Home-Bias in Online Fundraising:

An Analysis of International Reward-Based Crowdfunding

George Filatov

A thesis submitted in partial fulfilment of the requirements of

Sheffield Hallam University

for the degree of Doctor of Business Administration

October 2021

Candidate Declaration

I hereby declare that:

1. I have not been enrolled for another award of the University, or other academic or professional organisation, whilst undertaking my research degree.
2. None of the material contained in the thesis has been used in any other submission for an academic award.
3. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.
4. The work undertaken towards the thesis has been conducted in accordance with the SHU Principles of Integrity in Research and the SHU Research Ethics Policy.
5. The word count of the thesis is 71000.

Name	George Filatov
Date	October 2021
Award	Doctor of Business Administration (DBA)
Faculty	Sheffield Hallam University - Sheffield Business School
Director(s) of Studies	Samah Issa (SHU)
Supervisors	Barbara Scheck (MBS), Nigel Garrow (SHU)

Acknowledgements

I would like to express my sincere gratitude to my director of studies, Dr Samah Issa, for her continuous support and invaluable advice throughout all stages of this thesis. Her extensive constructive feedback, especially during the writing process, has not only been instrumental to improve the quality of this thesis but has also helped me to develop personally. Without her expertise, guidance, and encouragement, the completion of this thesis would not have been possible.

I also wish to express my profound gratitude to Prof. Dr Nigel Garrow for the numerous discussions, insightful comments, and invaluable suggestions. His advice has helped me to see complex matters from a different perspective and express my approach and findings more clearly in this thesis.

A debt of gratitude is also owed to the numerous lecturers from Sheffield Hallam University and Munich Business School that provided me valuable insights into scientific research, philosophy, and methodology. I would particularly like to thank Prof. Dr Barbara Scheck and Dr Ofer Dekel for helping me to develop the foundation for this research project.

I would also like to extend my deepest gratitude to my wife, Stephanie Vicentini, for giving me the daily encouragement, strength, and confidence in my abilities. Without her unremitting support and motivation, I would not have been able to persevere and complete this work.

Finally, I would like to thank my mother, Elena Filatov-Roth, and Dr Hermann Roth, without you none of this would have been possible.

Abstract

Home Bias is the recognized tendency of individuals to choose geographically proximate interaction partners. In business finance, Home Bias is to the detriment of both investors and entrepreneurs because it promotes an uneven distribution of capital and contributes to the Global Finance Gap. The aim of this thesis is to examine the existence of Home Bias in the emerging financing channel of reward-based crowdfunding. Crowdfunding, in general, is different from traditional financing because it shifts the entire fundraising process to a digital space on the internet. Moreover, it introduces new community-based trust mechanisms and eliminates some of the distance-related costs. The focus of this thesis lies on reward-based crowdfunding, which is currently the most popular, unrestricted and, therefore, most international form of crowdfunding.

To assess whether international reward-based crowdfunding is prone to Home Bias, this thesis employs a Negative Binomial regression model that examines the relationship between the count of crowdfunding project backers and their respective distance to entrepreneurs. The model builds on an aggregate data sample of 1,118,654 project-specific country-to-country investment observations (from 211,695 projects) that occurred on Kickstarter platform between 2009 and 2020, making it the largest and most up to date crowdfunding study.

Although large sample or "Big Data" models provide many advantages (e.g., higher representativeness), and have been commonly used in the crowdfunding literature, they however introduce some caveats that have been mostly ignored by previous research. One main issue that might distort results in Big Data models is that they are capable to identify marginally small patterns in the data that, although statistically significant in terms of p-values, might have little relevance in practice. Therefore, this thesis goes beyond the traditional analysis of statistical significance and devotes great attention to the assessment of different marginal effect sizes to identify the practical relevance of findings.

The thesis also investigates the effect of additional variables that may have potential effect on the count of backers namely GDP per capita of backers and entrepreneurs, project category, third-party endorsements, herding behaviour and Covid-19 pandemic.

The results suggest that although geographical distance appears to have a statistically significant negative influence on the count of backers, its practical effect is very small. This indicates that Home Bias has a comparably small relevance in international reward-based crowdfunding and that entrepreneurs should not overestimate its impact when planning their crowdfunding campaigns. Moreover, neither individual wealth of backers nor entrepreneurs, project category or global economic crises seem to affect the success of crowdfunding campaigns in a practically relevant manner. However, herding behaviour and third-party endorsements do seem to have a statistically and practically relevant influence on the count of backers and, therefore, should be considered in the planning of crowdfunding campaigns.

The overall findings of this thesis suggest that some of the prior research in crowdfunding might have overestimated the practical relevance of certain influencing factors (e.g., geographical distance and individual wealth), perhaps by focusing too much on statistical significance while ignoring the capability of Big Data models to identify marginally small and practically irrelevant patterns in the data.

Content

List of Tables and Figures

Tables

Figures

Abbreviations and Acronyms

CPS Chart	Coefficient P-Value Sample-Size Chart
EDA	Exploratory Data Analysis
GDP	Gross Domestic Product
GLM	Generalized Linear Model
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IMR	Internet-Mediated Research
IPO	Initial Public Offering
IQ	Intelligence Quotient
IRR	Incidence Rate Ratio
MEM	Marginal Effect at the Mean
MLE	Maximum Likelihood Estimation
NHST	Null Hypothesis Significance Testing
OLS	Ordinary Least Squares
PWL badge	"Projects We Love" badge awarded by Kickstarter
SMD	Standardized Mean Difference
SMEs	Small and Medium Sized Enterprises
SNT	Social Network Theory
VC	Venture Capital

Chapter One: Introduction

1.1 Access to Capital and the Global Finance Gap

Small- and medium-sized enterprises (SMEs) are considered the "backbone" of economies worldwide (Sibanda et al., 2018). They represent over 90% of all firms and account on average for more than 50% of total employment and GDP (International Finance Corporation, 2017). SMEs are important drivers for innovation, job creation, economic development and are positively linked to national prosperity (Ayyagari et al., 2016). Given their importance, most governments have strong interest in promoting the development of SMEs. For example, in 2020 the European Commission announced a new strategy specifically designed to support SMEs (European Comission, 2020). Some of the key actions include reducing regulatory burdens, improving market access as well as expanding the financing environment for new businesses.

However, research shows that, especially in the start-up phase, entrepreneurs confront multiple obstacles that jeopardize the growth and survival of their ventures (Anheier and Seibel, 1987; Steel and Webster, 1991; Aryeetey *et al.*, 1994; Gockel and Akoena, 2002; Abor and Quartey 2010). Among all obstacles, access to capital is one of the major hurdles for emerging entrepreneurs (Alibhai et al., 2017; Beck & Demirguc-Kunt, 2006; Carpenter & Petersen, 2002; World Bank, 2016a). Capital is required for daily operations, growth, and the development of new technologies. Studies have repeatedly found that the availability of capital at an early stage is crucial for the survival of young businesses (Mollick, 2014; Wiklund & Shepherd, 2005).

The central problem that dominates current public discourse is that start-ups are less likely to obtain bank loans or venture capital funding than large firms (Gompers & Lerner, 2004; Gorman & Sahlman, 1989; Hwang et al., 2019; Kortum & Lerner, 2000; Mollick & Robb, 2016). According to the *US Small Business Administration*, the average age of a firm receiving venture capital funding in the US is four years (Guenther et al. 2018). Accordingly, the probability of an entrepreneur receiving venture capital funding is as low as 0.0005% (Rao, 2013). Access to capital is a global problem, and the current global credit gap is estimated to range between 3 to 5 trillion USD (Ferrando et al., 2019; International Finance Corporation, 2017; World Bank, 2016a). The World Bank (2016) reports that more than 65 million enterprises (40% of all studied companies in 128 reviewed countries) are constrained by insufficient funding. The problem of fundraising appears to be most severe in emerging economies, where financial markets are often less developed (World Bank, 2016). Moreover, emerging economies often face additional challenges such as high inflation rates, unstable currency exchange rates, corruption, or political instability (International Finance Corporation, 2017). These factors demonstrate additional hurdles for investors to deal with emerging economies. According to the World Bank (2016), the East Asia and Pacific area accounts for the largest share of the total global finance gap (46%), followed by Latin America and the Caribbean (23%).

Due to inability to receive finance from professional institutional investors, many start-ups are dependent on internal funds or loans from family and friends (International Finance Corporation, 2017). These funds, however, are often insufficient to ensure the growth and survival of young firms (Hwang et al., 2019). Olawale and Garwe (2010) show that 75% of new companies in South Africa do not go beyond the start-up stage. Similarly, Von Broembsen et al. (2005) highlight that the probability of a new company surviving beyond 42 months is less likely in Africa than in any other place in the world. Enterprises that manage to survive the difficult start-up period are often smaller and grow considerably slower than comparable firms in developed economies (Hsieh and Klenow 2014).

Improving the access to capital for SMEs is a viable strategy to advance economic development and reduce poverty. Bruhn and Love (2009) show that new bank openings to unserved low-income groups in Mexico lead to a significant increase in businesses, which in turn lead to greater prosperity of the region. Accordingly, Dinh et al. (2010) find that the access to additional investment funds leads to a significantly higher employment growth. This is in line with Klapper et. al (2007), who show that countries with more developed financial markets have on average more entrepreneurs (Klapper et al., 2007).

Therefore, the development of new concepts to provide finance to young businesses has been one of the major priorities of the World Bank Group and other development institutions around the globe (Dinh et al., 2010; Hwang et al., 2019; World Bank, 2012, 2013). Different solutions have been discussed such as state-owned venture capital or micro-loans (Hwang et al., 2019; World Bank, 2016). However, the suggested solutions are often cumbersome, cost-intensive and difficult to implement (World Bank, 2016). Few of these efforts have led to systematic change in the funding problem of young firms (Hwang et al., 2019). This experience has promoted the need for better solutions for financing problems.

1.2 Home Bias in Investment Decisions

One central problem that contributes to the global finance gap is that financial resources are unevenly distributed in terms of geography (Hwang et al., 2019; Sorenson et al., 2016). Especially venture capital is often highly concentrated and largely unavailable outside of established business hubs (Mollick, 2014; Wiklund & Shepherd, 2005). The problem is particularly well illustrated in the global report of *Martin Prosperity Institute* on the geography of venture capital investments (Florida & King, 2016). The report states the top 10 business hubs account for more than 52% of the global venture capital funding.[1] A particular important role can be attributed to the United States (US). According to the report, the US alone accounts for approximately 70% of the total global venture capital. Moreover, the geographic concentration of venture capital persists also on the national level. Close to 80% of the US's venture capital funding is distributed among only five regional clusters: San Francisco (North Bay Area), Silicon Valley (South

[1] The top-10 business hubs are San Francisco, San Jose, Boston, New York, Los Angeles, San Diego, London, Washington D.C., Beijing and Seattle.

Bay Area), New England, New York Metro and Los Angeles/Orange Country (PwC & CB Insights, 2019). Overall, the San Francisco Bay Area (North and South) is responsible for more than a quarter ($11bn.) of all global venture capital investments (Florida & King, 2016).

Although remote investors are common in public equity markets, they are rather an exception for start-ups. According to Sorenson and Stuart (2005) the average distance between a venture capital investor and the target firm amounts to less than 70 miles. Venture capital investors rarely invest in distant companies and only deviate from this behaviour when they can invest together with a trusted third partner that resides near the target firm (Sorenson and Stuart, 2001). Therefore, entrepreneurs that find themselves remote from established business clusters often have few chances to receive venture capital funding for their ideas (Kim and Hann, 2013). Due to the uneven distribution of capital, research suggests that business success is often dependent on the geographical location of the firm (Ferrando et al., 2019; Hwang et al., 2019; Porter, 2000; World Bank, 2016a).

The described tendency to interact with close parties, either in the same country or same city, is a common phenomenon in investment decision making and referred to as the "Home Bias" (K. Kim & Hann, 2013; M. Lin & Viswanathan, 2016; Niemand et al., 2018). Home Bias is one of many cognitive biases that exist in investment decisions.[2] However, research on Home Bias enjoys special attention because of its ubiquitous presence. The concept has been documented in research on entrepreneurial finance, international trade as well as the purchasing behaviour of individuals (Ahearne et al., 2004; I. Cooper & Kaplanis, 1994; Coval & Moskowitz, 1999; Dziuda & Mondria, 2012; Graham et al., 2009; Karlsson & Nordén, 2007). In business financing, Home Bias is to the detriment of both entrepreneurs and investors (Chen et al., 2009; Mollick, 2013; T.

[2] See also "confirmation bias" (Nickerson, 1998), "anchoring bias" (Tversky and Kahneman, 1992) and "information bias" (Baron, Beattie, and Hershey, 1988).

Stuart & Sorenson, 2003). Entrepreneurs are unable to raise funds for their ventures if they are remote from established business hubs, while investors expose their portfolios to higher risk and miss potential business opportunities by focusing only on geographically proximate firms (Grauer & Hakansson, 1987).

1.3 The Emergence of Crowdfunding

Among all changes that occurred within the last century, the rise of the internet has been the most influential. It has changed how people work, socialize, create and share information (Niemand et al., 2018). Therefore, much of the economic growth in the last two decades can be attributed to the rise of the internet (Manyika & Roxburgh, 2011; World Bank, 2016b). According to the World Bank (2016b), the internet affects economic development by (1) introducing new ways of overcoming information problems; (2) making transactions faster, cheaper, and more convenient; and (3) spawning innovative business models that allow a more efficient connection of supply and demand, while using various types of scale economies. Overall, the internet contributes to making the world increasingly "flat" and more integrated (Friedman, 2005). Its significant influence is frequently compared to that of the agricultural- or industrial revolution (Bojanova, 2014; Sidhu & Doyle, 2016).

The financial sector has experienced numerous internet-driven innovations (Niemand et al., 2018). It is likely that the emergence of new financial instruments was also stimulated by the financial crisis as a response to the increasing mistrust into the traditional banking system. One particularly important change was introduced through the emergence of "crowdfunding" in the beginning of the 21st century that fundamentally changed how savers and borrowers, and investors and investees could interact (Surowiecki 2004).

The term crowdfunding describes the attempt of collecting financial resources from a large and unaffiliated crowd of investors via specialized digital platforms (H. Kim & Kim, 2017). Lambert & Schwienbacher (2010) were among the first to provide an official definition for the concept of crowdfunding in the scientific literature. They describe it as "*an open call, essentially through the Internet, for the provision of financial resources either in form of donation or in exchange for some form of reward and/or voting rights in order to support initiatives for specific purposes.*" According to Mollick (2014), *"crowdfunding refers to the efforts by entrepreneurial individuals and groups – cultural, social, and for-profit – to fund their ventures by drawing on relatively small contributions from a relatively large number of individuals using the internet, without standard financial intermediaries."*

Crowdfunding is fundamentally different from traditional fundraising concepts because the entire process of capital procurement takes place in a virtual environment on the internet. Business founders can use specialized crowdfunding platforms to describe their ideas and present it to a large and diversified crowd of potential investors (Mollick, 2014). Investors who like a specific idea can support the founder(s) with financial resources via the provided infrastructure of the platform. In return, the "backers" receive a reward which can be in the form of an early version of the new product itself, interest on the investment, or equity of the firm. [3] In general, crowdfunding projects can range from small creative intentions to social and entrepreneurial ventures seeking millions of dollars in capital.

