). Future researchers should explore in depth whether corporate venture capitalists change their investment strategies as we have proposed or whether the diverging empirical findings are due to differences between corporate venture capitalists from different industries.

8.5 Closure

This dissertation contributes significantly to the body of knowledge in the venture capital domain. We have summarized in a structured approach the current empirical analysis with a focus on network-related variables as the basis for our own empirical work. We have succeeded in shedding further light on the matching process between venture capitalists and startups by introducing a measure of a startup's distinctiveness and categorical distance, and by explaining how investment probabilities are moderated by a venture capitalist's attributes. Our empirical analysis shows that trade-offs between legitimacy versus distinctiveness and challenges to overcome categorical distance exist, but, ultimately, it is up to both sides of the venture capitalist - startup relationship to find the perfect match for individual preferences. We are confident that researchers in entrepreneurship and management, sociologist, and practitioners like venture capitalists and founders can build upon our findings in their future work.

Appendix A

Detailed regression models

On the next three pages we have included the results of the regression models including rare events correction and the full model with all independent variables and interaction terms.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
(Intercept)	4.99*** (0.35)	4.91*** (0.36)	5.68*** (0.38)	4.53*** (0.37)	5.59*** (0.35)	5.62*** (0.40)	4.96*** (0.36)	5.47*** (0.39)	5.32*** (0.37)	6.02*** (0.36)	5.93*** (0.40)	6.60*** (0.44)
State dummy CA	-0.26*** (0.03)	-0.25*** (0.03)	-0.26*** (0.03)	-0.26*** (0.03)	-0.25*** (0.03)	-0.25*** (0.03)	-0.26*** (0.03)	-0.27*** (0.03)	-0.27*** (0.03)	-0.26*** (0.03)	-0.26*** (0.03)	-0.26*** (0.03)
State dummy MA	-0.23*** (0.04)	-0.22*** (0.04)	-0.23*** (0.04)	-0.22*** (0.04)	-0.23*** (0.04)	-0.22*** (0.04)	-0.20*** (0.04)	-0.21*** (0.04)	-0.21*** (0.04)	-0.22*** (0.04)	-0.21*** (0.04)	-0.21*** (0.04)
State dummy NY	-0.15*** (0.04)	-0.13*** (0.04)	-0.15*** (0.04)	-0.14*** (0.04)	-0.14*** (0.04)	-0.13*** (0.04)	-0.15*** (0.04)	-0.16*** (0.04)	-0.15*** (0.04)	-0.16*** (0.04)	-0.14*** (0.04)	-0.14*** (0.04)
Startup age	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Investor count dummy 2	0.05 (0.04)	0.03 (0.04)	0.05 (0.04)	0.05 (0.04)	0.04 (0.04)	0.03 (0.04)	0.03 (0.04)	0.04 (0.04)	0.04 (0.04)	0.03 (0.04)	0.02 (0.04)	0.02 (0.04)
Investor count dummy 3	0.18*** (0.04)	0.17*** (0.04)	0.18*** (0.04)	0.18*** (0.04)	0.17*** (0.04)	0.17*** (0.04)	0.16*** (0.04)	0.17*** (0.04)	0.17*** (0.04)	0.15*** (0.04)	0.16*** (0.04)	0.16*** (0.04)
Investor count dummy 4	0.25*** (0.04)	0.25*** (0.04)	0.25*** (0.04)	0.26*** (0.04)	0.25*** (0.04)	0.25*** (0.04)	0.24*** (0.04)	0.25*** (0.04)	0.24*** (0.04)	0.23*** (0.04)	0.23*** (0.04)	0.23*** (0.04)
Investor count dummy 5	0.38*** (0.05)	0.38*** (0.05)	0.38*** (0.05)	0.39*** (0.05)	0.37*** (0.05)	0.38*** (0.05)	0.37*** (0.05)	0.38*** (0.05)	0.38*** (0.05)	0.36*** (0.05)	0.36*** (0.05)	0.36*** (0.05)
Investor count dummy 6	0.49*** (0.05)	0.49*** (0.05)	0.49*** (0.05)	0.49*** (0.05)	0.48*** (0.05)	0.49*** (0.05)	0.48*** (0.05)	0.49*** (0.05)	0.48*** (0.05)	0.47*** (0.05)	0.47*** (0.05)	0.47*** (0.05)
Investor count dummy 7	0.56*** (0.06)	0.58*** (0.06)	0.57*** (0.06)	0.57*** (0.06)	0.57*** (0.06)	0.58*** (0.06)	0.56*** (0.06)	0.56*** (0.06)	0.56*** (0.06)	0.55*** (0.06)	0.55*** (0.06)	0.56*** (0.06)
Investor count dummy 8	0.68*** (0.07)	0.68*** (0.07)	0.68*** (0.07)	0.69*** (0.07)	0.67*** (0.07)	0.68*** (0.07)	0.68*** (0.07)	0.68*** (0.07)	0.68*** (0.07)	0.65*** (0.07)	0.66*** (0.07)	0.66*** (0.07)
Investor count dummy 9	0.84*** (0.09)	0.85*** (0.09)	0.84*** (0.09)	0.85*** (0.09)	0.84*** (0.09)	0.85*** (0.09)	0.85*** (0.09)	0.84*** (0.09)	0.84*** (0.09)	0.83*** (0.09)	0.83*** (0.09)	0.83*** (0.09)
Investor count dummy 10	0.92*** (0.10)	0.92*** (0.10)	0.93*** (0.10)	0.93*** (0.10)	0.92*** (0.10)	0.92*** (0.10)	0.93*** (0.10)	0.93*** (0.10)	0.94*** (0.10)	0.91*** (0.10)	0.91*** (0.10)	0.92*** (0.10)
Investor count dummy 11	1.04*** (0.11)	1.06*** (0.11)	1.05*** (0.11)	1.05*** (0.11)	1.06*** (0.11)	1.06*** (0.11)	1.06*** (0.11)	1.05*** (0.11)	1.06*** (0.11)	1.05*** (0.11)	1.06*** (0.11)	1.06*** (0.11)
Investor count dummy 12	0.51*** (0.13)	0.51*** (0.13)	0.51*** (0.13)	0.51*** (0.13)	0.51*** (0.13)	0.51*** (0.13)	0.49*** (0.13)	0.50*** (0.13)	0.48*** (0.13)	0.49*** (0.13)	0.48*** (0.13)	0.48*** (0.13)
Investor count dummy 13	0.82*** (0.15)	0.83*** (0.15)	0.83*** (0.15)	0.84*** (0.15)	0.82*** (0.15)	0.84*** (0.15)	0.87*** (0.15)	0.84*** (0.15)	0.86*** (0.15)	0.82*** (0.15)	0.84*** (0.15)	0.84*** (0.15)
Investor count dummy 14	1.06*** (0.18)	1.06*** (0.18)	1.07*** (0.18)	1.08*** (0.18)	1.05*** (0.18)	1.06*** (0.18)	1.07*** (0.18)	1.08*** (0.18)	1.11*** (0.18)	1.05*** (0.18)	1.08*** (0.18)	1.08*** (0.18)
Investor count dummy 15	1.03*** (0.19)	1.03*** (0.19)	1.02*** (0.19)	1.03*** (0.19)	1.01*** (0.19)	1.03*** (0.19)	0.98*** (0.20)	0.98*** (0.19)	0.99*** (0.20)	0.97*** (0.19)	0.99*** (0.20)	0.99*** (0.20)
Investor count dummy 16	1.18*** (0.22)	1.15*** (0.22)	1.18*** (0.22)	1.18*** (0.22)	1.15*** (0.22)	1.15*** (0.22)	1.07*** (0.22)	1.11*** (0.22)	1.12*** (0.22)	1.09*** (0.22)	1.09*** (0.22)	1.09*** (0.22)
Investor count dummy 17	0.72* (0.30)	0.70* (0.30)	0.73* (0.30)	0.72* (0.30)	0.71* (0.30)	0.70* (0.30)	0.64* (0.30)	0.69* (0.30)	0.66* (0.30)	0.66* (0.30)	0.64* (0.30)	0.64* (0.30)
Investor count dummy 18	2.50* (1.07)	2.44* (1.07)	2.50* (1.07)	2.51* (1.07)	2.43* (1.08)	2.45* (1.07)	2.34* (1.07)	2.43* (1.07)	2.38* (1.08)	2.31* (1.07)	2.27* (1.07)	2.27* (1.07)
Investor count dummy 19	1.65*** (0.33)	1.68*** (0.33)	1.67*** (0.33)	1.71*** (0.33)	1.65*** (0.33)	1.70*** (0.33)	1.60*** (0.33)	1.59*** (0.33)	1.61*** (0.33)	1.56*** (0.33)	1.58*** (0.33)	1.59*** (0.33)
Investor count dummy 20	1.29 (0.70)	1.26 (0.71)	1.31 (0.70)	1.29 (0.71)	1.29 (0.70)	1.27 (0.71)	1.34 (0.71)	1.37 (0.70)	1.33 (0.72)	1.39* (0.70)	1.36 (0.71)	1.37 (0.71)
Investor count dummy 21	0.79 (0.49)	0.80 (0.49)	0.80 (0.49)	0.83 (0.49)	0.78 (0.49)	0.81 (0.49)	0.65 (0.49)	0.69 (0.49)	0.63 (0.49)	0.61 (0.48)	0.57 (0.48)	0.57 (0.48)
Investor count dummy 22	0.37 (0.43)	0.38 (0.43)	0.39 (0.43)	0.38 (0.44)	0.38 (0.43)	0.38 (0.43)	0.31 (0.43)	0.34 (0.43)	0.33 (0.44)	0.33 (0.43)	0.31 (0.43)	0.32 (0.43)
Investor count dummy 23	1.39** (0.43)	1.36* (0.43)	1.41** (0.43)	1.39* (0.44)	1.39** (0.43)	1.37* (0.43)	1.37* (0.43)	1.42** (0.43)	1.37* (0.44)	1.39* (0.43)	1.35* (0.43)	1.35* (0.43)