The fundamental idea of crowdfunding, raising funds from the general public, is not new. Back in 1885, Joseph Pulitzer used newspaper-led public fundraising to finance the construction of the monumental base for the Statue of Liberty (Frydrych et al., 2014). Similarly, non-profit organizations have been using public fundraising campaigns via newspapers, direct mailings, or television advertising successfully for many years (cp. campaigns of Greenpeace or Unicef). However, the central advantage of the modern internet-enabled crowdfunding approach is its speed and scope (Lansiti & Lakhani,

[3] The term "backer" is commonly used within the crowdfunding literature to refer to individuals that contribute financial resources to the realization of a certain crowdfunding project (see Kickstarter (2019c)).

2017). Crowdfunding websites such as Kickstarter.com or Indiegogo.com are typical examples of the emerging platform economy and act as global matchmakers for both entrepreneurs and funders (Mollick, 2013). They connect capital demand and supply more efficiently, reduce communication costs and introduce unprecedented possibilities for interaction and collaboration between investors and business providers (Blum & Goldfarb, 2006). These characteristics make crowdfunding a significant innovation that has the potential to fundamentally change capital markets and diminish the need for traditional intermediaries such as banks or venture capital firms (Niemand et al., 2018). By incorporating a large community of small and mostly private investors, crowdfunding platforms spread the risk among multiple stakeholders and significantly increase the scope of investor procurement. Scholars argue that crowdfunding platforms not only provide a more efficient but also more "democratic" access to capital that could greatly benefit new entrepreneurs and small business owners (K. Kim & Hann, 2013; Mollick & Robb, 2016).

Another important advantage of crowdfunding is that it builds on internet-enabled concepts to overcome information problems. In the past, even when search costs were low, transactions often did not take place when one party to a transaction had great informational advantage over the other.[4] This scenario is common and frequently referred to as the *Information Asymmetry Problem* (Spence, 1970). Crowdfunding, as well as other internet-based concepts, provide entirely new strategies to reduce information asymmetries between transaction partners (see section 2.3.2). This, in turn, leads to more trust and higher efficiency of markets. The beneficial characteristics of crowdfunding are expressed through the rapid growth of the industry. In 2014, the transaction volume of crowdfunding reached US$16,2 bn. (Massolution, 2015). The industry report of Massolution (2015) has predicted the transaction volume to double

[4] For example, a car salesman has typically more information about the condition of the car than the potential buyer.

and to reach US$34 bn. in 2020. The World Bank estimates that by 2025 crowdfunding will reach a yearly volume of US$90 - 96 billion (Niemand et al., 2018).

The crowdfunding platform Kickstarter provides numerous successful examples in which entrepreneurs were able to raise funds from a culturally diverse and geographically highly distributed crowd. In some cases, projects received support from more than 170 different countries (Kickstarter, 2016b). According to Kickstarter, the average distance between backers and founders is more than 2300 miles. This number is considerably higher than the reported distance in traditional venture capital financing of 70 to 300 miles (Sorenson and Stuart 2001, Cumming and Dai 2010).

A frequently used example in the literature for the potential of crowdfunding is the fundraising campaign of “Pebble”, an American smartwatch producer (Guenther 2018). After being denied by multiple local venture capital firms, Pebble used crowdfunding on the Kickstarter platform to raise more than US$20 million in 2015. Interestingly and surprisingly, more than 60% of the funding Pebble received came from non-American backers (Kickstarter, 2016a). Specifically, backers from the UK, Australia and Singapore contributed significantly to the success of the fundraising campaign. Pebble’s example attracts special attention because it stands in contradiction to the Home Bias theory which suggests that investors and founders are always co-located (Mason, 2007; Powell et al., 2002; Sorenson & Stuart, 2001; Zook, 2002).[5] This example has raised new hopes among scientists and entrepreneurs that crowdfunding bears the potential to alleviate some of the existing distance related frictions and provide new opportunities for businesses to raise capital. However, Pebble’s example must also be interpreted with caution as it does not allow any form of generalization for the industry. Therefore, profound scientific research is required to study the assumption that crowdfunding truly alleviates the Home Bias problem in business financing.

[5] Another famous example is “Oculus RV”, a company that was able to raise US$2.5 million via crowdfunding and was later acquired by Facebook for US$2 billion (Guenther et al. 2018).

1.4 Basic Models of Crowdfunding

Nowadays, more than 800 active crowdfunding platforms exist across the world (Frydrych et al., 2014).[6] The fundamental principle of these platforms is comparable, which is collecting funds from a large and unaffiliated crowd. However, a distinction is made based on the backers' primary motivation to participate and the corresponding reward for the investment (K. Kim & Hann, 2013). The scientific literature distinguishes between four basic models of crowdfunding: (1) *donation-based,* (2) *reward-based,* (3) *lending-based, and* (4) *equity-based models* (K. Kim & Hann, 2013; Lambert & Schwienbacher, 2010; Mollick, 2014, 2014).

1.4.1 The Donation-Based Model

The underlying concept of the donation-based crowdfunding model is known in charity campaigns of non-profit organizations (Langley & Leyshon, 2017). The aim is generally to help people that are facing a particularly difficult situation in life. Donors that engage into the donation-based model are usually not looking for a financial return or gain but are mostly philanthropists that find their reward in the satisfaction of knowing that their funds are going to be used for a good cause (Gerber & Hui, 2013). Accordingly, the receiver of funds does not need to repay the collected amount in any form.

The main advantages of donation-based crowdfunding, compared to traditional charity fundraising, are the easiness and the speed of campaign creation (Hui et al., 2012). In the past, it required great effort to find potential donors. Social welfare organizations had to use costly mailing campaigns, tv-advertisements, or direct face-to-face approaches in city-centres to attract attention. Crowdfunding platforms facilitate this process by providing a global meeting place for socially engaged people. These platforms

---

[6] Crunchbase.com, a platform for data on private and public firms, listed more than 2,000 companies related to the term "crowdfunding" in 2021.

offer space to describe the social project and provide the necessary infrastructure to guarantee secure transactions. Donation-based crowdfunding has significantly widened the scope of charity fundraising, allowing anyone to start a social campaign (Mollick & Robb, 2016). Examples of donation-focused crowdfunding platforms are *gofundme.com* and *crowdrise.com*.

1.4.2 The Reward-Based Model

Reward-based crowdfunding is currently the most famous and most widespread crowdfunding model (Frydrych et al., 2014; Kraus et al., 2016; Mollick, 2014). In contrast to the previously described donation-based type, reward-based crowdfunding platforms focus on entrepreneurship and the realization of concrete business ideas. Founders of reward-based crowdfunding campaigns usually offer different kinds of rewards to backers in relation to the amount funded (Langley & Leyshon, 2017). The rewards can range from a presale version of the product to a small symbolic reward such as a custom-made t-shirt or the mentioning of the backer's name. The reward is usually handed over after the successful realization of the idea. Backers are mostly private individuals that are interested in the product itself. Some of them also participate in projects because they like the idea and want to support entrepreneurship (Gerber & Hui, 2013). The most famous examples for the reward-based crowdfunding model are *Kickstarter.com* and *Indiegogo.com*.

1.4.3 The Lending-Based Model

In the lending-based model, also referred to as peer-to-peer lending, crowdfunding platforms resemble the activities of banks. They provide a virtual meeting place for lenders and borrowers. Private individuals and businesses can approach lending-based crowdfunding platforms for unsecured loans (M. Lin, Prabhala, & Viswanathan, 2013). Depending on the requested amount, the loan comes from one single or multiple private individuals. In return, the lenders receive a predefined interest on the loan, which is often higher than the market average due to the increased risk. Borrowers engage into this model of crowdfunding because the process to receive a loan is often faster and less bureaucratic than that of banks (Hui et al., 2012). Lending-based crowdfunding platforms also increase the probability to receive a loan by providing the access to a large community of potential lenders with different risk propensities (Langley & Leyshon, 2017). Backers participate in this model mostly in expectation of some sort of financial return. Examples for lending-based crowdfunding platforms are *prosper.com* and a*uxmoney.com*.

1.4.4 The Equity-Based Model

Entrepreneurs can use the equity-based model to promote their businesses. The major difference to other models lies in the type of reward. As the name indicates, investors receive equity shares of the company in exchange for their financial contributions. Due to its characteristics, equity crowdfunding can be considered as the most sophisticated version of crowdfunding (Hornuf & Schmitt, 2016).

The equity-based model strongly resembles traditional venture capital financing. The major difference is that the entire fundraising process takes place on the internet. Investors and entrepreneurs usually do not meet face-to-face. Moreover, investors usually do not have the possibility to conduct a due diligence prior to the investment. They are limited to the information that is publicly available on the specific project website (cp. www.seedrs.com). This information usually includes a business plan with key figures and a rough risk estimation. In comparison to other crowdfunding models, equity-based crowdfunding demonstrates the highest demands for disclosure. Nevertheless, these disclosure requirements are considerably lower than an investor would demand in a traditional venture capital fundraising context (cp. Vinturella & Erickson, 2013). Due to the one-to-many relationships and the transmission of company stakes, equity crowdfunding resembles, to a certain extent, initial public offerings (IPOs). In fact, the similarities make it difficult for countries to implement equity crowdfunding into their legislation (Hornuf & Schmitt, 2016). Some national financial supervisors have raised great concern that equity crowdfunding is not required to fulfil the official documentation requirements of IPOs. Therefore, some countries do not allow equity crowdfunding platforms to operate on their markets (Rose, 2019; Vasconcelos, 2013). Examples of equity-based crowdfunding platforms are *seedrs.com*, *crowdcube.com* and *crowdfunders.com*.

1.5 Research Gaps

Crowdfunding is a comparably new field of research and its underlying mechanisms have not yet been fully understood. Lin et al. (2014) highlight that most research is in working paper formats and a comparably small number of scientific studies have been published in journals (Lin et al., 2014). Despite the increasing interest in this industry, only a few authors have addressed the effect of Home Bias on international backing decisions to date (cp. Burtch et al., 2014; Agrawal et al., 2015; Guo et al., 2018; Niemand et al., 2018).

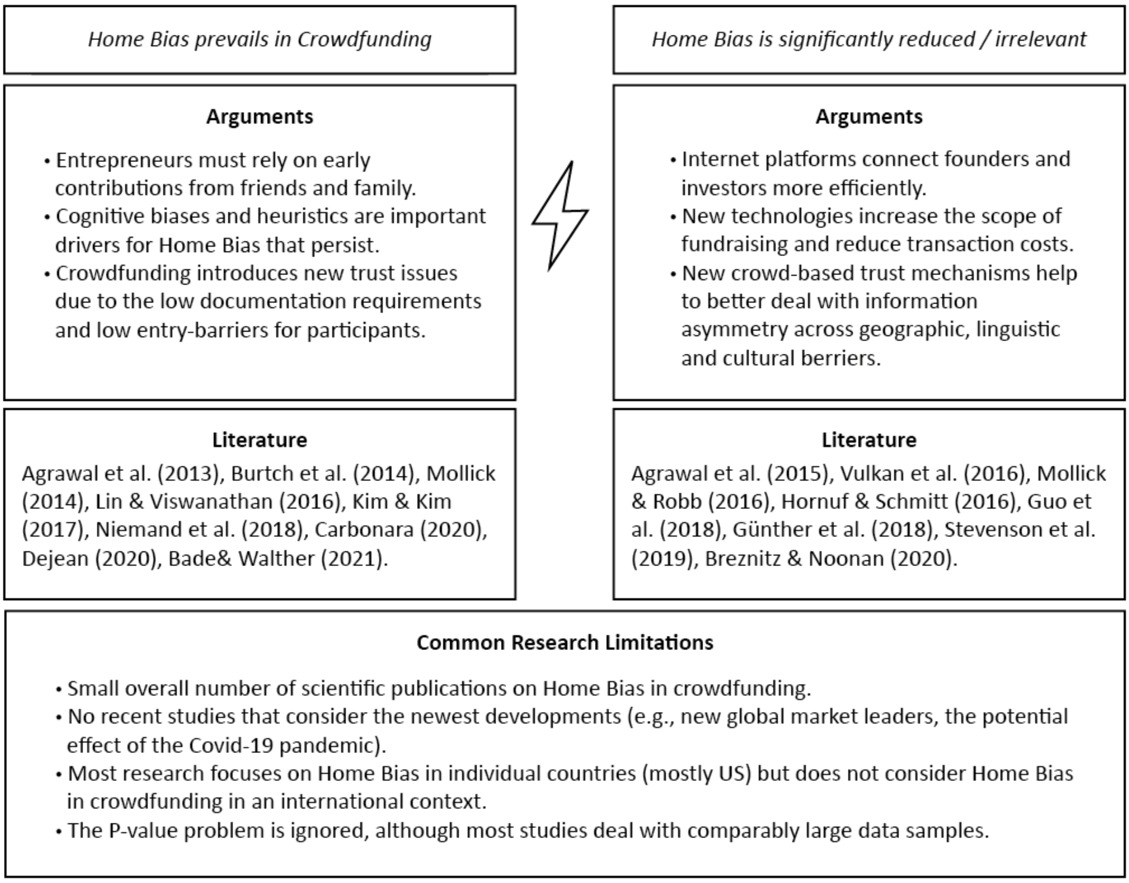
Within the scope of the available literature, a considerable disagreement exists on whether backers are negatively influenced by geographic distance to the entrepreneur when making decisions to support a specific project or not. On the one hand, some scholars argue that the digital nature of crowdfunding increases the scope of social networks and connects founders and investors more efficiently (Dekel et al., 2016; K. Kim & Hann, 2013; Mollick & Robb, 2016). Moreover, it reduces transaction costs and introduces new approaches to deal with the information asymmetry problem, which should render geographical distance irrelevant (Thierer et al., 2015).