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Industry distance	-1.52*** (0.07)	-1.61*** (0.07)	-1.53*** (0.07)	-1.68*** (0.07)	-1.46*** (0.07)	-1.60*** (0.07)	-1.09*** (0.08)	-1.13*** (0.08)	-1.02*** (0.08)	-1.01*** (0.08)	-0.92*** (0.08)	-0.92*** (0.08)
Prior investment dummy	5.84*** (0.17)	5.80*** (0.17)	5.83*** (0.17)	5.83*** (0.17)	5.81*** (0.17)	5.79*** (0.17)	5.82*** (0.17)	5.86*** (0.17)	5.84*** (0.17)	5.83*** (0.17)	5.81*** (0.17)	5.80*** (0.17)
Mean affiliation	0.04*** (0.00)	0.05*** (0.00)	0.04*** (0.00)	0.04*** (0.00)	0.05*** (0.00)	0.05*** (0.00)	0.05*** (0.00)	0.04*** (0.00)	0.04*** (0.00)	0.05*** (0.00)	0.05*** (0.00)	0.05*** (0.00)
Affiliate geographic distance	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Affiliate industry distance	-3.73*** (0.08)	-3.78*** (0.09)	-3.73*** (0.08)	-3.70*** (0.08)	-3.81*** (0.09)	-3.78*** (0.09)	-3.73*** (0.09)	-3.71*** (0.08)	-3.69*** (0.08)	-3.86*** (0.09)	-3.83*** (0.09)	-3.83*** (0.09)
VC age	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
General experience	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)
Industry experience	0.01*** (0.00)	0.02*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.02*** (0.00)	0.02*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.02*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Startup's distinctiveness		-0.16* (0.06)	-1.51*** (0.32)	-0.64** (0.23)	-0.28*** (0.09)	-1.78*** (0.38)	-0.11 (0.06)	-0.11 (0.06)	-0.09 (0.06)	-0.10 (0.06)	-0.08 (0.06)	-1.63*** (0.38)
Portfolio distinctiveness		-0.14 (0.13)	-1.45*** (0.33)			-1.26*** (0.34)	-0.17 (0.13)	-0.67 (0.39)			0.09 (0.39)	-1.02* (0.50)
Portfolio diversification		0.03*** (0.00)		0.02** (0.01)		0.02** (0.01)	0.03*** (0.00)		-0.03*** (0.01)		-0.02** (0.01)	-0.03** (0.01)
VC status		-0.14*** (0.01)			-0.17*** (0.01)	-0.16*** (0.01)	-0.14*** (0.01)			-0.31*** (0.02)	-0.31*** (0.02)	-0.33*** (0.02)
Startup's distinctiveness x Portfolio distinctiveness			2.87*** (0.66)			2.42*** (0.67)						2.45*** (0.68)
Startup's distinctiveness x Portfolio diversification				0.02* (0.01)		0.02 (0.01)						0.02 (0.01)
Startup's distinctiveness x VC status					0.06* (0.02)	0.05 (0.02)						0.04 (0.02)
Categorical distance							-0.03*** (0.00)	-0.03*** (0.01)	-0.09*** (0.01)	-0.04*** (0.00)	-0.09*** (0.01)	-0.09*** (0.01)
Categorical distance x Portfolio distinctiveness								0.02 (0.02)			-0.01 (0.02)	-0.01 (0.02)
Categorical distance x Portfolio diversification									0.00*** (0.00)		0.00*** (0.00)	0.00*** (0.00)
Categorical distance x VC status										0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
AIC	49559.09	49148.53	49539.52	49401.47	49284.26	49131.36	48950.76	49449.80	49093.82	49046.57	48711.44	48696.85
BIC	50106.15	49731.46	50113.48	49975.43	49858.22	49741.20	49542.66	50032.73	49676.76	49629.50	49330.25	49342.56
Log Likelihood	-24718.54	-24509.26	-24705.76	-24636.73	-24578.13	-24497.68	-24409.38	-24659.90	-24481.91	-24458.28	-24286.72	-24276.42
Deviance	49437.09	49018.53	49411.52	49273.47	49156.26	48995.36	48818.76	49319.80	48963.82	48916.57	48573.44	48552.85
Num. obs.	58000	58000	58000	58000	58000	58000	58000	58000	58000	58000	58000	58000

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

* The table shows the results of the logistic regression models with correction of coefficients for rare events. Standard errors are shown in parentheses. The unit of analysis is the venture capitalist - startup funding round dyad and the sample includes 58,000 observations from the period 2005 - 2015. The observations are half realized investments and half unrealized investments. The dependent variable *dummy actual data* is binary and equals 1 when the observation is a realized investment.

TABLE A.1: Regression output models with rare events correction*

Appendix B

Software used in research project

The complete analysis presented in this dissertation was done in R (R Core Team, 2016), version 3.3.1 from 21. June 2016, on a 64-bit machine running Windows 8. We used a number of packages in R, that were programmed by contributors to the open source environment. We have used the following packages for reformatting, analysis, and presentation of results. The list is in no particular order.

For reformatting of data and network analysis we used the packages *dplyr* (Wickham and Francois, 2016), *plyr* (Wickham, 2011), *data.table* (Dowle et al., 2015), *Matrix* (Bates and Maechler, 2016), *igraph* (Csardi and Nepusz, 2006), *tm* (Feinerer and Hornik, 2015; Feinerer, Hornik, and Meyer, 2008), *tnet* (Opsahl, 2009), *reshape* (Wickham, 2007), *reshape2* (Wickham, 2007), and *zoo* (Zeileis and Grothendieck, 2005).

We used the *doParallel* (Analytics and Weston, 2015a) package for parallel processing and *foreach* (Analytics and Weston, 2015b) for parallel loops. For statistical analysis the relevant packages were *Zelig* (Choirat et al., 2016; Imai, King, and Lau, 2008), *multiwayvcov* (Graham, Arai, and Hagstroemer, 2016), and *lmtest* (Zeileis and Hothorn, 2002).

For visualization like plotting, graphs, and descriptive statistics we used the packages *ggplot2* (Wickham, 2009), *effects* (Fox, 2003; Fox and Hong, 2009), *interplot* (Solt and Hu, 2015), *doBy* (Hoejsgaard and Halekoh, 2016), *stargazer* (Hlavac, 2015), *texreg* (Leifeld, 2013), *pastecs* (Grosjean and Ibanez, 2014), *grid* (R Core Team, 2016), *picante* (Kembel et al., 2010), and *scales* (Wickham, 2016). For geographic information retrieval, i.e., coordinates, and plotting the relevant packages are *geosphere* (Hijmans, 2016), *ggmap* (Kahle and Wickham, 2013), *maptools* (Bivand and Lewin-Koh, 2017), and *maps* (Becker et al., 2016). Furthermore, we used the packages *xlsx* (Dragulescu, 2014), and *survival* (Therneau, 2015; Therneau and Grambsch, 2000).

Appendix C

Supplementary analysis corporate venture capital

On the following page, the results of the logistic regression models with rare-events correction for the effect of corporate venture capital as described in section 7.3 are presented.

	Model 12	Model 13	Model 14	Model 15
(Intercept)	3.52*** (0.39)	3.55*** (0.39)	3.44*** (0.39)	3.47*** (0.39)
State dummy CA	-0.31*** (0.03)	-0.31*** (0.03)	-0.31*** (0.03)	-0.31*** (0.03)
State dummy MA	-0.17*** (0.05)	-0.17*** (0.05)	-0.17*** (0.05)	-0.17*** (0.05)
State dummy NY	-0.14** (0.04)	-0.14** (0.04)	-0.14** (0.04)	-0.14** (0.04)
Startup age	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Raised amount (log)	0.16*** (0.01)	0.16*** (0.01)	0.16*** (0.01)	0.16*** (0.01)
Geographical distance (log)	-0.19*** (0.00)	-0.19*** (0.00)	-0.19*** (0.00)	-0.19*** (0.00)
Geographical distance no data dummy	0.13 (0.13)	0.14 (0.13)	0.17 (0.13)	0.17 (0.13)
Industry distance	-1.09*** (0.09)	-1.09*** (0.09)	-1.08*** (0.09)	-1.08*** (0.09)
Prior investment dummy	6.08*** (0.20)	6.09*** (0.20)	6.08*** (0.20)	6.09*** (0.20)
Mean affiliation	0.04*** (0.00)	0.04*** (0.00)	0.04*** (0.00)	0.04*** (0.00)
Affiliate geographic distance	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Affiliate industry distance	-3.50*** (0.09)	-3.50*** (0.09)	-3.50*** (0.09)	-3.50*** (0.09)
VC age	-0.01 (0.00)	-0.01 (0.00)	-0.01 (0.00)	-0.01 (0.00)
General experience	-0.00** (0.00)	-0.00** (0.00)	-0.00** (0.00)	-0.00** (0.00)
Industry experience	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Portfolio distinctiveness	-0.27 (0.17)	-0.27 (0.17)	-0.26 (0.17)	-0.26 (0.17)
Portfolio diversification	0.04*** (0.00)	0.04*** (0.00)	0.04*** (0.00)	0.04*** (0.00)
VC status	-0.13*** (0.01)	-0.13*** (0.01)	-0.13*** (0.01)	-0.13*** (0.01)
Startup's distinctiveness	0.02 (0.08)	-0.04 (0.08)	0.02 (0.08)	-0.04 (0.08)
Categorical distance	-0.03*** (0.00)	-0.03*** (0.00)	-0.03*** (0.00)	-0.03*** (0.00)
CVC dummy	0.38*** (0.05)	0.04 (0.14)	1.03*** (0.15)	0.67*** (0.20)
Startup's distinctiveness x CVC dummy		0.72** (0.27)		0.80** (0.27)
Categorical distance x CVC dummy			-0.03*** (0.01)	-0.03*** (0.01)
AIC	37,217	37,211	37,198	37,191
BIC	37,801	37,804	37,790	37,792
Log Likelihood	-18,541	-18,537	-18,531	-18,526
Deviance	37,083	37,075	37,062	37,053
Num. obs.	44,992	44,992	44,992	44,992

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

* The table shows the results of the logistic regression models with correction of coefficients for rare events. Standard errors are shown in parentheses. Year dummies, round number dummies, and investor count dummies are included in all models. The unit of analysis is the venture capitalist - startup funding round dyad and the subsample includes 44,992 observations from the period 2005 - 2015. The observations are half realized investments and half unrealized investments. The dependent variable *dummy actual data* is binary and equals 1 when the observation is a realized investment. The dummy variable *CVC dummy* equals 1 if the investor is categorized as *corporate venture capital*.