On the other hand, some scholars provide evidence that Home Bias continues to matter in crowdfunding and that most transactions occur between parties that are geographically proximate to each other (Mollick, 2014; Kim & Kim, 2017; Niemand et al., 2018). These scholars argue that the new crowd-based trust mechanisms only work effectively at an advanced stage of the fundraising process. Entrepreneurs still face the challenge of attracting early-stage investors and must often rely on contributions from (mostly geographically proximate) family and friends to create a certain momentum of the campaign (Guo et al., 2018). Moreover, some authors suggest that crowdfunding introduces new trust issues due to the low entry-barriers for participants and low documentation requirements (Agrawal et al., 2013). Another argument as to why Home Bias might prevail in crowdfunding is that individuals do not always make rational or logical decisions (M. Lin & Viswanathan, 2016). Kahneman and Tversky (1981) showed that in a situation of uncertainty, people frequently use the strategy of attribute substitution. For example, the geographical proximity of a business partner might evoke "over-optimism" about the transaction (Hortaçsu et al., 2009; Lai & Teo, 2008; Strong & Xu, 2003). Such *cognitive biases* are often considered an important driver for Home Bias in traditional funding that might persist also in online fundraising (M. Lin & Viswanathan, 2016; Hortaçsu et al., 2009; Lai & Teo, 2008; Strong & Xu, 2003).

Figure 1 illustrates the current disagreement on the existence of Home Bias in the crowdfunding literature.

Figure 1: Disagreement on the Existence of Home Bias in Crowdfunding

This Figure shows the main disagreement in the literature on the existence of Home Bias in crowdfunding, the central arguments, as well as the research limitations.



The existing literature also highlights several shortcomings of current research on Home Bias in crowdfunding. One important shortcoming is that the current discussion on Home Bias is predominantly based on the findings of few early studies. Most notably is the work of Agrawal et al. (2013; 2015), Mollick (2013; 2014), Lin & Viswanathan (2016) and Burtch et al. (2014), which are the most cited authors in the literature. However, their research scope only covers the early years (prior to 2013) when the industry was highly fragmented and mostly a niche market. One important characteristic of digital markets is that they are subject to fast and continuous change. This dynamic change poses a great challenge to research because findings can quickly lose their validity. Since the above mentioned papers were published, the crowdfunding industry has undergone considerable consolidation, which has led to the emergence of new global market leaders (Ruhnau, 2019). Crowdfunding platforms such as Kickstarter and Indiegogo have integrated different areas of the world and created unprecedented transnational markets for venture capital. However, in recent years comparably few scientific studies (i.e., Guo et al, 2018; Günther et al, 2018; Niemand et al, 2018) have attempted to question or reassess the early findings of the above-mentioned authors.

Another considerable weakness is that most research examines the effect of Home Bias in individual countries or regions but does not assess it in an international context (K. Kim & Hann, 2013; M. Lin & Viswanathan, 2016; see Mollick, 2013, 2014). Particularly striking is that most studies focus exclusively on the US market and deliberately exclude projects or backers from foreign countries. Guo et al. (2018) suggest that the findings of such studies might be valid for American investors but lack the validity for other countries. The strong focus on the US is comprehensible because the US demonstrates a pioneering role in crowdfunding. However, national focus is no longer justified or contemporary because the industry is becoming increasingly international (Massolution, 2015; Niemand et al., 2018) and a growing number of individuals worldwide are using crowdfunding for corporate financing.

Particularly surprising is the lack of international studies in reward-based crowdfunding (cp., Table 2, p. 65), which is one of the most used and most extensive crowdfunding models that exist to date (Frydrych et al., 2014; Kraus et al., 2016; Mollick, 2014). Other crowdfunding variations, such as equity- or lending-based crowdfunding, are usually affected by national legislations, which limits their overall geographical scope (cp. also section 1.4). For example, Klaes (2017) identifies several potential legislative barriers to cross-border transactions that prevent equity-based crowdfunding platforms to operate across Europe. These barriers provide a possible explanation why researchers tend to find evidence of Home Bias in equity- and lending-based crowdfunding platforms (cp. Table 2, p. 65). The donation-based crowdfunding platforms, in turn, often lack the business context (K. Kim & Hann, 2013). The created crowdfunding campaigns usually focus on charitable purposes in which backers provide financial resources but do not expect anything in return. These platforms are not suitable for raising funds for product ideas or business ventures. Reward-based crowdfunding is currently the only model that offers both a business context and the possibility to launch wide ranging international fundraising campaigns. Therefore, further research on Home Bias in reward-based crowdfunding could help scientists and practitioners to better estimate the potential of crowdfunding to reduce the influence of Home Bias in business financing.

Although crowdfunding is repeatedly highlighted as an effective alternative to traditional corporate financing via recognized financial institutions (K. Kim & Hann, 2013; Mollick & Robb, 2016), little is known about its long-term reliability. For example, research is missing on how crowdfunding markets perform in times of global economic crises. Crowdfunding was spared from the 2007/8 financial crisis because it only became popular at a later stage.[7] Similarly, to date, no research has addressed how the global Covid-19 pandemic might have influenced the fundraising success of crowdfunding campaigns. It remains unclear whether the functionality of crowdfunding is limited during economic crises or whether it provides an effective and stable alternative when traditional funding (i.e., via financial institutions) is difficult to access. This question is highly relevant for theorists and practitioners because it can significantly influence the development of companies and, thereby, the course of economic crises. The impaired access to capital during the financial crisis is the dominant explanation for the Great Recession (Brunnermeier, 2009; Gorton, 2010). Therefore, it is interesting to investigate whether crowdfunding could offer an alternative funding source for companies during crises.

Another shortcoming of current research on crowdfunding concerns the methodological approach of existing studies. The new technologies of the internet, and advancements

[7] One assumption is that crowdfunding activities increased significantly in the last decade as a response to the impact the financial crise has left on the mainstream banking system. It has led many individuals to search for alternatives and, in particular, less financial intermediation between savers and borrowers, and investors and investees.

in information technologies in general, have allowed new possibilities for researchers. One important result is that they allow considerably faster and greater data collection. Big Data research that was previously reserved for professional scientists with large funding can now be realised by young scientists with ordinary computers. This shift offers many new benefits. For example, research can be replicated more easily and in significantly larger scope. Moreover, Big Data models allow incorporating a multitude of variables at the same time without losing estimation power. However, the new possibilities also introduce potential problems that are often overlooked by contemporary research. For example, M. Lin, Lucas, and Shmueli (2013) question whether the traditional approach of statistical significance testing is still valid for research in the era of Big Data. One problem is that Big Data models have the statistical power to identify marginally small patterns. These patterns might be statistically significant but irrelevant in practice (Chatfield, 1995). The so-called *p-value problem* in large samples is still insufficiently recognized by many scientists (Lin et al., 2013). Crowdfunding research is one of those areas that is greatly affected by the p-value problem because it often builds on comparatively large samples (cp. sample size of Burtch et al., 2014; Guo et al., 2018). In the current literature on crowdfunding, however, no research addresses this potential issue. None of the studies consider that the estimated low p-values might be an artefact of the large-sample size. By ignoring the capability of Big Data models to find particularly small patterns and relationships in the data might lead scientists to finding statistically significant results that are of little or no practical value.

1.6 Research Aim, Questions and Objectives

The aim of this thesis is to examine the existence of the Home Bias phenomenon in international reward-based crowdfunding. Home Bias is the proven tendency of individuals to choose geographically proximate interaction partners (Niemand et al., 2018). In business finance, it is considered an important problem because it promotes inefficient capital distribution and contributes to the Global Finance Gap (Coval &

Moskowitz, 1999; Tesar & Werner, 1995). Reward-based crowdfunding, which is the most popular and widespread crowdfunding concept for business financing (Statista, 2018), is often discussed as a potential approach to alleviate the effect of Home Bias and improve the allocation of capital globally (K. Kim & Hann, 2013; Mollick & Robb, 2016). However, current research provides inconsistent findings on whether backers in crowdfunding are influenced in their investment decisions by the geographical distance to entrepreneurs (cp. Dekel et al., 2016; K. Kim & Hann, 2013; Mollick & Robb, 2016; Mollick, 2014; Kim & Kim, 2017; Niemand et al., 2018). In particular, there is a lack of international studies that consider cross-border investment flows. To address the stated aim and fill the identified research gaps in the crowdfunding literature, this thesis poses the following research question:

Does geographical distance between backers and entrepreneurs affect the overall count of project supporters in international reward-based crowdfunding?

To pursue the stated research question, the following five objectives are defined:

Research Objective 1: *To construct the largest and most recent crowdfunding dataset to date.* This dataset will include information on the count of backers and the respective location data of the interacting parties. One benefit of large datasets is that they are commonly associated with higher representativeness, meaning that the findings are more likely to be generalizable (Kaplan et al., 2014). Another advantage of large datasets is that they allow the analysis of different subsamples of interest, while maintaining sufficiently large sample sizes for quantitative analysis (M. Lin, Lucas, & Shmueli, 2013). However, this thesis equally considers and evaluates potential problems that are associated with large datasets (cp. Research Objective 4). For example, different authors highlight that large datasets (or Big Data models) offer the statistical power to identify marginally small patterns in the data that might be statistically significant but irrelevant in practice (Singh Chawla, 2017; Lin et al., 2013; Chatfield, 1995).

The data for this thesis is extracted from Kickstarter platform, which is the global market leader for reward-based crowdfunding. Kickstarter meets the requirements of this thesis particularly well because it provides data on crowdfunding projects from more than 200 countries, making it the most international crowdfunding platform of the industry (Kickstarter, 2016b). By collecting information on all projects that were launched on Kickstarter between April 2009 to June 2020, this thesis incorporates not only the longest period but also the most recent data on crowdfunding that can be found in the literature to date.

Research Objective 2: *To use the collected data from Kickstarter to calculate the distance between backers and entrepreneurs.* For this purpose, the geographical coordinates (latitude and longitude) of the country capitals for each observed country-tuple (backer's and entrepreneur's home countries) are identified. Subsequently, the geographical coordinates are used to estimate the geodesic distance between the two interacting parties. [8, 9] The construction of the distance variable is a labour-intensive and time-consuming process, which partly explains why only few studies have focused on this research area in the past (cp. Guo et al., 2018). However, by knowing the geographical distance between backers and entrepreneurs in occurred crowdfunding transactions, this thesis can examine underlying patterns in the data through quantitative analysis (cp. Research Objective 3) and provide a potential answer to the question of whether Home Bias exists in reward-based crowdfunding on Kickstarter.

Research Objective 3: *To construct a Negative Binomial regression model that examines the effect of geographical distance (between backers and crowdfunding project creators) on the count of backing actions from a given country.* Negative Binomial regression is the most recommended approach in the literature to model count data (Cameron & Trivedi, 2010). The reason is that this category of regression models is specifically designed to handle non-normally distributed variables and can account for the special right-skewed distribution of the dependent variable (count of backers), which only allows discrete and non-negative values. Five additional control variables (GDP per

---

[8] More detailed explanation of the approach is provided in section 3.4.4.2.

[9] The geodesic distance is the shortest distance on the surface of an ellipsoidal model of the earth (Karney, 2013).

capita, third-party endorsements, project category, herding behaviour and Covid-19 pandemic) are included in the model that are likely to have a significant effect on the dependent variable and are highly interesting for both theorists and practitioners.

Research Objective 4: *To address potential caveats of dealing with Big Data models in crowdfunding research.* To identify the effect of increasing sample size on statistical key figures (such as p-values), coefficient/p-value/sample-size (CPS) charts are used. CPS charts are created by repeatedly drawing samples of increasing sizes, re-estimating the same statistical model, and plotting the p-values and coefficients on a chart. The created charts provide an insight into the statistical power of the model and can help to reveal potential biases (Lin et al., 2013). Following the identification of the effect of the sample size via CPS charts, this thesis has a strong focus on analysing effect sizes, which is recommended by current literature for dealing with Big Data models (Hryniewicz, 2018; M. Lin, Lucas, & Shmueli, 2013; Singh Chawla, 2017). Estimating effect sizes can help scientists to verify if the obtained results are not only statistically significant but also relevant in practice. To evaluate the practical relevance of the influence of distance (and other variables) on the count of backers, this thesis provides different measures for effect size including Incidence Rate Ratios (IRRs), Marginal Effect at the Mean (MEM) and Cohen's d.

Research Objective 5: *To employ a robustness test to evaluate the firmness of the findings in Research Objective 3 and 4.* This robustness test is an important extension of the preceding analysis because it examines whether the findings are potentially biased by the large number of US investors and entrepreneurs in the dataset. In the existing literature on crowdfunding, most studies focus exclusively on the US market and deliberately exclude projects or backers from foreign countries (cp. Mollick, 2013; Agrawal et al., 2013; Guo et al., 2018; Carbonara, 2020). The strong focus on the US is understandable because the US demonstrates a pioneering role in crowdfunding. However, this focus is not sustainable any longer because the crowdfunding industry is becoming increasingly international (cp. Guo et al., 2018).

Because of the described reasons, Guo et al. (2018) suggest that the findings of prior studies might be valid for American investors but lack the validity for other countries. By constructing the largest and most recent crowdfunding dataset to date (Research Objective 1), this thesis can exclude US participants (backers and entrepreneurs) from the dataset for robustness check and still provide sufficient data for a quantitative analysis. This analysis provides a new perspective on international crowdfunding and might reveal potential biases of prior research.

Figure 2 provides an overview of the described research aim, research question and the different research objectives that guide this thesis.

Figure 2: Overview Research Aim, Question and Objectives

This Figure provides an overview of the research aim, question, and objectives for this thesis.

Research Aim

Examine the existence of the Home Bias phenomenon in the emerging industry of international reward-based crowdfunding.

Research Question

Does geographical distance between backers and entrepreneurs affect the overall count of project supporters in international reward-based crowdfunding?

Research Objective 1

To construct the largest and most recent crowdfunding dataset to date.

Research Objective 2

To use the collected data from Kickstarter to calculate the distance between backers and entrepreneurs.

Research Objective 3

To construct a Negative Binomial regression model that examines the effect of geographical distance (between backers and crowdfunding project creators) on the count of backing actions from a given country.

Research Objective 4

To address potential caveats of dealing with Big Data models.

Research Objective 5

To employ a robustness test that excludes US participants to evaluate the firmness of the findings.

1.7 Contributions to Knowledge and Practice

This thesis makes multiple important contributions to theory and practice in the field of crowdfunding.