TABLE C.1: Regression output models supplementary analysis corporate venture capital*

Reference list

- Admati, Anat R. and Paul Pfleiderer (1994). "Robust financial contracting and the role of venture capitalists". In: *The Journal of Finance* 49.2, pp. 371–402.
- Aldrich, Howard E. and C. Marlene Fiol (1994). "Fools rush in? The institutional context of industry creation". In: *Academy of Management Review* 19.4, pp. 645–670.
- Aldrich, Howard E. and Diane Herker (1977). "Boundary spanning roles and organization structure". In: *Academy of Management Review* 2.2, pp. 217–230.
- Alexy, Oliver T., Joern H. Block, Philipp G. Sandner, and Anne L. J. Ter Wal (2012). "Social capital of venture capitalists and start-up funding". In: *Small Business Economics* 39.4, pp. 835–851.
- Alvarez–Garrido, Elisa and Gary Dushnitsky (2016). "Are entrepreneurial venture's innovation rates sensitive to investor complementary assets? Comparing biotech ventures backed by corporate and independent VCs". In: *Strategic Management Journal* 37.5, pp. 819–834.
- Amit, Raphael, James Brander, and Christoph Zott (1998). "Why do venture capital firms exist? Theory and canadian evidence". In: *Journal of Business Venturing* 13.6, pp. 441–466.
- Analytics, Revolution and Steve Weston (2015a). *doParallel: Foreach parallel adaptor for the 'parallel' package*.
- (2015b). *foreach: Provides foreach looping construct for R*.
- Arqué–Castells, Pere (2012). "How venture capitalists spur invention in Spain: Evidence from patent trajectories". In: *Research Policy* 41.5, pp. 897–912.
- Ashforth, Blake E. and Barrie W. Gibbs (1990). "The double-edge of organizational legitimation". In: *Organization Science* 1.2, pp. 177–194.
- Ashforth, Blake E. and Ronald H. Humphrey (1997). "The ubiquity and potency of labeling in organizations". In: *Organization Science* 8.1, pp. 43–58.
- Ashforth, Blake E. and Fred Mael (1989). "Social identity theory and the organization". In: *Academy of Management Review* 14.1, pp. 20–39.
- Audretsch, David B., Werner Bönte, and Prashanth Mahagaonkar (2012). "Financial signaling by innovative nascent ventures: The relevance of patents and prototypes". In: *Research Policy* 41.8, pp. 1407–1421.

- Audretsch, David B. and Maryann P. Feldman (2004). "Knowledge spillovers and the geography of innovation". In: *Handbook of regional and urban economics*. Ed. by J. Vernon Henderson and Jacques-François Thisse. Vol. 4. Handbooks in economics. Amsterdam and Boston: Elsevier North Holland, pp. 2713–2739.
- Autant-Bernard, Corinne, Pascal Billand, David Frachisse, and Nadine Massard (2007). "Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies". In: *Papers in Regional Science* 86.3, pp. 495–519.
- Bae, Jin-Hyun and Robert Salomon (2010). "Institutional distance in international business research". In: *The past, present and future of international business & management*. Ed. by Timothy M. Devinney, Torben Pedersen, and Laszlo Tihanyi. Vol. 23. Advances in international management. Bingley, UK: Emerald, pp. 327–349.
- Barney, Jay B., Lowell W. Busenitz, James O. Fiet, and Douglas D. Moesel (1996). "New venture teams' assessment of learning assistance from venture capital firms". In: *Journal of Business Venturing* 11.4, pp. 257–272.
- Barone, Michael J. (2005). "The interactive effects of mood and involvement on brand extension evaluations". In: *Journal of Consumer Psychology* 15.3, pp. 263–270.
- Bates, Douglas and Martin Maechler (2016). *Matrix: Sparse and dense matrix classes and methods*.
- Baum, Joel A. C. and Brian S. Silverman (2004). "Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups". In: *Journal of Business Venturing* 19.3, pp. 411–436.
- Becker, Richard A., Allen R. Wilks, Ray Brownrigg, Thomas P. Minka, and Alex Deckmyn (2016). *maps: Draw geographical maps*.
- Bell, Geoffrey G. and Akbar Zaheer (2007). "Geography, networks, and knowledge flow". In: *Organization Science* 18.6, pp. 955–972.
- Bengtsson, Ola and David H. Hsu (2015). "Ethnic matching in the U.S. venture capital market". In: *Journal of Business Venturing* 30.2, pp. 338–354.
- Benson, David and Rosemarie H. Ziedonis (2009). "Corporate venture capital as a window on new technologies: Implications for the performance of corporate investors when acquiring startups". In: *Organization Science* 20.2, pp. 329–351.
- Bergemann, Dirk and Ulrich Hege (1998). "Venture capital financing, moral hazard, and learning". In: *Journal of Banking & Finance* 22.6–8, pp. 703–735.
- Bernstein, Shai, Xavier Giroud, and Richard R. Townsend (2016). "The impact of venture capital monitoring". In: *The Journal of Finance* 71.4, pp. 1591–1622.

- Bertoni, Fabio, Massimo G. Colombo, and Luca Grilli (2011). "Venture capital financing and the growth of high-tech start-ups: Disentangling treatment from selection effects". In: *Research Policy* 40.7, pp. 1028–1043.
- Bettignies, Jean-Etienne de and Anne Duchêne (2015). "Product market competition and the financing of new ventures". In: *Management Science* 61.8, pp. 1849–1867.
- Bivand, Roger and Nicholas Lewin-Koh (2017). *maptools: Tools for reading and handling spatial objects*.
- Blau, Peter M. (1960). "Patterns of deviation in work groups". In: *Sociometry* 23.3, pp. 245–261.
- Bonacich, Phillip (1987). "Power and centrality: A family of measures". In: *American Journal of Sociology* 92.5, pp. 1170–1182.
- Bothner, Matthew S., Young-Kyu Kim, and Wonjae Lee (2015). "Primary status, complementary status, and organizational survival in the U.S. venture capital industry". In: *Social Science Research* 52, pp. 588–601.
- Bottazzi, Laura, Marco Da Rin, and Thomas Hellmann (2008). "Who are the active investors? Evidence from venture capital". In: *Journal of Financial Economics* 89.3, pp. 488–512.
- Brav, Alon and Paul A. Gompers (1997). "Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies". In: *The Journal of Finance* 52.5, pp. 1791–1821.
- Brewer, Marilynn B. (1991). "The social self: On being the same and different at the same time". In: *Personality and Social Psychology Bulletin* 17.5, pp. 475–482.
- Briley, Donnel A. and Robert S. Wyer (2002). "The effect of group membership salience on the avoidance of negative outcomes: Implications for social and consumer decisions". In: *Journal of Consumer Research* 29.3, pp. 400–415.
- Burago, Dmitri, Yuri D. Burago, and Sergei Ivanov (2001). *A course in metric geometry*. Vol. v. 33. Graduate studies in mathematics. Providence, RI: American Mathematical Society.
- Burt, Ronald S. (1992). *Structural holes: The social structure of competition*. Harvard University Press.
- (1997). "The contingent value of social capital". In: *Administrative Science Quarterly* 42.2, pp. 339–365.
- (2004). "Structural holes and good ideas". In: *American Journal of Sociology* 110.2, pp. 349–399.
- Burt, Ronald S., Martin Kilduff, and Stefano Tasselli (2013). "Social network analysis: Foundations and frontiers on advantage". In: *Annual Review of Psychology* 64, pp. 527–547.

- Busenitz, Lowell W., James O. Fiet, and Douglas D. Moesel (2004). "Reconsidering the venture capitalists' "value added" proposition: An interorganizational learning perspective". In: *Journal of Business Venturing* 19.6, pp. 787–807.
- Bygrave, William D. (1987). "Syndicated investments by venture capital firms: A networking perspective". In: *Journal of Business Venturing* 2.2, pp. 139–154.
- Bygrave, William D. and Jeffrey A. Timmons (1992). *Venture capital at the crossroads*. Harvard Business Press.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2012). "Robust inference with multiway clustering". In: *Journal of Business & Economic Statistics*.
- Carpenter, Mason A., Timothy G. Pollock, and Myleen M. Leary (2003). "Testing a model of reasoned risk-taking: Governance, the experience of principals and agents, and global strategy in high-technology IPO firms". In: *Strategic Management Journal* 24.9, pp. 803–820.
- Casamatta, Catherine (2003). "Financing and advising: optimal financial contracts with venture capitalists". In: *The Journal of Finance* 58.5, pp. 2059–2086.
- Casamatta, Catherine and Carole Haritchabalet (2007). "Experience, screening and syndication in venture capital investments". In: *Journal of Financial Intermediation* 16.3, pp. 368–398.
- Chan, Yuk-Shee (1983). "On the positive role of financial intermediation in allocation of venture capital in a market with imperfect information". In: *The Journal of Finance* 38.5, pp. 1543–1568.
- Chater, Nick and Paul M. B. Vitányi (2003). "The generalized universal law of generalization". In: *Journal of Mathematical Psychology* 47.3, pp. 346–369.
- Choirat, Christine, James Honaker, Kosuke Imai, Gary King, and Olivia Lau (2016). *Zelig: Everyone's statistical software*.
- Clercq, Dirk de and Dimo P. Dimov (2008). "Internal knowledge development and external knowledge access in venture capital investment Performance". In: *Journal of Management Studies* 45.3, pp. 585–612.
- Clercq, Dirk de and Harry J. Sapienza (2005). "When do venture capital firms learn from their portfolio companies?" In: *Entrepreneurship Theory and Practice* 29.4, pp. 517–535.
- Clercq, Dirk de, Harry J. Sapienza, and Akbar Zaheer (2008). "Firm and group influences on venture capital firms' involvement in new ventures". In: *Journal of Management Studies* 45.7, pp. 1169–1194.
- Cochrane, John H. (2005). "The risk and return of venture capital". In: *Journal of Financial Economics* 75.1, pp. 3–52.

- Cohen, Wesley M. and Daniel A. Levinthal (1990). "Absorptive capacity: A new perspective on learning and innovation". In: *Administrative Science Quarterly*, pp. 128–152.
- Cox, David R. (1958). "The regression analysis of binary sequences". In: *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 215–242.
- Csardi, Gabor and Tamas Nepusz (2006). "The igraph software package for complex network research". In: *InterJournal Complex Systems*, p. 1695.
- Cumming, Douglas J. (2005). "Agency costs, institutions, learning, and taxation in venture capital contracting". In: *Journal of Business Venturing* 20.5, pp. 573–622.
- (2006). "Adverse selection and capital structure: Evidence from venture capital". In: *Entrepreneurship Theory and Practice* 30.2, pp. 155–183.
- Cumming, Douglas J. and Na Dai (2013). "Why do entrepreneurs switch lead venture capitalists?" In: *Entrepreneurship Theory and Practice* 37.5, pp. 999–1017.
- Cumming, Douglas J., Daniel Schmidt, and Uwe Walz (2010). "Legality and venture capital governance around the world". In: *Journal of Business Venturing* 25.1, pp. 54–72.
- Dai, Na, Hoje Jo, and Sul Kassiech (2012). "Cross-border venture capital investments in Asia: Selection and exit performance". In: *Journal of Business Venturing* 27.6, pp. 666–684.
- Davila, Antonio, George Foster, and Mahendra Gupta (2003). "Venture capital financing and the growth of startup firms". In: *Journal of Business Venturing* 18.6, pp. 689–708.
- Devigne, David, Sophie Manigart, and Mike Wright (2016). "Escalation of commitment in venture capital decision making: Differentiating between domestic and international investors". In: *Journal of Business Venturing* 31.3, pp. 253–271.
- DiMaggio, Paul J. and Walter W. Powell (1983). "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields". In: *American Sociological Review* 48.2, pp. 147–160.
- eds. (1991). *The new institutionalism in organizational analysis*. Chicago: University of Chicago Press.
- Dimov, Dimo P. and Dirk de Clercq (2006). "Venture capital investment strategy and portfolio failure rate: A longitudinal study". In: *Entrepreneurship Theory and Practice* 30.2, pp. 207–223.
- Dimov, Dimo P. and Hana Milanov (2010). "The interplay of need and opportunity in venture capital investment syndication". In: *Journal of Business Venturing* 25.4, pp. 331–348.

- Dimov, Dimo P., Dean A. Shepherd, and Kathleen M. Sutcliffe (2007). "Requisite expertise, firm reputation, and status in venture capital investment allocation decisions". In: *Journal of Business Venturing* 22.4, pp. 481–502.
- Dobrev, Stanislav D., Tai-Young Kim, and Michael T. Hannan (2001). "Dynamics of niche width and resource partitioning". In: *American Journal of Sociology* 106.5, pp. 1299–1337.
- Dowle, Matt, Arun Srinivasan, Tom Short, Eduard Antonyan, Jan Gorecki, and Steve Lianoglou (2015). *data.table: Extension of data.frame*.
- Dowling, John and Jeffrey Pfeffer (1975). "Organizational legitimacy: Social values and organizational behavior". In: *The Pacific Sociological Review* 18.1, pp. 122–136.
- Dragulescu, Adrian A. (2014). *xlsx: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files*.
- Drogendijk, Rian and Arjen Slangen (2006). "Hofstede, Schwartz, or managerial perceptions? The effects of different cultural distance measures on establishment mode choices by multinational enterprises". In: *International Business Review* 15.4, pp. 361–380.
- Drover, Will, Matthew S. Wood, and Yves Fassin (2014). "Take the money or run? Investors' ethical reputation and entrepreneurs' willingness to partner". In: *Journal of Business Venturing* 29.6, pp. 723–740.
- Dubuisson, Marie-Pierre and Anil K. Jain (1994). "A modified Hausdorff distance for object matching". In: *12th international conference on pattern recognition*. Vol. 1, pp. 566–568.
- Dushnitsky, Gary (2012). "Corporate venture capital in the twenty-first century: An integral part of firms' innovation toolkit". In: *The Oxford Handbook of Venture Capital*. Ed. by Douglas J. Cumming. Oxford University Press.
- Dushnitsky, Gary and Michael J. Lenox (2005). "When do incumbents learn from entrepreneurial ventures? Corporate venture capital and investing firm innovation rates". In: *Research Policy* 34.5, pp. 615–639.
- Dushnitsky, Gary and Zur Shapira (2010). "Entrepreneurial finance meets organizational reality: Comparing investment practices and performance of corporate and independent venture capitalists". In: *Strategic Management Journal* 31.9, pp. 990–1017.
- Echols, Ann and Wenpin Tsai (2005). "Niche and performance: the moderating role of network embeddedness". In: *Strategic Management Journal* 26.3, pp. 219–238.
- Elfring, Tom and Willem Hulsink (2003). "Networks in entrepreneurship: The case of high-technology firms". In: *Small Business Economics* 21.4, pp. 409–422.

- Elsbach, Kimberly D. and Roderick M. Kramer (2003). "Assessing creativity in hollywood pitch meetings: Evidence for a dual-process model of creativity judgments". In: *Academy of Management Journal* 46.3, pp. 283–301.
- Engel, Dirk and Max Keilbach (2007). "Firm-level implications of early stage venture capital investment — An empirical investigation". In: *Journal of Empirical Finance* 14.2, pp. 150–167.
- Feinerer, Ingo and Kurt Hornik (2015). *tm: Text mining package*.
- Feinerer, Ingo, Kurt Hornik, and David Meyer (2008). "Text mining infrastructure in R". In: *Journal of Statistical Software* 25.5, pp. 1–54.
- Fennell, Mary L. and Jeffrey A. Alexander (1987). "Organizational boundary spanning in institutionalized environments". In: *Academy of Management Journal* 30.3, pp. 456–476.
- Fernhaber, Stephanie A. and Dan Li (2013). "International exposure through network relationships: Implications for new venture internationalization". In: *Journal of Business Venturing* 28.2, pp. 316–334.
- Fernhaber, Stephanie A. and Patricia P. McDougall-Covin (2009). "Venture capitalists as catalysts to new venture internationalization: The impact of their knowledge and reputation resources". In: *Entrepreneurship Theory and Practice* 33.1, pp. 277–295.
- Ferrary, Michel (2010). "Syndication of venture capital investment: The art of resource pooling". In: *Entrepreneurship Theory and Practice* 34.5, pp. 885–907.
- Fischer, Harald M. and Timothy G. Pollock (2004). "Effects of social capital and power on surviving transformational change: The case of initial public offerings". In: *Academy of Management Journal* 47.4, pp. 463–481.
- Fisher, Greg, Suresh Kotha, and Amrita Lahiri (2016). "Changing with the times: An integrated view of identity, legitimacy, and new venture life cycles". In: *Academy of Management Review* 41.3, pp. 383–409.
- Fitza, Markus A., Sharon F. Matusik, and Elaine Mosakowski (2009). "Do VCs matter? The importance of owners on performance variance in start-up firms". In: *Strategic Management Journal* 30.4, pp. 387–404.
- Fox, John (2003). "Effect displays in R for generalised linear models". In: *Journal of Statistical Software* 8.15, pp. 1–27.
- Fox, John and Jangman Hong (2009). "Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package". In: *Journal of Statistical Software* 32.1, pp. 1–24.
- Fulghieri, Paolo and Merih Sevilir (2009). "Size and focus of a venture capitalist's portfolio". In: *Review of Financial Studies* 22.11, pp. 4643–4680.
- Gaur, Ajai S. and Jane W. Lu (2007). "Ownership strategies and survival of foreign subsidiaries: Impacts of institutional distance and experience". In: *Journal of Management* 33.1, pp. 84–110.