The basis for these contributions is introducing the most comprehensive dataset yet collected in crowdfunding research. In total, the dataset comprises 1,118,654 observations from 211,695 crowdfunding projects that represent more than 44 million individual backing actions.[10] By covering all projects that were launched on Kickstarter

[10] For comparison, Guo et al. (2018), who also examine geographical data extracted from Kickstarter, use a sample of 136,234 crowdfunding projects.

between April 2009 to June 2020, this thesis incorporates not only the longest period but also the most recent data on crowdfunding.[11] The extensive scope of the dataset allows to provide the most up-to-date overview of the industry and an insight into cross-border backing activities, which to date have received insufficient attention in existing literature (cp. Guo et al., 2018). Although some preceding work has examined the existence of Home Bias in crowdfunding, most of the published studies are restricted to the analysis of individual countries (cp. Mollick, 2013; Agrawal et al., 2013; Guo et al., 2018; Carbonara, 2020). Particularly salient is the frequent focus on crowdfunding in the US (K. Kim & Hann, 2013; M. Lin & Viswanathan, 2016; see Mollick, 2013, 2014). The reason for this focus is that many crowdfunding platforms originate from the US (i.e., Kickstarter and Indiegogo), making it the most mature market for crowdfunding. However, national focus is no longer justified or contemporary because the industry is becoming increasingly international and a growing number of individuals worldwide are using crowdfunding for corporate financing (Massolution, 2015; Niemand et al., 2018). Guo et al. (2018) recognize this knowledge gap and suggest that further research should incorporate data from different cultures and regions. This thesis addresses this weakness by incorporating crowdfunding projects from more than 200 different countries, making it the most comprehensive analysis of international reward-based crowdfunding to date. The provided findings are important contributions to knowledge because they shed a new light on the potential of crowdfunding to reduce the Global Finance Gap problem (as described in section 1.1.) and promote higher economic welfare globally.

In terms of methodology, this thesis makes a valuable contribution to current knowledge and practice by introducing an adapted version of the Negative Binomial regression to model the count of backers (from a specific country) as a function of geographical distance and other relevant variables. This is the first study to use the count of backers

[11] For comparison, the most recent crowdfunding study of Walther (2021) includes data on crowdfunding projects only until January 2019.

as a dependent variable to examine the effect of Home Bias in reward-based crowdfunding.[12] The advantage of this approach is that the dependent variable is in easily comprehensible units (count of backers) with high practical relevance. Entrepreneurs are interested in the expected count of backers from the specific countries because it helps them to design their crowdfunding projects more effectively (cp. Kickstarter 2017). For example, knowing where most backers will come from allows project creators to focus on the relevant countries (i.e., in terms of language and currency), design more efficient marketing activities (i.e., by targeting the respective countries) and better estimate the prospective shipping costs. The findings of this thesis, thereby, provide valuable insights for both scientists and practitioners.

Another important contribution of this thesis, in association with the dependent variable, is that it introduces a new approach in Home Bias research to counterbalance the potential influence of backers that maintain a social connection to the entrepreneur (i.e., family and friends). The literature suggests that these backers often do not make their decisions objectively but are influenced by the personal relationship to the entrepreneur which, in turn, leads to a higher willingness to support a specific project (Guo et al., 2018; Kuppuswamy and Bayus, 2013). Moreover, it can be assumed that most of these backers share the same country of residency as the entrepreneur. The result is that crowdfunding projects might receive comparably more backers from the same country of the project founder than from other countries. This, in turn, might distort the outcome of crowdfunding research in favour of the existence of Home Bias. To account for this potential distortion, this thesis is the first to introduce the threshold of 150 (backers) to limit the count of backers that share the same country of origin as

[12] Prior studies used pledge results (Guo et al., 2018), probability of transaction (Lin & Viswanathan, 2013, Agrawal et al., 2011), or funding success (Mollick, 2013) as the dependent variable.

the entrepreneur.[13] The used threshold is based on Dunbar's (1992; 1993) findings that individuals are usually unable to maintain a stable social relationship with more than 150 different people at the same time (for more information see section 3.4.4.1).

In the process of analysing the influence of geographical distance, this thesis also examines the effect of five additional control variables (GDP per capita, third-party endorsements, project category, herding behaviour and Covid-19 pandemic) on the count of backers in international reward-based crowdfunding. The particular contributions to knowledge and practice resulting from this analysis are explained for each variable in more detail below:

- First, the model controls for the GDP per capita of both backers' and entrepreneurs' home countries. The analysis of this variable can help scientists and policymakers to answer the question of whether reward-based crowdfunding improves the general distribution of financial resources by promoting investment flows from high-income to low-income countries. The general assumption of this thesis is that backers in countries with high GDP per capita are more likely to participate in crowdfunding due to their surplus in financial resources. Simultaneously, projects in countries with low GDP per capita might tend to attract more funds because backers might be influenced by a philanthropic component that motivates them to support entrepreneurs in economically weak regions of the world (cp. Burtch et al., 2014). Although some prior research has examined the influence of GDP on fundraising performance in lending-based crowdfunding (e.g., M. Lin & Viswanathan, 2016; Burtch et al., 2014), no research exists in the field of reward-based crowdfunding. Moreover, no prior research studies the influence of the GDP per capita (as opposed to GDP),

[13] Some of the prior research estimates the influence of family and friends by studying the social networks of the founders online (cp., Agrawal et al., 2013, Kuppuswamy & Bayus, 2013). However, this approach is no longer feasible because it would violate the ethical principles of this thesis. This is because the privacy policy of platform operators strictly forbids the automated collection of personal user data (cp. Kickstarter 2018).

which is considered a better measurement of individual wealth (OECD, 2009). This thesis is the first to examine the effect of GDP per capita (of backers' and entrepreneurs' home countries) on the count of backers in international reward-based crowdfunding.

- The model also controls for third-party endorsements, referred to as "Projects We Love" (PWL) on Kickstarter, and examines how this variable influences the count of backers. This variable is particularly important for entrepreneurs as it might provide them with an insight on whether it is a rewarding strategy to strive for a third-party endorsement or whether they should rather use their (often limited) time and resources to pursue other more effective quality signals instead to attract interest in their projects. Prior research provides first evidence that third-party endorsements do have a positive influence on the success of crowdfunding campaigns (see Mollick, 2014; Qiu, 2013). However, the existing research focuses exclusively on American crowdfunding participants and does not indicate whether third-party endorsements have a similar impact on backers from other countries. Moreover, an assessment of the concrete effect size of such endorsements is missing that would enable entrepreneurs to assess its advantage in relation to the required effort to attain the signal of quality for their projects. Besides this contribution to practice, this thesis also makes a valuable contribution to knowledge because it is the first to examine the influence of third-party endorsements in an international context in reward-based crowdfunding. Furthermore, it is the first to provide a comprehensible metric for the actual effect size of third-party endorsements (as the number of additional backers) that could help entrepreneurs to better assess the value of this quality signal when developing their crowdfunding campaigns.
- Another variable is included in the model to control for the different project categories. In this respect, it builds on the findings of K. Kim and Hann (2013) and Guo et al. (2018) who showed that the chosen project category can significantly affect crowdfunding performance. For example, both studies show that technology-related ventures tend to perform exceptionally well in

crowdfunding, whereas projects related to “Food” and “Theatre” perform considerably worse. This thesis is the first to examine the influence of project category on the count of backers in an international context. By knowing the expected count of backers for each specific project category, entrepreneurs can have a better estimate of the suitability of reward-based crowdfunding for their fundraising intentions.

- A fourth variable is included in the model to control for the effect of herding behaviour. Herding behaviour is an important trust mechanism that can help market participants to deal with information asymmetry and, thereby, significantly affect the success of crowdfunding projects. To estimate the potential effect of herding behaviour, this thesis constructs a dummy variable that specifies if a project reached a specific threshold of total backers and can therefore be considered a “Large Project”. This thesis uses a threshold of 107 backers, which is the median for the total count of backers in crowdfunding projects across the sample. The method introduced in this thesis is a new approach to account for herding behaviour in crowdfunding research. It provides researchers with a tool to control for herding behaviour especially when additional information on the development of crowdfunding projects over time is not available.
- A fifth and final variable is added to the model to examine the effect of the Covid-19 pandemic on crowdfunding projects. This analysis is particularly interesting for both scientists and practitioners because no prior research has addressed how global economic crises might affect the crowdfunding industry. This is because crowdfunding became popular only after the financial crisis in 2008 and the Covid-19 pandemic is still too recent to have been considered by previous research. Theoretically, the crowdfunding industry could be affected in both ways by the pandemic. On the one hand, the crowdfunding market could have attracted lots of interest from backers and entrepreneurs during Covid-19 pandemic because it demonstrates an alternative fundraising source that can be accessed via the internet from the comfort of one’s own home. It provides

entrepreneurs an unprecedentedly fast and wide-reaching possibility to raise capital for their business ventures. Moreover, the increasing need for capital to ensure their survival during the economic crisis caused by the pandemic as well as the decreasing trust in and difficulty to access traditional financiers (caused by the global financial crisis in 2008) could have motivated many entrepreneurs and business owners to explore new paths and take advantage of this alternative form of fundraising. On the other hand, the crowdfunding industry might be negatively affected by the Covid-19 pandemic due to the many company closures and mass layoffs (International Labour Organization, 2020). The widespread economic uncertainty might limit the willingness of many private backers, which are the centrepiece of the crowdfunding concept, to take the additional risk of investing into other businesses. This thesis is the first to examine this issue and provide a preliminary insight into how a global economic crisis could affect the crowdfunding industry.

This thesis also makes a valuable contribution to crowdfunding research by addressing potential biases that have remained unnoticed or unexplored by existing literature. For example, while many studies in crowdfunding research use comparatively large data samples (cp. Agrawal et al., 2015; Burtch et al., 2014; Guo et al., 2018; Gallemore et al., 2019), none of them have considered the potential problems that can result from such big samples. The p-value problem, which is associated with Big Data, is a recognized issue in other research fields such as statistics and medical research (Kühberger et al., 2015; B. Thompson, 2007; N. Thompson et al., 2021), but remains vastly ignored in current crowdfunding literature. This thesis is the first to raise this issue and to highlight that in Big Data models, researchers must be careful with inferring practical relevance from statistical significance of their findings. Moreover, this thesis describes different approaches and tools that could help future scientists in crowdfunding research to identify potential problems related to sample size (i.e., through CPS charts) and verify the practical relevance of their findings (i.e., by using effect sizes and confidence intervals). By reporting p-values, confidence intervals and effect sizes, the “Big Three” (Hatcher, 2013), this thesis provides the most comprehensive analysis of the statistical and practical significance of Home Bias (and other variables) in international reward-based crowdfunding.

Another bias that is first addressed by this thesis relates to the high dominance of US data in current crowdfunding research. Studies typically focus exclusively on crowdfunding in the US or use data samples that are highly dominated by US backers and entrepreneurs (cp. Guo et al, 2018). To date, none of the research in crowdfunding has examined the impact of this potential bias on their results. The explanation for this might be the fact that most studies have not been able to obtain sufficient data on crowdfunding projects from other countries. The extensive data sample of this thesis, however, allows to exclude US participants (backers and entrepreneurs) at a later stage in the analysis for robustness check and still provides sufficient data for a quantitative analysis. By re-estimating the model developed in this thesis on a subsample that excludes US participants, this thesis is the first to examine the potential bias introduced by the high dominance of US related data.

1.8 Research Scope

Figure 3 summarizes the scope of this thesis and facilitates the categorization of this research into the existing stream of crowdfunding literature. This thesis focuses on the reward-based model, which is currently the most international and widespread crowdfunding model for business financing. The Kickstarter platform is chosen because it is the international market leader for reward-based crowdfunding (cp. section 3.6.1). In terms of methodology, this thesis is the first to use the Negative Binomial regression model to examine the influence of different influencing factors (e.g., distance) on the count of backers. This research also differentiates itself from most previous studies because it examines crowdfunding in an international context. Moreover, this research uses the most extensive data sample, by covering all projects that were launched in the time frame from April 2009 to June 2020.

Figure 3: Research Scope

This Figure summarizes the scope of the research and describes how it is positioned in the crowdfunding literature.

Crowdfunding Model	**Reward-Based**	*Others (e.g., Lending-Based)*
Crowdfunding Platform	**Kickstarter**	*Others (e.g., Indiegogo)*
Methodology	**Negative Binomial**	*Others (e.g., Logistic Regression)*
Scope	**International**	*National*
Time Frame	**2009-2020**	*Other years*

1.9 Thesis Structure

Chapter one provides the context for this thesis by explaining how the Global Finance Gap inhibits businesses, especially start-ups, in their development and how solving (or alleviating) this problem could lead to greater global prosperity. Home Bias is identified as a one major contributor to the Global Finance Gap because it promotes an unequal distribution of capital, which is particularly detrimental to businesses outside of established business regions (i.e., Silicon Valley). This insight justifies the overall focus of this thesis on Home Bias. Crowdfunding is introduced as a new approach for business fundraising that, due to its digital nature, bears the potential to reduce (or eliminate) the influence of Home Bias. The first chapter outlines the main research gaps in the literature on Home Bias in crowdfunding, establishing the research aim and objectives for this thesis. The chapter also introduces the intended contributions of this thesis to both knowledge and practice.

Chapter two provides a review of the relevant literature on Home Bias and crowdfunding. It explores the current knowledge on Home Bias including the recognized factors that promote and alleviate it. In this context, different reasons and theories are presented as to why the emerging concept of reward-based crowdfunding could provide a potential solution to the Home Bias problem in business financing. Subsequently, the relevant literature on the existence of Home Bias in the crowdfunding industry is reviewed. Overall, the chapter discloses the ongoing controversy on whether crowdfunding can alleviate the influence of Home Bias in business financing. Simultaneously, it highlights the current research gaps and provides justification for the focus of this thesis.

Chapter three begins with explaining the philosophical position of the author. The intention of this section is to reveal potential biases that might influence the research. Subsequently, the research methodology is presented. Considering the relevant literature, the Negative Binomial regression model that this thesis uses to examine the influence of Home Bias in crowdfunding is described. The variables and the process of data collection is explained in detail. The chapter also describes how to interpret the model results. Special attention is given to potential biases that can arise when dealing with Big Data samples. In this context, the meaning of statistical and practical significance is discussed critically. Moreover, different approaches recommended by the current statistics literature are presented that should help scientists to ensure the practical validity of their findings when dealing with Big Data models. The chapter concludes with a description of the used data and an exploratory data analysis.

Chapter four presents the results of the Negative Binomial regression. The results are initially discussed in terms of their statistical significance and implications. Additionally, considering the interpretation recommendations for Big Data models (in Chapter Three), the results are analysed in detail for their practical effect sizes. This analysis provides a different perspective on the preliminary findings and reveals a potential bias that might not be unique to this thesis but has potentially been underestimated in prior studies on crowdfunding. The chapter concludes with an additional robustness test to evaluate if the findings are influenced by the large number of US backers and US entrepreneurs in the dataset.

The thesis ends with Chapter Five which provides the overall conclusion of this research on whether Home Bias exists in international reward-based crowdfunding. The chapter recapitulates the approach, most important findings, their implications, and potential limitations of this research. Moreover, suggestions for future research are provided.

Chapter Two: Literature Review

2.1 Introduction

This chapter provides an overview of the underlying theoretical framework for this thesis. Central attention is dedicated to the Home Bias phenomenon, which is the tendency of individuals to choose geographically proximate interaction partners. Different theories are presented that provide possible explanations and drivers for Home Bias.

Building on this knowledge, the chapter also reviews the literature on potential factors that might affect the Home Bias issue and may alleviate the influence of distance. A special focus is paid to crowdfunding literature where the current state of knowledge on Home Bias is critically reviewed.