- Gerasymenko, Violetta, Dirk de Clercq, and Harry J. Sapienza (2015). "Changing the business model: Effects of venture capital firms and outside CEOs on portfolio Company Performance". In: *Strategic Entrepreneurship Journal* 9.1, pp. 79–98.
- Ghoshal, Sumantra (1987). "Global strategy: An organizing framework". In: *Strategic Management Journal* 8.5, pp. 425–440.
- Glynn, Mary Ann (2000). "When cymbals become symbols: Conflict over organizational identity within a symphony orchestra". In: *Organization Science* 11.3, pp. 285–298.
- Goldberg, Amir, V. Govind Manian, Will Monroe, Christopher Potts, and Sameer B. Srivastava (2016). "Fitting in or standing out? The tradeoffs of structural and cultural embeddedness". In: *American Sociological Review* 81.6, pp. 1190–1222.
- Gompers, Paul A. (1995). "Optimal investment, monitoring, and the staging of venture capital". In: *Journal of Finance* 50.5, pp. 1461–1489.
- (1996). "Grandstanding in the venture capital industry". In: *Journal of Financial Economics* 42.1, pp. 133–156.
- Gompers, Paul A., Anna Kovner, and Josh Lerner (2009). "Specialization and success: Evidence from venture capital". In: *Journal of Economics & Management Strategy* 18.3, pp. 817–844.
- Gompers, Paul A. and Josh Lerner (2000). "The determinants of corporate venture capital success: Organizational structure, incentives, and complementarities". In: *Concentrated corporate ownership*. Chicago, IL: University of Chicago Press, pp. 17–54.
- (2001). "The venture capital revolution". In: *The Journal of Economic Perspectives* 15.2, pp. 145–168.
- Gorman, Michael and William A. Sahlman (1989). "What do venture capitalists do?" In: *Journal of Business Venturing* 4.4, pp. 231–248.
- Gould, Roger V. (2002). "The origins of status hierarchies: A formal theory and empirical test". In: *American Journal of Sociology* 107.5, pp. 1143–1178.
- Graham, Nathaniel, Mahmood Arai, and Bjoern Hagstroemer (2016). *multiwaycov: Multi-way standard error clustering*.
- Grosjean, Philippe and Frederic Ibanez (2014). *pastecs: Package for analysis of space-time ecological series*.
- Gruber, Martin J. and Jerold B. Warner (1977). "Bankruptcy costs: Some evidence". In: *The Journal of Finance* 32.2, pp. 337–347.
- Gu, Qian and Xiaohui Lu (2014). "Unraveling the mechanisms of reputation and alliance formation: A study of venture capital syndication in China". In: *Strategic Management Journal* 35.5, pp. 739–750.

- Gulati, Ranjay (1995). "Social structure and alliance formation patterns: A longitudinal analysis". In: *Administrative Science Quarterly* 40.4, pp. 619–652.
- Gulati, Ranjay and Monica C. Higgins (2003). "Which ties matter when? The contingent effects of interorganizational partnerships on IPO success". In: *Strategic Management Journal* 24.2, pp. 127–144.
- Guler, Isin and Mauro F. Guillen (2010). "Home country networks and foreign expansion: Evidence from the venture capital industry". In: *Academy of Management Journal* 53.2, pp. 390–410.
- Gupta, Anil K. and Harry J. Sapienza (1992). "Determinants of venture capital firms' preferences regarding the industry diversity and geographic scope of their investments". In: *Journal of Business Venturing* 7.5, pp. 347–362.
- Hall, John and Charles W. Hofer (1993). "Venture capitalists' decision criteria in new venture evaluation". In: *Journal of Business Venturing* 8.1, pp. 25–42.
- Hallen, Benjamin L. (2008). "The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments?" In: *Administrative Science Quarterly* 53.4, pp. 685–718.
- Hallen, Benjamin L., Riitta Katila, and Jeff D. Rosenberger (2014). "How do social defenses work? A resource-dependence lens on technology ventures, venture capital investors, and corporate relationships". In: *Academy of Management Journal* 57.4, pp. 1078–1101.
- Hannan, Michael T., Amir Goldberg, and Balázs Kovács (2016). "What does it mean to span cultural boundaries? Variety and atypicality in cultural consumption". In: *American Sociological Review* 81.2, pp. 215–241.
- Harrell, Frank E. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis*. Second edition. Springer Series in Statistics. Cham: Springer.
- Hellmann, Thomas and Manju Puri (2002). "Venture capital and the professionalization of start-up firms: Empirical evidence". In: *The Journal of Finance* 57.1, pp. 169–197.
- Helmreich, James E. (2016). "Regression modeling strategies with applications to linear models, logistic and ordinal regression and survival analysis". In: *Journal of Statistical Software* 70.1, pp. 1–3.
- Higgins, Monica C. and Ranjay Gulati (2003). "Getting off to a good start: The effects of upper echelon affiliations on underwriter prestige". In: *Organization Science* 14.3, pp. 244–263.
- Hijmans, Robert J. (2016). *geosphere: Spherical trigonometry*.

- Hite, Julie M. and William S. Hesterly (2001). "The evolution of firm networks: From emergence to early growth of the firm". In: *Strategic Management Journal* 22.3, pp. 275–286.
- Hlavac, Marek (2015). *stargazer: Well-formatted regression and summary statistics tables*. Cambridge, USA.
- Hochberg, Yael V., Alexander Ljungqvist, and Yang Lu (2007). "Whom you know matters: Venture capital networks and investment performance". In: *The Journal of Finance* 62.1, pp. 251–301.
- Hoejsgaard, Soeren and Ulrich Halekoh (2016). *doBy: Groupwise statistics, LSmeans, linear contrasts, utilities*.
- Hoenen, Sebastian, Christos Kolympiris, Wilfred Schoenmakers, and Nicholas Kalaitzandonakes (2014). "The diminishing signaling value of patents between early rounds of venture capital financing". In: *Research Policy* 43.6, pp. 956–989.
- Hofstede, Geert (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: SAGE.
- Hogg, Michael A. and Deborah J. Terry (2000). "Social identity and self-categorization processes in organizational contexts". In: *Academy of Management Review* 25.1, pp. 121–140.
- Hollander, Edwin P. (1958). "Conformity, status, and idiosyncrasy credit". In: *Psychological Review* 65.2, pp. 117–127.
- (1960). "Competence and conformity in the acceptance of influence". In: *The Journal of Abnormal and Social Psychology* 61.3, pp. 365–369.
- Hopp, Christian and Christian Lukas (2014). "A signaling perspective on partner selection in venture capital syndicates". In: *Entrepreneurship Theory and Practice* 38.3, pp. 635–670.
- Hsu, David H. (2004). "What do entrepreneurs pay for venture capital affiliation?" In: *The Journal of Finance* 59.4, pp. 1805–1844.
- (2006). "Venture capitalists and cooperative start-up commercialization strategy". In: *Management Science* 52.2, pp. 204–219.
- Hsu, Greta, Michael T. Hannan, and Özgecan Koçak (2009). "Multiple category memberships in markets: An integrative theory and two empirical tests". In: *American Sociological Review* 74.1, pp. 150–169.
- Hung, Wen-Liang and Miin-Shen Yang (2004). "Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance". In: *Pattern Recognition Letters* 25.14, pp. 1603–1611.
- Huttenlocher, Daniel P., Gregory A. Klanderman, and William J. Rucklidge (1993). "Comparing images using the Hausdorff distance". In: *IEEE Transactions on pattern analysis and machine intelligence* 15.9, pp. 850–863.
- Imai, Kosuke, Gary King, and Olivia Lau (2008). "Toward a common framework for statistical analysis and development". In: *Journal of Computational Graphics and Statistics* 17.4, pp. 892–913.

- Imbens, Guido W. (1992). "An efficient method of moments estimator for discrete choice models with choice-based sampling". In: *Econometrica* 60.5, pp. 1187–1214.
- Imbens, Guido W. and Tony Lancaster (1996). "Efficient estimation and stratified sampling". In: *Journal of Econometrics* 74.2, pp. 289–318.
- Jääskeläinen, Mikko and Markku Maula (2014). "Do networks of financial intermediaries help reduce local bias? Evidence from cross-border venture capital exits". In: *Journal of Business Venturing* 29.5, pp. 704–721.
- Jääskeläinen, Mikko, Markku Maula, and Tuukka Seppä (2006). "Allocation of attention to portfolio companies and the performance of venture capital firms". In: *Entrepreneurship Theory and Practice* 30.2, pp. 185–206.
- Jaccard, Paul (1901). *Etude comparative de la distribution florale dans une portion des Alpes et du Jura*. Impr. Corbaz.
- Jain, Bharat A., Narayanan Jayaraman, and Omesh Kini (2008). "The path-to-profitability of Internet IPO firms". In: *Journal of Business Venturing* 23.2, pp. 165–194.
- Jenkins, Richard (2000). "Categorization: Identity, social process and epistemology". In: *Current Sociology* 48.3, pp. 7–25.
- Jensen, Michael (2003). "The role of network resources in market entry: Commercial banks' entry into investment banking, 1991-1997". In: *Administrative Science Quarterly* 48.3, pp. 466–497.
- Jensen, Michael and Aradhana Roy (2008). "Staging exchange partner choices: When do status and reputation matter?" In: *Academy of Management Journal* 51.3, pp. 495–516.
- Johnston, Josée and Shyon Baumann (2007). "Democracy versus distinction: A study of omnivorousness in gourmet food writing". In: *American Journal of Sociology* 113.1, pp. 165–204.
- Kahle, David and Hadley Wickham (2013). "ggmap: Spatial visualization with ggplot2". In: *The R Journal* 5.1, pp. 144–161.
- Kaplan, Steven N. and Josh Lerner (2010). "It ain't broke: The past, present, and future of venture capital". In: *Journal of Applied Corporate Finance* 22.2, pp. 36–47.
- (2015). "Venture capital data: Opportunities and challenges". In: *Measuring entrepreneurial businesses: Current knowledge and challenges*. University of Chicago Press.
- Kaplan, Steven N. and Per Strömberg (2001). "Venture capitalists as principals: Contracting, screening, and monitoring". In: *American Economic Review* 91.2, pp. 426–430.
- Kembel, Steven W., Peter D. Cowan, Matthew R. Helmus, William K. Cornwell, Helene Morlon, David D. Ackerly, Simon P. Blomberg, and Campbell O. Webb (2010). "Picante: R tools for integrating phylogenies and ecology". In: *Bioinformatics* 26, pp. 1463–1464.

- Kennedy, Mark Thomas (2008). "Getting counted: Markets, media, and reality". In: *American Sociological Review* 73.2, pp. 270–295.
- Kiebzak, Stephen, Greg Rafert, and Catherine E. Tucker (2016). "The effect of patent litigation and patent assertion entities on entrepreneurial activity". In: *Research Policy* 45.1, pp. 218–231.
- King, Gary and Langche Zeng (2001). "Logistic regression in rare events data". In: *Political Analysis* 9, pp. 137–163.
- Kogut, Bruce and Harbir Singh (1988). "The effect of national culture on the choice of entry mode". In: *Journal of International Business Studies* 19.3, pp. 411–432.
- Kogut, Bruce, Pietro Urso, and Gordon Walker (2007). "Emergent properties of a new financial market: American venture capital syndication, 1960–2005". In: *Management Science* 53.7, pp. 1181–1198.
- Kolympiris, Christos, Nicholas Kalaitzandonakes, and Douglas Miller (2011). "Spatial collocation and venture capital in the US biotechnology industry". In: *Research Policy* 40.9, pp. 1188–1199.
- Kortum, Samuel and Josh Lerner (2000). "Assessing the contribution of venture capital to innovation". In: *RAND journal of Economics*, pp. 674–692.
- Kostova, Tatiana and Srilata Zaheer (1999). "Organizational legitimacy under conditions of complexity: The case of the multinational enterprise". In: *Academy of Management Review* 24.1, pp. 64–81.
- Kovács, Balázs and Michael T. Hannan (2015). "Conceptual spaces and the consequences of category spanning". In: *Sociological Science* 2, pp. 252–286.
- Kraus, Alan and Robert H. Litzenberger (1973). "A state–preference model of optimal financial leverage". In: *The Journal of Finance* 28.4, pp. 911–922.
- Lahr, Henry and Andrea Mina (2016). "Venture capital investments and the technological performance of portfolio firms". In: *Research Policy* 45.1, pp. 303–318.
- Lechner, Christian, Michael Dowling, and Isabell Welppe (2006). "Firm networks and firm development: The role of the relational mix". In: *Journal of Business Venturing* 21.4, pp. 514–540.
- Lee, Peggy M., Timothy G. Pollock, and Kyuho Jin (2011). "The contingent value of venture capitalist reputation". In: *Strategic Organization* 9.1, pp. 33–69.
- Lee, Peggy M. and Sunil Wahal (2004). "Grandstanding, certification and the underpricing of venture capital backed IPOs". In: *Journal of Financial Economics* 73.2, pp. 375–407.
- Leifeld, Philip (2013). "texreg: Conversion of statistical model output in R to HTML tables". In: *Journal of Statistical Software* 55.8, pp. 1–24.

- Lerner, Josh (1994). "The syndication of venture capital investments". In: *Financial Management* 23.3, pp. 16–27.
- (1995). "Venture capitalists and the oversight of private firms". In: *The Journal of Finance* 50.1, pp. 301–318.
- Lincoln, James R., Mitsuyo Hanada, and Jon Olson (1981). "Cultural orientations and individual reactions to organizations: A study of employees of Japanese-owned firms". In: *Administrative Science Quarterly* 26.1, p. 93.
- LiPuma, Joseph A. and Sarah Park (2014). "Venture capitalists' risk mitigation of portfolio company internationalization". In: *Entrepreneurship Theory and Practice* 38.5, pp. 1183–1205.
- Liu, Yu and Markku Maula (2016). "Local partnering in foreign ventures: Uncertainty, experiential learning, and syndication in cross-border venture capital investments". In: *Academy of Management Journal* 59.4, pp. 1407–1429.
- Loken, Barbara, Lawrence W. Barsalou, and Christopher Joiner (2008). "Categorization theory and research in consumer psychology". In: *Handbook of consumer psychology*. Ed. by Curtis P. Haugtvedt, Paul M. Herr, and Frank R. Kardes. Marketing and consumer psychology series. New York: Psychology Press, pp. 133–165.
- Lounsbury, Michael and Mary Ann Glynn (2001). "Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources". In: *Strategic Management Journal* 22.6-7, pp. 545–564.
- Lungeanu, Razvan and Edward J. Zajac (2016). "Venture capital ownership as a contingent resource: How owner-firm fit influences IPO outcomes". In: *Academy of Management Journal* 59.3, pp. 930–955.
- Lynn, Freda B., Joel M. Podolny, and Lin Tao (2009). "A sociological (de)construction of the relationship between status and quality". In: *American Journal of Sociology* 115.3, pp. 755–804.
- Ma, Dali, Mooweon Rhee, and Daegyung Yang (2013). "Power source mismatch and the effectiveness of interorganizational relations: The case of venture capital syndication". In: *Academy of Management Journal* 56.3, pp. 711–734.
- Magretta, Joan (2002). "Why business models matter". In: *Harvard Business Review* 80.5, pp. 86–92.
- Manev, Ivan M. and William B. Stevenson (2001). "Nationality, cultural distance, and expatriate status: Effects on the managerial network in a multinational enterprise". In: *Journal of International Business Studies* 32.2, pp. 285–303.
- Manigart, Sophie, Andy Lockett, Miguel Meuleman, Mike Wright, Hans Landström, Hans Bruining, Philippe Desbrières, and Ulrich Hommel (2006). "Venture capitalists' decision to syndicate". In: *Entrepreneurship Theory and Practice* 30.2, pp. 131–153.

- Mann, Ronald J. and Thomas W. Sager (2007). "Patents, venture capital, and software start-ups". In: *Research Policy* 36.2, pp. 193–208.
- Martens, Martin L., Jennifer E. Jennings, and P. Devereaux Jennings (2007). "Do the stories they tell get them the money they need? The role of entrepreneurial narratives in resource acquisition". In: *Academy of Management Journal* 50.5, pp. 1107–1132.
- Matusik, Sharon F. and Markus A. Fitza (2012). "Diversification in the venture capital industry: Leveraging knowledge under uncertainty". In: *Strategic Management Journal* 33.4, pp. 407–426.
- Maurer, Indre and Mark Ebers (2006). "Dynamics of social capital and their performance implications: Lessons from biotechnology start-ups". In: *Administrative Science Quarterly* 51.2, pp. 262–292.
- Merton, Robert K. (1968). "The Matthew effect in science". In: *Science* 159.3810, pp. 56–63.
- Meuleman, Miguel, Andy Lockett, Sophie Manigart, and Mike Wright (2010). "Partner selection decisions in interfirm collaborations: The paradox of relational embeddedness". In: *Journal of Management Studies* 47.6, pp. 995–1019.
- Meyer, John W. and Brian Rowan (1977). "Institutionalized organizations: Formal structure as myth and ceremony". In: *American Journal of Sociology*, pp. 340–363.
- Meyers-Levy, Joan and Alice M. Tybout (1989). "Schema congruity as a basis for product evaluation". In: *Journal of Consumer Research* 16.1, pp. 39–54.
- Milanov, Hana and Dean A. Shepherd (2013). "The importance of the first relationship: The ongoing influence of initial network on future status". In: *Strategic Management Journal* 34.6, pp. 727–750.
- Monika and A. K. Sharma (2015). "Venture capitalists' investment decision criteria for new ventures: A review". In: *Procedia - Social and Behavioral Sciences* 189, pp. 465–470.
- Morosini, Piero, Scott Shane, and Harbir Singh (1998). "National cultural distance and cross-border acquisition performance". In: *Journal of International Business Studies* 29.1, pp. 137–158.
- Murnieks, Charles Y., J. Michael Haynie, Robert E. Wiltbank, and Troy Harting (2011). "'I like how you think': Similarity as an interaction bias in the investor–entrepreneur dyad". In: *Journal of Management Studies* 48.7, pp. 1533–1561.
- Nagarajan, Anuradha and Will Mitchell (1998). "Evolutionary diffusion: Internal and external methods used to acquire encompassing, complementary, and incremental technological changes in the lithotripsy industry". In: *Strategic Management Journal*, pp. 1063–1077.