2.2 Home Bias Theory

2.2.1 The Home Bias Phenomenon

The Home Bias theory describes the tendency of individuals and institutions to interact with physically proximate partners (Niemand et al., 2018). It is first described by French and Poterba (1991) in portfolio investment decisions. The authors observe that despite the recognized benefits of international diversification, most investors hold their wealth in domestic assets. For example, French and Poterba (1991) show that more than 98% of the Japanese equity investments are held in domestic firms. Similarly, American equity portfolios exhibit less than 6% of their investments in foreign assets (French & Poterba, 1991). Moreover, the authors emphasize that the tendency to invest into national markets is due to the investors' own choice, rather than due to institutional constraints (see also Baltzer et al., 2012).

Subsequent research reveals that Home Bias is not limited to international portfolio investment decisions but can also be found in different contexts such as international trade, consumer purchasing behaviour and sport betting (Karlsson & Nordén, 2007; Noton, 2015; Petersen & Rajan, 2002; Staněk, 2017). Bernanke and Rogoff (2001) describe Home Bias as one of the six major puzzles in international macroeconomics.

Since French and Poterba (1991), different publications have confirmed the existence of Home Bias in business financing at national and international level. For example, Coval and Moskowitz (1999) show that American investment managers exhibit a strong preference for locally headquartered firms. A greater geographical distance between investors and target firms is found to lower the chances of collaboration. Sorenson and Stuart (2001) find that the average distance between venture capital investors and target firms is less than 70 miles (see also Cumming & Dai, 2010). Although distant investors are common in publicly listed companies, research suggests that start-ups and small businesses receive their financial resources almost exclusively from local providers. According to Amel and Brevoort (2005), more than 90% of small business lending in the US comes from local bank branches.

Home Bias is considered a suboptimal market behaviour that can negatively affect both investors and entrepreneurs (French & Poterba, 1991, Coval & Moskowitz, 1999; M. Lin & Viswanathan, 2016). For example, investors that focus exclusively on firms in their domestic market are highly dependent on the performance of the local economy. This dependence leads to an increased volatility risk because an economic downturn in the home country has a direct effect on the entire portfolio performance (Coval & Moskowitz, 1999; Tesar & Werner, 1995). Moreover, the financial losses might last longer because investors must rely on the national economy to recover (M. Lin & Viswanathan, 2016). The focus on one country also leads to increased opportunity costs, because investors can miss chances in faster growing economies (Petersen & Rajan, 2002; Tesar & Werner, 1995).

Entrepreneurs are equally negatively affected by Home Bias because it reduces the availability of potential financiers and makes the access to capital more difficult (Niemand et al., 2018). The search for funding sources is particularly difficult for SMEs that are located apart from established business clusters or find themselves in countries with insufficiently developed financial structures (World Bank, 2012). The missing competition among investors can also lead to unfavourable financing conditions and affect firms in their general competitiveness (Guiso et al., 2002). The lack of capital restricts business activities of companies and is considered one of the major hurdles for development (World Bank, 2012). Entrepreneurship becomes highly dependent on the local conditions of the financial industry (Bruhn & Love, 2009; Dinh et al., 2010; Klapper et al., 2007).

2.2.2 Drivers for Home Bias

Research provides different explanations for the Home Bias phenomenon (Glassman & Riddick, 2001; Lewis, 1999; Strong & Xu, 2003). However, most authors agree that besides culture dependent preferences for regional products (e.g. music) and location specific goods (e.g. concert tickets), the Home Bias phenomenon is mainly caused by the failure of actors to establish a relationship of trust (Blum & Goldfarb, 2006; Huberman, 2001; M. Lin & Viswanathan, 2016; McPherson et al., 2001). In this context, great importance is attributed to the *Information Asymmetry Problem,* which provides a possible explanation why people prefer geographically close interaction partners (Ahlers et al., 2015; Dziuda & Mondria, 2012; Kang et al., 2017; K. Kim & Hann, 2013).

The term information asymmetry describes the situation in which two interacting market participants do not have access to the same amount or quality of information (Akerlof, 1970). Information asymmetries are ubiquitous and can be found in most markets where products or services are traded. Akerlof (1970) is one of the first authors to describe the problem of information asymmetry and to study its effect on markets. The author illustrates the information asymmetry problem by using the example of a used car purchase. A person interested in buying a used car typically compares different options in terms of quality and price. In most cases, salespeople have additional information about the product that is not immediately available to the buyer.[14] For example, the salespeople might know about previous accidents or certain malfunctions of the car and could withhold this information from the buyer to achieve the highest possible sales price. The buyer has often few objective sources of information about the current condition of the car and must rely on the provided information. Akerlof uses this example to explain that markets with heterogeneous quality of goods and limited information sources are prone to the *adverse selection* problem. The adverse selection problem describes the scenario in which a privileged market participant (i.e., the car salesman) exploits their position at the expense of other market participants (i.e., selling a low-quality car at the price of a high-quality car). The information asymmetry and the associated adverse selection problem can lead to a scenario in which market participants fear to make a fair deal and therefore withdraw from the interaction, diminishing the volume of trade in the market (Akerlof, 1970). On the long-term, if no regulatory measures exist, this behaviour will lead to the collapse of the market.

According to Spence (1973), many of the observed market inefficiencies can be explained through the information asymmetry problem and the resulting lack of confidence in the interaction. Consequently, they are also considered one of the major drivers for the Home Bias phenomenon. In business financing, information asymmetries are common because entrepreneurs usually have better knowledge about their company's condition, weaknesses and opportunities than do investors. The information asymmetry problem is particularly common in early stage fundraising where public or historical information on the target firm is scarce (Bernstein et al., 2014). At the same time, the adverse selection problem can occur in business fundraising because entrepreneurs have a strong interest in concealing their weaknesses as they aim for the highest possible company valuation.

[14] In certain cases, such as the insurance industry, also the buyer can have an information advantage.

The literature identifies three different factors that lead to the choice of geographically proximate investment targets amid information asymmetry: (1) Reduction of Transaction Costs, (2) Reliance on Social Networks, and (3) Cognitive Biases.

2.2.2.1 Reduction of Transaction Costs

One reason for choosing geographically proximate interaction partners, especially in the pre-internet era, is that the exchange of information can be facilitated. For example, investing in close firms allows personal meetings with the founders, constant performance monitoring and the attainment of tacit knowledge (Jaffe et al., 1993; Zucker et al., 1994). Moreover, investors are usually already familiar with the environmental conditions such as the state of the economy or the legal regulations. This knowledge can facilitate the enforcement of contracts and reduce different forms of research costs (Hortacsu et al., 2006). On the contrary, investing in distant firms often requires additional time and effort to study the legal circumstances or to deal with additional obstacles such as the currency exchange risk (Lewis, 1999). The choice of geographically proximate target firms can therefore be driven by the rational intention of investors to reduce different forms of transaction costs (Ahearne et al., 2004; Glassman & Riddick, 2001; Rowland, 1999).

2.2.2.2 Reliance on Social Networks

Another reason for choosing geographically proximate target firms is that investors can rely on their personal social network (Zane & DeCarolis, 2016; Shane & Stuart, 2002). Two explanations are provided in the literature why social networks are important for investors in the context of business financing. First, research has found that investors typically identify promising investment targets through referrals from entrepreneurs they have previously sponsored, fellow venture capitalists, family members, or other professional contacts (Fried & Hisrich, 1994; Hsu, 2004; Shane & Stuart, 2002). Second, research shows that investors particularly prefer to invest in companies that they have discovered through their personal network because it allows them to acquire additional information about the entrepreneur's reliability and integrity that otherwise would be unavailable to them (Zane & DeCarolis, 2016). For example, when the investor's trusted contacts offer assessments of an entrepreneur, these evaluations often lack the perception of bias that discredits information provided directly by the entrepreneur (T. E. Stuart & Sorenson, 2005). At the same time, close contacts have an incentive to provide accurate and complete information, as otherwise they jeopardize the perception of their personal credibility and integrity (J. S. Coleman, 2000). Overall, these mechanisms contribute to a higher quality of information that can be obtained through social networks and help to alleviate the information asymmetry. However, Sorenson (2018) highlights one important problem of this approach. If investors rely heavily on their social networks to find potential investment targets, industries will tend to become and remain geographically concentrated, even when firms do not benefit from this clustering (Sorenson, 2018). The reason for this is that strong social relationships tend to be local (Agrawal et al., 2010; Wellman et al., 1996). Moreover, relying on personal contacts can also be to the detriment of newcomers and entrepreneurs that are not part of an established network. Some researchers see a potential problem if resources are distributed based on social relationships rather than the principles of demand and supply or the actual quality of the business idea. For example, Mollick and Robb (2016) argue that the allocation of financial resources is and has historically always been a profoundly undemocratic process. They argue that the decision which businesses will receive the necessary funding is mostly made by a small elite of highly connected white men (Mollick and Robb 2016). Entrepreneurs that manage to receive funding are often themselves part of this elitist group. Studies confirm this tendency by showing that only 2.7% of the venture capital (VC) backed companies have a female CEO (P. G. Greene et al., 2014). Similarly, Timothy Bates and Bradford (2008) find that only a small fraction of VC-funded start-ups were founded by African Americans in the US. Different authors present evidence that lending- or investment decisions are often biased by factors such as gender, race and social background (Agrawal et al., 2011; T. Bates & Bradford, 2007; Blanchflower et al., 1998; Brush et al., 2014; Cavalluzzo et al., 1999; Chen et al., 2009; S.

Coleman & Robb, 2009; Patricia G. Greene et al., 2001; Harrison & Mason, 2007; Holmquist & Carter, 2009; Kenney, 1999; Miller, 2010; T. Stuart & Sorenson, 2003).

2.2.2.3 Cognitive Biases

Kahneman and Tversky (2008) provide a potential link between missing trust and the Home Bias phenomenon by arguing that the inability to evaluate the trustworthiness of a business partner motivates people to rely on *heuristic techniques*. Heuristic techniques are mental shortcuts of problem solving in scenarios where finding an optimal solution is difficult (Myers, 2009). An important characteristic of heuristic techniques is that they are often not logical or rational, but instead provide a fast and temporarily satisfactory solution to the problem (Kahneman, 2013). Kahneman and Tversky (1981) linked heuristics to cognitive biases and showed that in situation of uncertainty, people frequently use the strategy of *attribute substitution*. For example, the geographical proximity of a business partner is used as a substitute for trust and can engender "over-optimism" about the transaction (Hortaçsu et al., 2009; Lai & Teo, 2008; Strong & Xu, 2003). This assumption itself is a cognitive bias, meaning a systematic error in the use of heuristics, since the geographical proximity cannot necessarily be linked to the trustworthiness of a business partner (Ahlers et al., 2015). The approach of attribute substitution provides a potential explanation for the Home Bias phenomenon in the situation of high information asymmetry.

2.2.3 Signalling- and Screening Theory

Different authors have studied the possibilities of actors to reduce the information asymmetry problem and establish a relationship of trust. One important approach is provided by the theories of Spence (1973) and Stiglitz (1974) on *signalling* and *screening*, respectively.[15] The fundamental idea of signalling is that information holders can use certain objective quality signals to distinguish themselves from fraudsters or less qualitative product providers. An important requirement is that the signalling costs, which for instance can be measured by the effort, money or dedicated time to attain the signal, must be reciprocally proportional to the quality of the service or product (Spence 1973). Spence (1973) illustrates the signalling theory on the example of the job market by arguing that job applicants, which are usually difficult to evaluate for employers in advance, can invest their time in attaining a university certificate that could distinguish them from less qualified candidates. In this example, Spence explains that higher qualified applicants will find it significantly easier (i.e., less time and effort) to attain a university certificate than their less qualified counterparts. At the same time, some applicants will refrain from pursuing a graduation certificate because it requires more effort than they are willing to accept. Thereby, the university certificate can serve as an objective and reliable signal of quality.

While signalling describes the possibilities of information holders (e.g., the entrepreneurs) to objectively communicate quality, screening refers to the efforts of the underinformed parties (i.e., investors) to evaluate the quality of the transaction partner (Spence, 1973). In business financing, a growing body of literature studies potential quality signals that investors can use to choose their investment targets. For example, research has found that many successful investors focus on the founders, their professional background and former success stories (Bernstein et al., 2014; Boeuf et al., 2014; Colombo et al., 2015; Mollick, 2013). Moreover, the degree of preparedness of the pitch is often used as an indicator for the competence of the entrepreneur (Mollick 2014; Frydrych et al. 2014; Chen et al., 2009). The techniques of signalling and screening provide possible explanations on how markets with supposedly high information asymmetries can work successfully. However, the current distribution of venture capital

[15] Akerlof (1970), Spence (1973) and Stiglitz (1974) received the Nobel Prize in Economics for their collective findings on the information asymmetry problem and signalling theory.

is still considerably uneven (see PwC & CB Insights, 2019; Florida & King, 2016; Mollick, 2014; Wiklund & Shepherd, 2005). This fact suggests that most of the hitherto identified quality signals and strategies are insufficient to outweigh the drivers for the Home Bias as described in section 2.2.2. Start-ups continue to be dependent on local finance providers and are largely unable to attract distant investors (World Bank 2016a, Alibhai et al., 2017).

2.3 The Potential of Crowdfunding

2.3.1 The Rise of the Internet

The internet has introduced a fundamental shift in how people communicate and interact (Brynjolfsson et al., 2009; Cairncross, 1998; Choi & Bell, 2011; Forman et al., 2009). One important characteristic of the internet is that it greatly alleviates search and interaction costs, allowing an unprecedented speed and scope of information exchange (Blum & Goldfarb, 2006). Therefore, Agrawal et al. (2010) argue that the internet is also eliminating or greatly reducing the influence of geographical distance in certain business areas. Some popular examples include the impact of the internet on the significant increase in international retail transactions and the cross-border development of software. Referring to these examples, Friedman (2005) famously hypothesized that the world is becoming increasingly “flat” and integrated.

Crowdfunding is one example of how technological innovations have disrupted traditional markets (Newman et al., 2016). It has fundamentally changed the financial sector (Cumming et al., 2019; Mollick & Robb, 2016). The major advantage of crowdfunding is that it relocates the fundraising process into a virtual space. Thereby, it significantly increases the potential scope and speed of fundraising activities. These new characteristics have led to the assumption that crowdfunding could also potentially lower the relevance of geographical distance between investors and entrepreneurs (K. Kim & Hann, 2013). Theoretically, any individual who has access to the internet can participate in crowdfunding on the donor or receiver side. Entrepreneurs can approach a large and highly diversified group of investors, independent of their physical location, and investors can use crowdfunding platforms to search for promising business ideas from across the globe. The new possibilities have led to an increased interest into the industry and motivated different scientists to study if crowdfunding can solve some of the distance related challenges in business financing (Lin & Viswanathan, 2016; Agrawal et al., 2013).