- Nahapiet, Janine and Sumantra Ghoshal (1998). "Social capital, intellectual capital, and the organizational advantage". In: *Academy of Management Review* 23.2, pp. 242–266.
- Nam, Dae-il, Haemin Dennis Park, and Jonathan D. Arthurs (2014). "Looking attractive until you sell: Earnings management, lockup expiration, and venture capitalists". In: *Journal of Management Studies* 51.8, pp. 1286–1310.
- Navis, Chad and Mary Ann Glynn (2011). "Legitimate distinctiveness and the entrepreneurial identity: Influence on investor judgments of new venture plausibility". In: *Academy of Management Review* 36.3, pp. 479–499.
- Newman, Daniel A. (2003). "Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques". In: *Organizational Research Methods* 6.3, pp. 328–362.
- Norton, Edgar and Bernard H. Tenenbaum (1993). "Specialization versus diversification as a venture capital investment strategy". In: *Journal of Business Venturing* 8.5, pp. 431–442.
- Oh, Hongseok, Myung-Ho Chung, and Giuseppe Labianca (2004). "Group social capital and group effectiveness: The role of informal socializing ties". In: *Academy of Management Journal* 47.6, pp. 860–875.
- Ollivier, Michèle (2008). "Modes of openness to cultural diversity: Humanist, populist, practical, and indifferent". In: *Poetics* 36.2-3, pp. 120–147.
- Opsahl, Tore (2009). *Structure and evolution of weighted networks*. University of London (Queen Mary College), London, UK.
- Ormans, Laurent (12. September 2016). *50 journals used in FT research rank*.
- Ozmel, Umit and Isin Guler (2015). "Small fish, big fish: The performance effects of the relative standing in partners' affiliate portfolios". In: *Strategic Management Journal* 36.13, pp. 2039–2057.
- Ozmel, Umit, Jeffrey J. Reuer, and Ranjay Gulati (2013). "Signals across multiple networks: How venture capital and alliance networks affect interorganizational collaboration". In: *Academy of Management Journal* 56.3, pp. 852–866.
- Pahnke, Emily Cox, Riitta Katila, and Kathleen M. Eisenhardt (2015). "Who takes you to the dance? How partners' institutional logics influence innovation in young firms". In: *Administrative Science Quarterly* 60.4, pp. 596–633.
- Pahnke, Emily Cox, Rory McDonald, Dan Wang, and Benjamin L. Hallen (2015). "Exposed: Venture capital, competitor ties, and entrepreneurial innovation". In: *Academy of Management Journal* 58.5, pp. 1334–1360.

- Park, Haemin Dennis and H. Kevin Steensma (2013). "The selection and nurturing effects of corporate investors on new venture innovativeness". In: *Strategic Entrepreneurship Journal* 7.4, pp. 311–330.
- Park, Haemin Dennis and Daniel Tzabbar (2016). "Venture capital, CEOs' sources of power, and innovation novelty at different life stages of a new venture". In: *Organization Science* 27.2, pp. 336–353.
- Perkmann, Markus and André Spicer (2010). "What are business models? Developing a theory of performative representations". In: *Research in the Sociology of Organizations* 29, pp. 269–279.
- Petersen, Mitchell A. (2009). "Estimating standard errors in finance panel data sets: Comparing approaches". In: *Review of Financial Studies* 22.1, pp. 435–480.
- Peterson, Richard A. (1992). "Understanding audience segmentation: From elite and mass to omnivore and univore". In: *Poetics* 21.4, pp. 243–258.
- Petkova, Antoaneta P., Violina P. Rindova, and Anil K. Gupta (2013). "No news is bad news: Sensegiving activities, media attention, and venture capital funding of new technology organizations". In: *Organization Science* 24.3, pp. 865–888.
- Petkova, Antoaneta P., Anu Wadhwa, Xin Yao, and Sanjay Jain (2014). "Reputation and decision making under ambiguity: a study of US venture capital firms' investments in the emerging clean energy sector". In: *Academy of Management Journal* 57.2, pp. 422–448.
- Phillips, Damon J., Catherine J. Turco, and Ezra W. Zuckerman (2013). "Betrayal as market barrier: Identity-based limits to diversification among high-status corporate law firms". In: *American Journal of Sociology* 118.4, pp. 1023–1054.
- Phillips, Damon J. and Ezra W. Zuckerman (2001). "Middle-status conformity: Theoretical restatement and empirical demonstration in two markets". In: *American Journal of Sociology* 107.2, pp. 379–429.
- Podolny, Joel M. (1993). "A status-based model of market competition". In: *American Journal of Sociology* 98.4, pp. 829–872.
- (1994). "Market uncertainty and the social character of economic exchange". In: *Administrative Science Quarterly* 39, pp. 458–483.
- (2001). "Networks as the pipes and prisms of the market". In: *American Journal of Sociology* 107.1, pp. 33–60.
- (2005). *Status signals: A sociological study of market competition*. Princeton, N.J. and Woodstock: Princeton University Press.
- Podolny, Joel M. and James N. Baron (1997). "Resources and relationships: Social networks and mobility in the workplace". In: *American Sociological Review* 62.5, pp. 673–693.

- Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan (1996). "Networks, knowledge, and niches: Competition in the worldwide semiconductor industry, 1984-1991". In: *American Journal of Sociology* 102.3, pp. 659–689.
- Pollock, Timothy G., Guoli Chen, Eric M. Jackson, and Donald C. Hambrick (2010). "How much prestige is enough? Assessing the value of multiple types of high-status affiliates for young firms". In: *Journal of Business Venturing* 25.1, pp. 6–23.
- Pollock, Timothy G., Peggy M. Lee, Kyuho Jin, and Kisha Lashley (2015). "(Un)tangled: Exploring the asymmetric coevolution of new venture capital firms' reputation and status". In: *Administrative Science Quarterly* 60.3, pp. 482–517.
- Ponds, Roderik, Frank van Oort, and Koen Frenken (2007). "The geographical and institutional proximity of research collaboration". In: *Papers in Regional Science* 86.3, pp. 423–443.
- Pontikes, Elizabeth G. (2012). "Two sides of the same coin: How category ambiguity affects multiple audience evaluations". In: *Administrative Science Quarterly* 57.1, pp. 81–118.
- Popov, Alexander and Peter Roosenboom (2012). "Venture capital and patented innovation: evidence from Europe". In: *Economic Policy* 27.71, pp. 447–482.
- Powell, Walter W., Kenneth W. Koput, and Laurel Smith-Doerr (1996). "Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology". In: *Administrative Science Quarterly* 41.1, pp. 116–145.
- Powell, Walter W., Douglas R. White, Kenneth W. Koput, and Jason Owen-Smith (2005). "Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences". In: *American Journal of Sociology* 110.4, pp. 1132–1205.
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria.
- Rao, M. V. Hayagreeva (1994). "The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895–1912". In: *Strategic Management Journal* 15.S1, pp. 29–44.
- Rider, Christopher I. (2009). "Constraints on the control benefits of brokerage: A study of placement agents in US venture capital fundraising". In: *Administrative Science Quarterly* 54.4, pp. 575–601.
- (2012). "How employees' prior affiliations constrain organizational network change: A study of U.S. venture capital and private equity". In: *Administrative Science Quarterly* 57.3, pp. 453–483.
- Robert J. Hijmans (2016). *geosphere: Spherical trigonometry*.

- Rosenbusch, Nina, Jan Brinckmann, and Verena Müller (2013). "Does acquiring venture capital pay off for the funded firms? A meta-analysis on the relationship between venture capital investment and funded firm financial performance". In: *Journal of Business Venturing* 28.3, pp. 335–353.
- Rosenkopf, Lori and Atul Nerkar (2001). "Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry". In: *Strategic Management Journal* 22.4, pp. 287–306.
- Rubin, Donald B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- (1996). "Multiple imputation after 18+ years". In: *Journal of the American Statistical Association* 91.434, pp. 473–489.
- Ruhnka, John C., Howard D. Feldman, and Thomas J. Dean (1992). "The "living dead" phenomenon in venture capital investments". In: *Journal of Business Venturing* 7.2, pp. 137–155.
- Ruhnka, John C. and John E. Young (1991). "Some hypotheses about risk in venture capital investing". In: *Journal of Business Venturing* 6.2, pp. 115–133.
- Sahlman, William A. (1990). "The structure and governance of venture-capital organizations". In: *Journal of Financial Economics* 27.2, pp. 473–521.
- Salomon, Robert and Zheyang Wu (2012). "Institutional distance and local isomorphism strategy". In: *Journal of International Business Studies* 43.4, pp. 343–367.
- Sandberg, William R. (1986). *The role of strategy, industry structure, and the entrepreneur: New venture performance*. MA: Lexington Books.
- Sapienza, Harry J. (1992). "When do venture capitalists add value?" In: *Journal of Business Venturing* 7.1, pp. 9–27.
- Sapienza, Harry J., Sophie Manigart, and Wim Vermeir (1996). "Venture capitalist governance and value added in four countries". In: *Journal of Business Venturing* 11.6, pp. 439–469.
- Schafer, Joseph L. and John W. Graham (2002). "Missing data: Our view of the state of the art". In: *Psychological methods* 7.2, pp. 147–177.
- Schwartz, Shalom H. (1994). "Beyond individualism-collectivism: New cultural dimensions of values". In: *Individualism and collectivism: Theory, method, and applications*. Ed. by Uichol Kim, Harry C. Triandis, Cigem Kagitcibasi, Sang-Chin Choi, and Gene Yoon. Newbury Park, CA: SAGE, pp. 85–119.
- Scott, W. Richard (2013). *Institutions and organizations: Ideas, interests, and identities*. SAGE Publications.
- Scott, W. Richard and John W. Meyer (1983). *Organizational environments: Ritual and rationality*. Beverly Hills: SAGE.