2.3.2 Signalling and Screening in Crowdfunding

The increasing affordability of technology combined with innovative electronic marketplaces (i.e., Amazon, Shopify or Patreon) have created entirely new possibilities for individuals to engage in the digital world. Launching a website, webstore, or crowdfunding campaign became increasingly cheaper and easier (Mollick, 2013). The new possibilities led to a significant increase in available business partners and raised new issues in terms of trust (Agrawal et al., 2013). Crowdfunding, and especially the reward-based model, is a prime example for this development. Unlike in traditional fundraising, crowdfunding campaigns can be set up relatively quickly and at low costs. Campaign creators in crowdfunding are not obliged to provide a business plan, financial data, or a detailed risk assessment. Creators only present some key characteristics of their product or business idea and potential investors must evaluate these ideas based on this limited information (Niemand et al., 2018). The entire fundraising process takes place on the internet and the interacting parties never meet in person. Due to the described characteristics, the reward-based crowdfunding model is often considered the fundraising approach with the lowest information disclosure requirements for entrepreneurs and as a market that is particularly prone to the information asymmetry problem (Ahlers et al., 2015; Dziuda & Mondria, 2012). According to Akerlof (1970), markets that demonstrate these characteristics are threatened to collapse on the long term. The crowdfunding market, however, has existed for more than a decade and continues to expand globally (Niemand et al., 2018).

A growing body of literature suggests that backers in crowdfunding do not make poorer decisions than professional investors, although they have significantly less information at their disposal and usually less experience in the evaluation of companies (Mollick & Robb, 2016). According to Mollick and Kuppuswamy (2014), the general crowd is particularly good at identifying promising business ideas as more than 90% of the companies remain ongoing ventures after the successful termination of the crowdfunding campaign. One possible explanation for the post-campaign success is that crowdfunding also helps entrepreneurs to raise awareness and build a large and loyal customer base. For this reason, many studies examine the decision process of backers more in detail and try to understand the underlying mechanisms that enable the market to work successfully (Bernstein et al., 2014; Boeuf et al., 2014; Colombo et al., 2015; Mollick, 2013).

Early research found that backers partly have similar approaches for identifying promising business ideas as do traditional investors. For example, Marom and Sade (2013) show that backers equally focus on the information about the founders, their professional background and former success stories. Mollick (2014) finds that backers often assess how well the entrepreneurs have prepared the pitch. In this context, a high amount of spelling errors in the project description or a missing campaign video are considered a sign of bad preparedness and can lower the chances of successful fundraising. Xu et al. (2014) find that frequent updates about the company's progress have a positive influence on the success of the crowdfunding campaign (see also Kuppuswamy & Bayus, 2014).

The described strategies resemble the approaches of traditional (offline) investors (Mollick 2014; Frydrych et al. 2014; Chen et al., 2009). However, an increasing number of studies also finds important differences. For example, backers seem to be considerably more affected by emotional factors than traditional investors. According to Frydrych et al. (2014), the narrative, meaning the establishment of a convincing and compelling product story plays a significant role in successful crowdfunding. The scientists find that projects that only provide a factual product description have lower probabilities of achieving the funding target. Therefore, the authors conclude that successful fundraising requires a campaign video that “touches the heart” of backers and tells an emotional and engaging story about the product or entrepreneur (Wheat et al., 2013). Also, Kuppuswamy and Bayus (2014) suggest that entrepreneurs should try to awaken emotions and excitement, especially in the final stage of the fundraising if they aim to increase the overall funding. Davis et al. (2017) study the influence of the campaign video more in detail. The researchers find in their experiment with 102 participants that the performance of a crowdfunding campaign is highly dependent on two factors: The perceived passion of the entrepreneur and the perceived product creativity. This is an interesting aspect because it is partly in contradiction to the finding of Chen et al. (2009) that in traditional venture capital pitches preparedness, not passion, positively impacts the investment decisions.

The specific literature on crowdfunding suggests many potential quality signals that backers and entrepreneurs use to alleviate the information asymmetry problem (see Table 1). However, these signals do not always meet Akerlof’s (1970) traditional definition, which states that the costs of attaining a signal must be inversely proportional to the quality of the product. Quality signals that do not meet this requirement can be easily copied by other market participants and are threatened to lose their value on the long term.[16] In the literature, mainly three quality signals have been repeatedly suggested that meet Akerlof’s requirements and provide a plausible explanation how the crowdfunding industry can work successfully. These signals are: (1) crowd-based trust mechanisms, (2) trustworthy intermediaries, and (3) reputation signalling through social networks (see also Agrawal et al., 2013).

[16] One example is Mollick’s (2013) finding that the number of images and videos in the campaign description is a quality signal that positively influences the crowdfunding success. This finding has been disproved by subsequent research when Frydrych et al. (2014) showed that images and videos became industry wide standards and do not demonstrate a positive influence on the crowdfunding performance anylonger.

Table 1: Quality Signals in Crowdfunding

This table provides an overview of the research on quality signals in crowdfunding.

Authors	Finding
Marom and Sade (2013)	Information about the founder is most important.
Mollick (2014)	Perceived preparedness of the entrepreneur is important.
Xu et al. (2014)	Frequent updates are important.
Frydrych et al. (2014)	Emotional and engaging story is important for success.
Kuppuswamy & Bayus (2014)	Emotions are important.
Davis et al. (2017)	Perceived passion and product creativity are important.
Mollick (2014)	Choice of target is important for success. Project goals that are perceived too high or too low are more likely to fail.
Krishnan et al. (2015)	Fixed funding goal projects are more likely to succeed than variable funding goal projects.
Colombo et al. (2015)	Choice of rewards is important. They should be creative, tangible and fairly priced.
Wheat et al. (2013)	Rewards should have a personal connection to the project or a public acknowledgment of the contribution.

2.3.2.1 Crowd-based Trust Mechanisms

The internet introduces new crowd-based evaluation techniques that can complement or act as a substitute for formal enforcement mechanisms (Ahlers et al 2015). For example, Tucker and Zhang (2011) demonstrate that reporting sales information of a product has important effects on customer choices. Customers prefer to purchase products that have demonstrably been bought by other individuals. Another common strategy of digital marketplaces is to provide customer-based rating and feedback systems. For example, eBay provides buyers the possibility to rate and comment on the service of the vendor. If vendors provide a high-quality experience, they receive a good rating that is visible to all future customers. New customers prefer to buy from vendors that have already proven to be a reliable business partner and are often willing to accept a price premium (Agrawal et al., 2013).

Feedback-based trust mechanisms are effective signals because they fulfil the requirements of Akerlof's (1970) signalling theory. The costs of attaining a positive rating are inversely proportional to the general quality of the vendor. According to Thierer et al. (2015) the use of real-time feedback mechanisms provide a solution to the information asymmetry problem that many regulators considered insurmountable in offline markets. Different authors highlight that crowd-based trust mechanisms can help to increase the volume of transactions between impersonal agents over a wider variety of goods and across geographic, linguistic and cultural barriers (Ahlers et al. 2015, Mollick and Robb 2016).

Crowd-based rating mechanisms are frequently associated with the theory of the "wisdom of the crowd" (Clauss et al., 2018). The theory assumes that the collective estimate of a group of individuals is often better or more precise than that of a single expert (Hertwig, 2012). An explanation for this phenomenon is that by taking the average opinion of a general crowd, the idiosyncratic noise associated with each individual judgment can be reduced (Yi et al., 2012). In crowdfunding, platforms usually display the count of backers that have already contributed to a project. This count is a good indicator for subsequent backers to estimate the crowd's general opinion on the campaign.

In the literature, multiple scientists study how crowd-based mechanisms influence the behaviour of backers. Zhang and Liu (2012) were among the first scientists to find evidence of social herding on Prosper.com, a major lending-based crowdfunding platform. The authors find that backers engage in observational learning, using lending decisions by other peers to infer the borrowers' creditworthiness. A similar behaviour is observed by Colombo et al. (2015) in reward-based crowdfunding. The researchers find that early accumulation of backers encourages additional participation by others. Once a sufficient amount of backers has been achieved, new reinforcement mechanisms such as word-of-mouth and observational learning take over (Colombo et al., 2015). New backers can also benefit from the accumulated feedback, questions and discussions that provide additional information to reduce uncertainty. The reinforcement mechanism in crowdfunding has been documented in multiple studies and the scientific literature agrees that the illustration of the invested amount serves as a reliable signal of quality for many backers (Colombo et al., 2015; Gerber & Hui, 2013; Zhang & Liu, 2012). According to Mollick and Nanda (2016), crowd-based trust mechanisms are one possible explanation why crowdfunding platforms are good at identifying fraudulent campaigns. Moreover, the authors suggest that crowd-based trust mechanisms can lower the incidence of "false negatives", meaning the proportion of good ideas that did not receive the necessary funding because of missing trust.

The crowd's behaviour can serve as a reliable quality signal, however, it does not contribute limitlessly to the success of the crowdfunding campaign. Burtch et al. (2013) observe that the willingness of backers to fund a certain project can change as soon as the funding target is achieved or a sufficient amount of people has demonstrated their commitment to help. Burtch et. al describe this behaviour as the "crowding-out" effect. According to the authors, backers are less likely to participate if a campaign has reached its goal because they experience a decreased marginal utility. In other words, backers have the feeling that their additional contribution is less important to the project and therefore refrain from the interaction. This finding is in accordance with Mollick (2014), who shows that projects usually succeed by narrow margins or fail by large amounts (Frydrych et al., 2014). The explanation for this observation is that backers desire to "make a difference".

The crowd-based trust mechanisms explain how projects can gain momentum after a certain amount of investment was reached. However, they do not explain the decision process of early investors that help creators to reduce the initial level of uncertainty (Colombo et al. 2015). According to Kuppuswamy and Bayus (2013), entrepreneurs in crowdfunding are still dependent on local contacts to initiate a certain momentum of the campaign. In their research, the scientists find that crowdfunding cycles typically follow a “U”-shape, meaning that projects receive most of the funds in the first and last weeks of the funding cycle. The authors also highlight that the first and last contributions are largely coming from family members or close friends of the founder. Also Agrawal et al. (2013) consider family and friends as the major reason for the initial rise in contributions. These findings call into question whether crowdfunding truly gives entrepreneurs access to new, distant investors or whether it is simply a medium to pool an entrepreneur's available network more efficiently.

2.3.2.2 Trustworthy Intermediaries

Another effective approach to signal trustworthiness is to use endorsements of trusted third parties. For example, in job markets, universities can serve as a trustworthy intermediary between employers and potential employees (Spence, 1973). In traditional fundraising investors often rely on articles in magazines, recommendations from industry experts or celebrities to evaluate the trustworthiness of a company (Moritz et al. 2014).

The internet introduced new intermediaries that serve as a signal of quality and can induce trust building. For example, digital marketplaces can use established social network platforms such as Facebook, Twitter or LinkedIn to validate user profiles when moral hazard is a concern (Agrawal et al., 2013). They offer an additional source of information about the founders and thus contribute to more transparency.

According to Mollick (2013), third-party endorsements play an important role in crowdfunding and can significantly contribute to trust building. One of the most effective quality signals is found to be the direct recommendation of the crowdfunding platform operators (Mollick 2014). For example, most crowdfunding platforms provide a section where they display their favourite projects.[17] These projects often receive an additional badge that states: “Projects We Love” (PWL) or “Team’s Favourite”. Research suggests that being mentioned as a favourite project has a significant positive effect on the overall campaign success (Qiu 2013). Third-party endorsements have the greatest, effect when compared to other forms of advertisement (Mollick, 2014; Qiu, 2013). However, the existing research focuses exclusively on the US-market and does not consider the influence of third-party endorsements on project performance in an international context. It remains an open question whether the badge is sufficient to reduce the information asymmetry and help entrepreneurs to attract distant investors.

2.3.2.3 Reputation Signalling

The internet introduced new possibilities for individuals to develop and maintain social relationships (Colombo et al. 2015). Moreover, social platforms such as *Facebook*, *LinkedIn* or *Twitter* disclose many of the previously unobservable relationships, which led to an increased interest in social network theory (SNT) in digital platforms (Borgatti et al., 2014). In crowdfunding research, the study of social relationships enjoys increasing popularity and scientists question to which degree the entrepreneur’s *social capital* influences fundraising success (Colombo et al., 2015; M. Lin & Viswanathan, 2016; Zvilichovsky et al., 2013).

Social capital is defined as “the sum of the actual and potential resources embedded within, available through, and derived from the social contacts of an individual or an organization” (Colombo et al. 2015). In crowdfunding research, scientists usually distinguish additionally between “internal” and “external” social capital (Colombo et al., 2015). External social capital refers to the founder’s contacts outside of the crowdfunding community. This includes the number of family members and friends,

[17] One example: www.kickstarter.com/discover/pwl

which can be estimated through the number of Facebook contacts (Zvilichovsky et al. 2013).[18] Internal social capital, in turn, relates to the relationships developed inside the crowdfunding community. The salient question is whether entrepreneurs rely on their external social capital, or whether crowdfunding promotes the emergence of new relationships within the crowdfunding platform (internal social capital).

Kuppuswamy and Bayus (2013) find in their research that early contributions in crowdfunding are mostly achieved through the entrepreneur's external capital. These findings are in line with the observations of Agrawal et al. (2013), who show that on "SellaBand", a leading crowdfunding platform for musicians, especially the contributions of family and friends are crucial to develop a certain momentum of funding in the initial phase. This finding is confirmed by Indiegogo, a large reward-based crowdfunding platform, stating that entrepreneurs should be able to raise approximately 30% from their private network to attract new backers (Indiegogo 2015).

Colombo et al. (2015), on the contrary, find that the effect of external social capital on early contributions is significantly smaller in magnitude than the effect of internal social capital. The researchers argue that crowdfunding platforms appear to progressively develop into environments rich in social interactions, norms, and behaviours. These communities facilitate the generation and observation of additional information about entrepreneurs and the viability of their initiatives (see also Cheng et al., 2011). Lin et al. (2013) find that individuals that can establish friendships within the crowdfunding community can significantly increase the probability of successful funding.

Zvilichovsky et al. (2013) find that the feeling of belonging to a community can also lead to reciprocal behaviour, meaning that founders return the favour by investing into projects from their own backers. Moreover, the authors find that entrepreneurs that supported other crowdfunding campaigns in the past, show significantly higher success

[18] In many cases crowdfunding platforms are directly linked to popular social networks (e.g. Facebook). This connection allows to study the social ties of crowdfunding participants.

rates, attract more backers and collect more funds than entrepreneurs that did not engage in the community before their campaign. The researchers also find that former project founders tend to invest into other projects at a significantly higher rate. Overall, the researchers conclude that acting on both sides of the market, being a backer and founder, is a rewarding strategy.[19] This finding is in accordance with the Kickstarter statistic that more than 70% of the successful creators were themselves backers of other projects in the past (Kickstarter 2019c).