- Sgourev, Stoyan V. and Niek Althuisen (2014). ““Notable” or “not able”:
When are acts of inconsistency rewarded?” In: *American Sociological Review* 79.2, pp. 282–302.
- Shane, Scott and Daniel Cable (2002). “Network ties, reputation, and the
financing of new ventures”. In: *Management Science* 48.3, pp. 364–381.
- Sharpe, William F. (1964). “Capital asset prices: A theory of market equilibrium
under conditions of risk”. In: *The Journal of Finance* 19.3, pp. 425–
442.
- Shenkar, Oded (2001). “Cultural distance revisited: Towards a more rig-
orous conceptualization and measurement of cultural differences”. In:
Journal of International Business Studies 32.3, pp. 519–535.
- Shepard, Roger N. (1987). “Toward a universal law of generalization for
psychological science”. In: *Science* 237.4820, pp. 1317–1323.
- Shepherd, Dean A. (1999). “Venture capitalists’ assessment of new venture
survival”. In: *Management Science* 45.5, pp. 621–632.
- Shipilov, Andrew V. and Stan Xiao Li (2008). “Can you have your cake and
eat it too? Structural holes’ influence on status accumulation and mar-
ket performance in collaborative networks”. In: *Administrative Science
Quarterly* 53.1, pp. 73–108.
- Short, Jeremy (2009). “The art of writing a review article”. In: *Journal of Man-
agement* 35.6, pp. 1312–1317.
- Shrader, Rodney C. and Mark Simon (1997). “Corporate versus indepen-
dent new ventures: Resource, strategy, and performance differences”.
In: *Journal of Business Venturing* 12.1, pp. 47–66.
- Simon, Herbert A. (1955). “A behavioral model of rational choice”. In: *The
Quarterly Journal of Economics*, pp. 99–118.
- Singh, Jasjit (2005). “Collaborative networks as determinants of knowledge
diffusion patterns”. In: *Management Science* 51.5, pp. 756–770.
- Solt, Frederick and Yue Hu (2015). *interplot: Plot the effects of variables in in-
teraction terms*.
- Sorensen, Jesper B. and Toby E. Stuart (2000). “Aging, obsolescence, and
organizational innovation”. In: *Administrative Science Quarterly* 45.1,
pp. 81–112.
- Sorensen, Morten (2007). “How smart is smart money? A two-sided match-
ing model of Venture Capital”. In: *The Journal of Finance* 62.6, pp. 2725–
2762.
- Sorenson, Olav and Toby E. Stuart (2001). “Syndication networks and the
spatial distribution of venture capital investments”. In: *American Journal
of Sociology* 106.6, pp. 1546–1588.
- (2008). “Bringing the context back in: Settings and the search for syndi-
cate partners in venture capital investment networks”. In: *Administrative
Science Quarterly* 53.2, pp. 266–294.

- Stam, Wouter, Souren Arzlanian, and Tom Elfring (2014). "Social capital of entrepreneurs and small firm performance: A meta-analysis of contextual and methodological moderators". In: *Journal of Business Venturing* 29.1, pp. 152–173.
- Stinchcombe, Arthur L. (1965). "Social structure and organizations". In: *Handbook of organizations*. Ed. by James G. March. Chicago: Rand-McNally, pp. 142–193.
- Stuart, Toby E. (1998). "Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry". In: *Administrative Science Quarterly* 43.3, pp. 668–698.
- Stuart, Toby E., Ha Hoang, and Ralph C. Hybels (1999). "Interorganizational endorsements and the performance of entrepreneurial ventures". In: *Administrative Science Quarterly* 44.2, pp. 315–349.
- Stuart, Toby E. and Joel M. Podolny (1996). "Local search and the evolution of technological capabilities". In: *Strategic Management Journal* 17.S1, pp. 21–38.
- Suchman, M. C. (1995). "Managing legitimacy: Strategic and institutional approaches". In: *Academy of Management Review* 20.3, pp. 571–610.
- Sykes, Hollister B. (1990). "Corporate venture capital: Strategies for success". In: *Journal of Business Venturing* 5.1, pp. 37–47.
- Tenenbaum, Joshua B. and Thomas L. Griffiths (2001). "Generalization, similarity, and Bayesian inference". In: *Behavioral and Brain Sciences* 24.04, pp. 629–640.
- Ter Wal, Anne L. J., Oliver T. Alexy, Joern H. Block, and Philipp G. Sandner (2016). "The best of both worlds: The benefits of open-specialized and closed-diverse syndication networks for new ventures' success". In: *Administrative Science Quarterly* 61.3, pp. 393–432.
- Therneau, Terry M. (2015). *A package for survival analysis in S*.
- Therneau, Terry M. and Patricia M. Grambsch (2000). *Modeling survival data: Extending the Cox model*. New York: Springer.
- Trapido, D. (2007). "Competitive embeddedness and the emergence of interfirm cooperation". In: *Social Forces* 86.1, pp. 165–191.
- Tripsas, Mary (2009). "Technology, identity, and inertia through the lens of "the digital photography company"". In: *Organization Science* 20.2, pp. 441–460.
- Tushman, Michael L. (1977). "Special boundary roles in the innovation process". In: *Administrative Science Quarterly* 22.4, pp. 587–605.
- Tushman, Michael L. and Philip Anderson (1986). "Technological discontinuities and organizational environments". In: *Administrative Science Quarterly*, pp. 439–465.
- Ueda, Masako and Masayuki Hirukawa (2008). "Venture capital and industrial 'innovation'". In: *Available at SSRN 1242693*.

- van Tulder, Rob (2010). "The past, present and future of managing distance: Stakeholders and development". In: *The past, present and future of international business & management*. Ed. by Timothy M. Devinney, Torben Pedersen, and Laszlo Tihanyi. Vol. 23. Advances in international management. Bingley, UK: Emerald, pp. 137–165.
- Vergara, Marcos, Claudio A. Bonilla, and Jean P. Sepulveda (2016). "The complementarity effect: Effort and sharing in the entrepreneur and venture capital contract". In: *European Journal of Operational Research* 254.3, pp. 1017–1025.
- Walker, Strother H. and David B. Duncan (1967). "Estimation of the probability of an event as a function of several independent variables". In: *Biometrika* 54.1-2, pp. 167–179.
- Walske, Jennifer M. and Andrew L. Zacharakis (2009). "Genetically engineered: Why some venture capital firms are more successful than others". In: *Entrepreneurship Theory and Practice* 33.1, pp. 297–318.
- Washington, Marvin and Edward J. Zajac (2005). "Status evolution and competition: Theory and evidence". In: *Academy of Management Journal* 48.2, pp. 282–296.
- Wasserman, Noam (2003). "Founder-CEO succession and the paradox of entrepreneurial success". In: *Organization Science* 14.2, pp. 149–172.
- Wickham, Hadley (2007). "Reshaping data with the reshape package". In: *Journal of Statistical Software* 21.12.
- (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
- (2011). "The split-apply-combine strategy for data analysis". In: *Journal of Statistical Software* 40.1, pp. 1–29.
- (2016). *scales: Scale functions for visualization*.
- Wickham, Hadley and Romain Francois (2016). *dplyr: A grammar of data manipulation*.
- Wijbenga, Frits H., Theo JBM Postma, and Rebecca Stratling (2007). "The influence of the venture capitalist's governance activities on the entrepreneurial firm's control systems and performance". In: *Entrepreneurship Theory and Practice* 31.2, pp. 257–277.
- Wilson, Robert (1968). "The theory of syndicates". In: *Econometrica: journal of the Econometric Society*, pp. 119–132.
- Wooldridge, Jeffrey M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.
- Wright, Mike and Andy Lockett (2003). "The structure and management of alliances: Syndication in the venture capital industry". In: *Journal of Management Studies* 40.8, pp. 2073–2102.

- Wry, Tyler and Michael Lounsbury (2013). "Contextualizing the categorical imperative: Category linkages, technology focus, and resource acquisition in nanotechnology entrepreneurship". In: *Special Issue: Institutions, Entrepreneurs, Community* 28.1, pp. 117–133.
- Wry, Tyler, Michael Lounsbury, and P. Devereaux Jennings (2014). "Hybrid vigor: Securing venture capital by spanning categories in nanotechnology". In: *Academy of Management Journal* 57.5, pp. 1309–1333.
- Wuebker, Robert, Nina Hampl, and Rolf Wüstenhagen (2015). "The strength of strong ties in an emerging industry: Experimental Evidence of the Effects of Status Hierarchies and personal ties in venture capitalist decision making". In: *Strategic Entrepreneurship Journal* 9.2, pp. 167–187.
- Xu, Dean and Oded Shenkar (2002). "Institutional distance and the multinational enterprise". In: *Academy of Management Review* 27.4, pp. 608–618.
- Ypma, Tjalling J. (1995). "Historical development of the Newton–Raphson method". In: *SIAM Review* 37.4, pp. 531–551.
- Zacharakis, Andrew L. and G. Dale Meyer (1998). "A lack of insight: Do venture capitalists really understand their own decision process?" In: *Journal of Business Venturing* 13.1, pp. 57–76.
- Zeileis, Achim and Gabor Grothendieck (2005). "zoo: S3 infrastructure for regular and irregular time series". In: *Journal of Statistical Software* 14.6, pp. 1–27.
- Zeileis, Achim and Torsten Hothorn (2002). "Diagnostic checking in regression relationships". In: *R News* 2.3, pp. 7–10.
- Zhelyazkov, P. I. and Ranjay Gulati (2016). "After the break-up: The relational and reputational consequences of withdrawals from venture capital syndicates". In: *Academy of Management Journal* 59.1, pp. 277–301.
- Zott, Christoph and Quy Nguyen Huy (2007). "How entrepreneurs use symbolic management to acquire resources". In: *Administrative Science Quarterly* 52.1, pp. 70–105.
- Zucker, Lynne G., Michael R. Darby, and Jeff Armstrong (1994). *Intellectual capital and the firm: The technology of geographically localized knowledge spillovers*. Cambridge, MA: National Bureau of Economic Research.
- Zuckerman, Ezra W. (1999). "The categorical imperative: Securities analysts and the illegitimacy discount". In: *American Journal of Sociology* 104.5, pp. 1398–1438.
- (2015). "Optimal distinctiveness revisited: An integrative framework for understanding the balance between differentiation and conformity in individual and organizational identities". In: *Handbook of Organizational Identity*.
- Zuckerman, Ezra W., Tai-Young Kim, Kalinda Ukanwa, and James von Rittmann (2003). "Robust identities or non-entities? Typecasting in

the feature-film labor market". In: *American Journal of Sociology* 108.5, pp. 1018-1073.