Colombo et al. (2015) additionally distinguish between two main facets of reciprocity. First, founders that received funding from a community member feel obliged to return the favour by investing into their project. This process is described as *specific reciprocity*. Second, by supporting multiple projects, a backer can become a recognized community member, which in turn promotes *generalized reciprocity*. Generalized reciprocity refers to the situation in which entrepreneurs receive strong support from other unrelated backers because they have demonstrated to be a valuable member to the crowdfunding community in the past. Kickstarter acknowledges this phenomenon and promotes generalized reciprocity by showing the number of backed projects of each user on their individual profile page. Every community member can see whether the person that is asking for funding has been a generous community member in the past. Another method to employ generalized reciprocity is to use a reciprocity clause. For example, Kickstarter introduced the "Kicking-It-Forward" concept in which project creators commit themselves to reinvest 5% of their profits to future crowdfunding campaigns. The demonstration of the commitment to the community is often appreciated by other backers and can improve the likelihood of crowdfunding success (Colombo et al. 2015).

In summary, most research suggests that crowdfunding platforms are communities of like-minded people with strong entrepreneurial spirit. These communities can develop

[19] A similar effect is observed in other social media platforms such as Instagram. Users usually follow their own followers (Teng, Yeh, and Chuang (2015)).

own dynamics (i.e., through specific and generalized reciprocity) in which entrepreneurs can receive support from a more diverse and globally distributed group of unrelated investors.

2.3.3 The Changes Brought by Crowdfunding

The literature considers information asymmetry and the associated lack of trust into a transaction as the main cause for Home Bias (Blum & Goldfarb, 2006; Huberman, 2001; M. Lin & Viswanathan, 2016; McPherson et al., 2001). Amid information asymmetry, different drivers incentivise investors to choose geographically proximate interaction partners. For example, by investing in geographically proximate firms, investors can reduce interaction costs or obtain additional information through shared local contacts (Fried & Hisrich, 1994; Hsu, 2004; Shane & Stuart, 2002; Zane & DeCarolis, 2016). Crowdfunding, however, seems to eliminate some of these drivers that favour the choice of proximate interaction partners. First, crowdfunding equalizes transaction costs because all participants must use the same tools for communication and interaction. These web-based tools are location independent and provide the same possibilities for all crowdfunding participants across national borders (H. Kim & Kim, 2017). Second, crowdfunding facilitates the incorporation of social networks. Investors can easily identify shared contacts and see which projects their friends are backing (Kickstarter Support, 2020). Finally, crowdfunding (and the internet in general) introduces new quality signals, such as crowd-based trust mechanisms and reputation signalling (see section 2.3.2), that can serve as effective measures to alleviate information asymmetry and to build trust into a transaction partner (Thierer et al., 2015). Therefore, investors do not need to rely on heuristic techniques, which are often considered an important driver for cognitive biases. For example, Cook & Parigi (2016) show that the strategic application of web-based trust mechanisms can lead to a reduction of homophily bias on AirBnb, a peer-to-peer online marketplace for accommodations. Similarly, Hortaçsu et al. (2009) find that feedback-based trust mechanisms can reduce prejudice and promote geographical independence on Ebay, a large digital auction platform.

In the crowdfunding literature, different authors examine its potential to reduce investment biases and to “democratize” the access to capital (Dekel et al., 2016; K. Kim & Hann, 2013; Mollick & Robb, 2016). For example, Mollick (2013) compares the amount of female founders that successfully raised capital via crowdfunding to the amount of female founders that received funds from traditional venture capital firms. The author finds that female founders in crowdfunding are present in at least 21% of the successful projects. This number is fifteen times higher than the number of female founders that were successfully backed in a comparable listing by venture-capital firms. Barasinska and Schäfer (2014) address a similar question by examining the data of a German peer-to-peer lending platform. The researchers find that the gender of the project creator does not affect the success probability of the crowdfunding campaign. According to Frydrych et al. (2014), Colombo et al. (2015) and Marom and Sade (2013), women in crowdfunding achieve higher success rates than men. Two different explanations are provided for this observation. First, women are often perceived more trustworthy than men. Second, women often receive proportionally more support from other women that are active on the platform. The current findings on the role of gender suggest optimism that crowdfunding is significantly different from traditional fundraising approaches and might provide new opportunities to previously discriminated groups. Because of this reason, many authors assume that crowdfunding can also alleviate the Home Bias problem in business financing (e.g., Mollick, 2013; Agrawal et al. 2015; Stevenson et al. 2019). However, this assumption is not supported by all authors. Some research provides evidence that Home Bias prevails (e.g., Mendes-Da-Silva et al. 2016; Gallemore et al. 2019; Bade & Walther, 2021). Due to the existing controversary, additional research is required to evaluate the full potential of crowdfunding to alleviate the influence of geography in business finance.

2.4 Home Bias in Crowdfunding

Crowdfunding is a comparably new industry, therefore, research on Home Bias is still in its infancy (Breznitz & Noonan, 2020). Moreover, the existing research is divided on whether geographical distance continues to influence the decision of backers to fund a certain project. Table 2 (p. 65) recapitulates the most important studies on Home Bias in the crowdfunding literature.

Agrawal et al. (2010) were among the first scientists to publish a working paper on the role of geographic distance in crowdfunding. The researchers examine the data of "SellaBand", a former fundraising platform for musicians, and find that the average distance between artists and investors is more than 3,000 miles. The researchers conclude that the geographic distribution of crowdfunding capital is different from that of traditional capital. After controlling for social relationships (friends and family), they find that Home Bias in funding activity is eliminated (Agrawal et al., 2015). The study provides first evidence for a reduced role of geographical proximity in crowdfunding and is often used by advocates as a proof that crowdfunding can reduce distance related frictions. However, this study demonstrates one considerable weakness because it focuses on music, a product that especially in the digital world has only marginal reproduction and transaction costs. Music can be shared and evaluated immediately. Therefore, it should be questioned to which degree the findings of this study are transferable to other crowdfunding models in which entrepreneurs offer more physical products. Surprisingly, the same researchers come to an entirely contradictory result in a consecutive study. After analysing more than 27.000 projects on Kickstarter, a major reward-based crowdfunding platform, they conclude that crowdfunding follows a surprisingly similar geographic pattern as traditional venture capital funding (Agrawal et al., 2013). In the US, Backers seem to prefer physically close entrepreneurs.

Table 2: Literature Overview

This table provides an overview for the research conducted on home bias in crowdfunding. "+++": Strong support for Home Bias, "++": Alleviated Home Bias, "+": No Home Bias.

Author	Focus	Model	Size	Methodology	Home Bias
Agrawal et al. (2013)	US	Reward	27403	Quantitative, Correlation Analysis	+++
Agrawal et al. (2015)	International	Equity	18827	Quantitative, Probability Model	+
Bade & Walther (2021)	DE	Equity	63691	Quantitative, Linear Regression Model	+++
Breznitz & Noonan (2020)	US and CA	Reward	35514	Quantitative, Spacial Error Regression	++
Burtch et al. (2014)	International	Lending	25530	Quantitative, Poisson psuedo-maxmium Likelihood Regression	++
Carbonara (2020)	US	Reward	792	Quantitative, Linear Regression Model	+++
Cumming et al. (2019)	UK	Equity	266	Quantitative, Probability and Negative Binomial Model	++
Dejean (2020)	FR	Reward	14800	Quantitative, Linear Regression Model	+++
Gallemore et al. (2019)	US	Reward	134098	Quantitative, Logistic Regression Model	+++
Günther et al. (2018)	AU/International	Equity	487	Quantitative, Logistic Regression Model	++
Guo et al. (2018)	International/US	Reward	136234	Quantitative, Probability Model	++
Hornuf & Schmitt (2016)	DE	Equity	20460	Quantitative, Probability Model	++
Kang et al. (2017)	CN	Reward	6494	Quantitative, Linear Regression Model	++
Kim & Hann (2013)	US	Reward	9120	Quantitative, Linear Regression Model	++
Kim & Kim (2017)	US	Lending	6059	Quantitative, Logistic Regression Model	+++
Lin & Viswanathan (2016)	US	Lending	4358	Quantitative, Logistic and Linear Model	+++
Mendes-Da-Silva et al. (2016)	BR	Reward	1835	Quantitative, Linear Regression Model	+++
Mollick (2013)	US	Reward	2101	Quantitative, Logistic Regression Model	++
Mollick (2014)	US	Reward	48500	Quantitative, Logistic Regression Model	+++
Niemand et al. (2018)	Europe	Equity	792	Quantitative, Logistic Regression Model	+++
Vigneron (2020)	FR	Reward	4861	Quantitative, Logistic Regression Model	++
Vulkan et al. (2016)	UK	Equity	64831	Quantitative, Probability Model	+
Stevenson et al. (2019)	US	Reward	48500	Multiple Quantitative methods	+
Mollick & Robb (2016)	US	Reward	1536	Quantitative, Gini-Coefficient	+

Mollick (2013) studies the role of geography in crowdfunding and finds evidence that reward-based crowdfunding in the US is less concentrated than conventional venture capital funding. The author concludes that crowdfunding might alleviate some of the geographical frictions. However, Mollick limits their research to projects with a funding volume of below US$5,000 and excludes larger projects. This limitation is a central weakness of the study that calls its validity into question because a significant number of projects on Kickstarter exceeds this low funding volume.[20] In subsequent research, Mollick (2014) finds that the project mix of founders and their success probability echoes the cultural products of the cities in which they are based. Concrete examples are that Nashville shows an increased number of successful projects related to music, Los Angeles to film, San Francisco to technology, games and design products. Mollick (2014), therefore, suggests that geography might play an important role in the success of crowdfunding projects. However, this analysis focuses exclusively on the geography of crowdfunding in the US, meaning that only participants located in the US (backers and entrepreneurs) are included in the sample.

K. Kim and Hann (2013) show that difficult access to capital from local banks in the US encourages entrepreneurs to rely more on reward-based crowdfunding. In this regard, they provide evidence that especially small cities that are remote from large venture capital firms appear to get a disproportionate benefit from this new fundraising tool. Overall, their study provides evidence that web-enabled crowdfunding has the potential to democratize access to capital. According to the researchers, crowdfunding represents a viable option for entrepreneurs that struggle to receive capital from traditional "offline" financiers. The authors observe that especially technology-related ventures tend to be geographically independent. This is different from traditional venture capital investments that are known to be predominantly geographically concentrated (e.g., Silicon Valley). Due to this observation, the authors conclude that reward-based

[20] In the dataset of this thesis, more than 47% of the projects collected more than US$ 5.000.

crowdfunding might become a more viable option especially for technology entrepreneurs located outside the three centres of venture capital activity namely San Francisco, Boston and New York.

Another noteworthy study on the influence of geography is presented by Burtch et al. (2014). The scientists examine the remarkable size of three million individual transactions on Kiva.org to determine the effect of distance and culture on capital allocation decisions. Kiva.org is primarily a donation-based crowdfunding platform which, however, incorporates characteristics of micro-lending. The authors state that on the one hand, Kiva.org can connect individuals from more than 190 countries. On the other hand, however, they show evidence that lenders tend to prefer culturally similar and geographically proximate borrowers. The research of Burtch et al. (2014) is an important contribution to the understanding of the dynamics and peculiarities of the crowdfunding industry, because it is one of only few global studies. However, it must be noted that Kiva.org is a crowdfunding platform that focuses explicitly on charitable social lending. Most borrowers are individuals in need rather than entrepreneurs with innovative business ideas. Therefore, the applicability of these findings to reward-based crowdfunding markets might not be simple or intuitive.

M. Lin and Viswanathan (2016) find that despite the web-based nature of crowdfunding, Home Bias still persist. The researchers provide evidence that lenders on Prosper.com, a major lending-based crowdfunding platform in the US, do prefer to interact with borrowers from the same state. In this context, the authors highlight that a series of statistical tests consistently refuted a purely economic explanation for the Home Bias. The scientists assume that Home Bias occurs mostly due to cognitive biases. Although these findings of are an interesting contribution to the research on Home Bias in crowdfunding, one major limitation is that they focus exclusively on the US. Moreover, lending-based crowdfunding is considerably different from the reward-based model in terms of concept and the backers' motivation. First, it is significantly more regulated, which makes cross-border transactions more difficult (Hornuf & Schmitt, 2016). Second, backers in reward-based crowdfunding frequently support projects because they like the idea or want to support entrepreneurship, rather than to realize profit on their investment (Gerber & Hui, 2013). Therefore, backers in the reward-based model might be affected by different cognitive factors.

Hornuf & Schmitt (2016) study Home Bias in equity-based crowdfunding in Germany. They find that Home Bias continues to matter in online investment decisions even when controlling for family and friends of the founder. However, the expression of the phenomenon differs among different platforms and investor types. The authors suggest that platform design is important for attracting backers more prone to Home Bias. For example, the higher the minimum investment threshold of a platform is, the larger the Home Bias.

Mendes-Da-Silva et al. (2016) study crowdfunding for music related projects in Brazil. The authors find that most pledges are received from within a 50-km radius of the project founders. Moreover, most funds are received from individuals located within a 5km radius. This finding stands in contradiction to the earlier mentioned findings of Agrawal et al. (2015) who found that music projects on SellaBand typically show an average distance of more than 3,000 miles.

Vulkan et al. (2016) examine 64,831 investments on one of the largest equity-based crowdfunding platforms in the UK. They find that although most backers are located in the London area, a considerable geographical dispersion of backers exists across the country. Therefore, the authors conclude that crowdfunding could mitigate the effect that distance has on traditional fundraising efforts.

Guenther et al. (2018) analyse investment flows in equity-based crowdfunding and find that geographic distance is negatively correlated with investment probability for all home country investors in Australia. However, the authors find that overseas investors are not to sensitive to distance. One important limitation of this finding is that the data sample is comparably small. Overall, only 34 overseas investors were included in the study.

Niemand et al. (2018) study the influential factors in 792 equity-crowdfunding projects with a special focus on cross-border investments. They find that participants largely chose projects located in their home country, compared to a geographically distant country or a neighbouring country. The researchers argue that despite the new possibilities of the internet, investors still seem to perceive information gathering as more difficult when it occurs across national borders. Another explanation that the researchers provide is that investors often exhibit a lack of trust towards foreign companies. The "sense of nationalism" or "cultural biases" are considered an explanation for the preference of local projects. This is in accordance with the earlier findings of Burtch et al. (2014) who argue that individual investors are prone to rely on elementary evaluation criteria, such as cultural similarity, in the situation of high-information asymmetry. H. Kim and Kim (2017), who analyse the funding behaviour in equity-crowdfunding on Prosper.com, come to a similar conclusion. They find that despite the potential of crowdfunding to reduce transaction-costs, physical distance remains relevant in web-based fundraising. However, all of the mentioned studies focus on equity-based crowdfunding which is significantly more regulated and restricted than the reward-based model. The findings are, therefore, not necessarily transferable.

Guo et al. (2018) study the dynamics of Home Bias in crowdfunding investment on Kickstarter, a major reward-based crowdfunding site. It is one of the few studies that focuses on reward-based crowdfunding in an international context. Their research suggests that the average distance between investors and founders increases gradually from 3605 km to 4229 km as the funding progresses. Moreover, they find that the effect of Home Bias can vary among different product categories. In which "Food" and "Technology" demonstrate the two extreme examples. The researchers conclude their study with the statement that investors' behaviour demonstrates significant Home Bias. This final statement appears somewhat exaggerated and harsh, given the fact that their estimated average distance is considerably higher than the average distance reported by Sorenson and Stuart (2001) and Cumming and Dai (2010) for traditional venture capital funding (between 70 and 300 miles). Moreover, the finding that certain product categories seem not to be affected by the Home Bias (e.g., technology products) receives insufficient attention. Instead, the scientists highlight that projects related to food and theatre are highly influenced by distance. This finding is somewhat self-explanatory as food and theatre projects can often be associated with “local events” that require physical presence, i.e., opening of a new restaurant or showing a performance at a local theatre (see also Blum and Goldfarb, 2006). Another limitation that the authors admit is that the data sample is highly unbalanced. Most founders and investors are from the US. Therefore, the results might reflect mostly the investment preferences of American backers. A robustness test that excludes the US is missing to validate the findings.

Gallemore et al. (2019) examine the geography of 134,098 crowdfunding campaigns on Indiegogo, a major reward-based crowdfunding platform. The authors find that geographical context mediates the relationship between resources and success. They show that rural areas have lower success rates than urban areas, and affluent areas have the highest success rates. Overall, their findings suggest that Home Bias prevails and does not democratize the access to finance. However, Stevenson et al. (2019) present exactly the opposite finding to Gallemore et al. (2019). The authors study 48,500 projects on Kickstarter and conclude that crowdfunding does augment national and regional funding patterns by re-allocating funding to industries that VCs typically do not fund. The authors highlight that crowdfunding is unlocking new growth opportunities especially for entrepreneurs that are located in underserviced funding regions. However, both studies focus exclusively on projects and investors from the US, which is a major limitation.

According to Breznitz and Noonan (2020), who study data extracted from Kickstarter, crowdfunding alleviates Home Bias but does not fully eliminate it. The authors state that although crowdfunding can expand the geographic reach of fundraising, a project’s crowd is not necessarily global. However, the authors include only two countries into their analysis (US and Canada), which reduces the validity of their findings for a global context.

A recent study by Bade and Walther (2021), however, speaks again in favour for the existence of Home Bias in crowdfunding. The authors examine drivers of investment probability in equity-based crowdfunding using a hand-collected data set of 94 projects that were published on the *Companisto* platform until January 2019. Their findings suggest that investors allocate more attention to campaigns for which they have information advantages, such as local campaigns, due to their limited capacity to process information. Such behaviour may eventually amplify information asymmetry and local preferences. However, the study focuses exclusively on crowdfunding projects from Germany (and German investors), which is a major limitation.

As seen in Table 2 (p. 65) and discussed in this section, the current research is highly divided on whether crowdfunding can affect Home Bias and reduce the influence of geographical distance on backers' investment decisions. This inconsistency of findings justifies the need for additional research and is an important driver for this thesis. Moreover, the direct comparison of the literature in Table 2 reveals three characteristics that appear to be common to crowdfunding research. First, many studies use a quantitative research methodology in which Probability-, Logistic- or Linear Regression models are the most common. Second, many studies use comparably large data samples. Third, most of the research focuses on the US or on individual countries. Only few studies consider Home Bias in an international context.

Although this thesis shares some commonalities with previous research (i.e., quantitative approach and Big Data), it introduces several interesting extensions. For example, by using Negative Binomial regression, this thesis proposes a different quantitative approach to measure the effect of Home Bias, which is based on the observed count of backers and their distance to the entrepreneurs. Moreover, this thesis is the first to examine the potential problems of Big Data samples and how they might lead to misleading findings in crowdfunding research. Furthermore, this thesis builds on the most recent and most international data sample that can be found in the literature to date, addressing the limitations of previous studies that focus only on individual countries.

2.5 Summary on Literature Review

Chapter two summarizes the most important findings in the literature on Home Bias in business finance. First, it explains why Home Bias is to the detriment of both investors and entrepreneurs and how investing only in geographically proximate firms can contribute to a reduced efficiency of the global economy.

The literature review of this chapter shows that many authors consider information asymmetry an important explanation for Home Bias. In this context, the desire to reduce transaction costs, the reliance on social networks (to obtain additional information about the firm), as well as cognitive biases can drive investors to choosing geographically proximate investment targets. Simultaneously, the literature review also shows that several authors consider crowdfunding a potential solution to the Home Bias problem because it can reduce or eliminate some of the described drivers for Home Bias. For example, crowdfunding introduces new possibilities to reduce transaction costs, to connect investors and entrepreneurs more efficiently across geographic, linguistic, and cultural barriers, and to remove some of the former trust issues through new crowd-based trust mechanisms (e.g., real-time feedback).

However, an interesting observation that results from the literature review in this chapter is also that comparably few studies examine the relevance of Home Bias in international crowdfunding empirically. The existing research has divided opinions on whether geographical distance between investors and entrepreneurs is still relevant in digital fundraising. Moreover, this chapter shows that most existing studies on Home Bias in crowdfunding are no longer up to date or show considerable weaknesses. For example, many studies focus on the relevance of distance in individual countries (mostly US) and do not consider crowdfunding in an international context. Moreover, the existing research does not consider new influencing factors (i.e., the Covid-19 pandemic) or was conducted on crowdfunding platforms that do not exist any longer (e.g., SellaBand).

This thesis aims to close the identified research gaps and examine the existence of Home Bias in the emerging industry of international reward-based crowdfunding. The following chapter (Chapter Three) describes the research approach and data sources used in this thesis to achieve this aim.

Chapter Three: Methodology and Data

3.1 Introduction

This chapter explains the research approach, methodology and the underlying data of this thesis in detail.

First, considering different ontological and epistemological dimensions, the philosophical position of the author is described justifying why this thesis follows a pragmatic approach. Additionally, the research methodology is explained in detail providing justifications as to why this thesis can be describes as “nomothetic”, “deductive” and “quantitative”. The chronological process of the Negative Binomial regression model development is also discussed in detail as well as the justification of the model’s suitability for this thesis. Moreover, this chapter provides explanations of the chosen level of analysis, variables, and model architecture. Special attention is also dedicated to discussing potential interpretation problems in Big Data models and how this thesis is addressing these issues. The final section of this chapter portrays the process of data collection and modification.

3.2 Research Paradigms

An ongoing debate in social science concerns the choice of the “right” approach to research (Johnson & Duberley, 2000). Over time, different research methodologies have prevailed. They mostly depend on the researcher’s personal standpoint to truth and reality (Burrell & Morgan, 1979). A central question within this debate concerns how society can assess the relevance and value of research findings that have resulted from different research methodologies (Johnson & Duberley, 2000; McAuley et al., 2014). The understanding of this discussion presupposes the comprehension of Kuhn’s (1996) theory of “paradigms”. Therefore, the following parts of this section introduces and uses Kuhn’s theory as a fundamental concept to compare the existing research methodologies.

The word paradigm derives from the Greek word "*paradeigma"* and can be translated as "*pattern*", "*model*" or "*plan*" (Johnson & Duberley, 2000). Kuhn (1996) suggests that every advanced scientific discipline relies upon a specific paradigm that defines "*what to study (relevance of social phenomena), why to study (…) and how to study (through which methods)*" (Della Porta & Keating, 2008). Therefore, a paradigm describes a certain set of values, beliefs and assumptions that enables scientists to make sense of reality and to distinguish between valid and invalid knowledge (Bird, 2013).

Kuhn (1996) uses the concept of paradigms to describe the development of science. The author argues that scientific progress does not merely occur through the gradual accumulation of knowledge, but rather results from alternating normal and extraordinary phases of discoveries (Bird, 2013). In the normal phase, science resembles a puzzle-solving type of research, where scientists try to answer questions with ever-greater precision that arise within the boundaries of the predominant paradigm. However, during this phase researchers also occasionally discover anomalies that seem to violate their prevailing set of beliefs. A significant accumulation of these anomalies can lead to a crisis within the scientific community. If the crisis is based on sufficient evidence, it will cause a *paradigm-shift*, meaning an entire re-thinking of the former believes and assumptions (Kuhn, 1996). As a famous example for a paradigm-shift, Kuhn uses the transition from the geocentric to the heliocentric model of the universe.

Kuhn's basic concept of paradigms still plays an important role in the contemporary understanding of knowledge and research philosophy (Johnson & Duberley, 2000). However, there seems to be a prevailing disagreement on the current paradigm in the area of social science (Berger & Luckmann, 1967; McAuley et al., 2014; Oberheim et al., 2015). While some scholars argue that social scientists should follow an identical approach to research as natural scientists, others disagree with this assumption by emphasizing that the "object" of observation is an actual human being who can act, think and adapt their behaviour based to the situation (Burrell & Morgan, 1979; McAuley et al., 2014). Therefore, some scholars demand an entirely different approach to research in order to do better justice to the extraordinary circumstances of social sciences (Johnson & Duberley, 2000).

The central problem of the debate is that the arguing scientists are often themselves advocates of entirely different paradigms, meaning they rely upon fundamentally diverse assumptions regarding the nature of "reality" (Burrell & Morgan, 1979). They have different conceptions of truth, address different problems, use different methodologies and often give little recognition to the work of competing paradigms (Oberheim et al., 2015). Kuhn describes this problem as the "*incommensurability*" of paradigms (Bird, 2013; Kuhn, 1996). Accordingly, Johnson and Duberley (2000) emphasize the problem of incommensurability by stating that "*a paradigm cannot be compared or criticized from the standpoint of an alternative paradigm, since the proponents of competing paradigms practice their trades in different worlds (…). Practicing in different worlds, the two groups of scientists see different things when they look from the same point in the same direction*". The incommensurability of paradigms manifests itself in cross-paradigmatic disputes on justified knowledge and the legitimate approach to research. Each of the scientists strives to work towards truthfulness, integrity and authenticity. However, this task proves to be difficult as each scientist has a personal understanding of truth and reality.

The aim of this section is to explain the research paradigm of this thesis. This clarification is necessary, because the thesis makes certain assumptions about how reality is constructed and what constitutes valid research. These assumptions might deviate from other research, where scientists adopt a different paradigm. From a philosophical perspective, different paradigms or conceptions of reality can coexist at any given time and it is difficult to privilege one approach over the other (Burrell & Morgan, 1979; Oberheim et al., 2015). This coexistence, however, can lead to understanding problems when scientists attempt to evaluate, comment, or criticize the work of others. A better understanding of coexisting paradigms helps scientists to be humbler about the validity of their research findings. Moreover, it can lead to a higher appreciation of the work of other scientists.

A common approach in philosophy is to classify the existing paradigms among their "ontological" and "epistemological" orientation (Corbetta, 2003; Della Porta & Keating, 2008). According to McAuley, Duberley & Johnson (2014) any scientific endeavour is underpinned, whether consciously or not, by a positioning along these two dimensions. The following sections 3.2.1. to 3.2.3 provide a general introduction into the terms of ontology and epistemology, describe the scope of competing paradigms, and explain two concrete examples (positivism and postmodernism). Section 3.2.4. builds on the comprehension of the preceding sections and describes the ontological and epistemological position of this thesis.

3.2.1 Ontology

Ontology is the "science of being" and concerns itself with the question of how people understand the nature of reality (Johnson & Duberley, 2000). In general, the academic literature distinguishes between two contrary dimensions: the *objectivist/realist* position and the *subjectivist/nominalist* position (Burrell & Morgan, 1979; McAuley et al., 2014). A person that follows the realist conception believes that it is possible to discover nature's given structures, objects and concepts. For a realist, entities or phenomena such as "organizations" or "taxonomies" do exist in reality and are entirely independent from human perception or cognition. On the contrary, a person that follows the subjectivist conception would describe these entities as a projection or reification of the consciousness that cannot exist independent from the act of knowing (McAuley et al., 2014). In simpler terms, this means that the realist believes in an ultimate, nature given constructs that can be studied and discovered, whereas the subjectivist considers any form of knowledge rather a human attempt to structure and simplify the world.

3.2.2 Epistemology

Epistemology is the "science of knowledge" and concerns itself with the question of what constitutes justified knowledge and how it can be obtained (Johnson & Duberley, 2000). If ontology is "the way people see the world", epistemology is "the way people validate knowledge". Like ontology, the academic literature distinguishes between the *objectivist* and the *subjectivist/idealist* epistemological positions (Burrell & Morgan, 1979; McAuley et al., 2014). Epistemological objectivists are sense driven. They consider something as true as long as they can see, hear, touch, smell or taste it (McAuley et al., 2014). Empirical evidence, metric evaluation criteria and quantitative statistics often play a key role in the research approach and serve as a necessary requirement to validate knowledge (Burrell & Morgan, 1979). Another important characteristic of the objectivist is the strong conviction that the truth is detached from the researcher (Johnson & Duberley, 2000). Values, feelings and beliefs do not and should not have any influence on the truth, because it exists independently from the interest in or awareness of it (Pratt, 1998). Subjectivists, on the contrary, disagree with the idea of an independent truth. They argue that researchers do not discover but rather create new knowledge through the intensive examination of a topic (Johnson & Duberley, 2000). Knowledge is thereby the individual interpretation of the truth by the researcher and can vary from one consciousness to another. Peikoff (1971) states that "*the virtually infallible sign of the subjectivist is his refusal to say, of a statement he accepts: "It is true"; instead, he says: "It is true—for me (or for us)." There is no truth, only truth relative to an individual or a group—truth for me, for you, for him, for her, for us, for them*" (Rand & Binswanger, 1988). Subjectivists argue that complete objectivity does not exist. Moreover, they deny the existence of a neutral observational language which would enable scientists to present knowledge in a neutral and completely objective manner (Berger & Luckmann, 1967; Burr, 2003). According to subjectivists, knowledge is created by the human mind and always influenced, at least to a certain degree, by the learner's values and individual understandings of the world (McAuley et al., 2014). Due to their conception of knowledge, subjectivists are rather interested in narratives than numbers and statistics. They focus on emotions, beliefs, and prior experiences of people in order to understand their worldview and particular situation. Therefore, subjectivists prefer to undertake qualitative research by studying individual cases, feelings and perceptions (Johnson & Duberley, 2000).

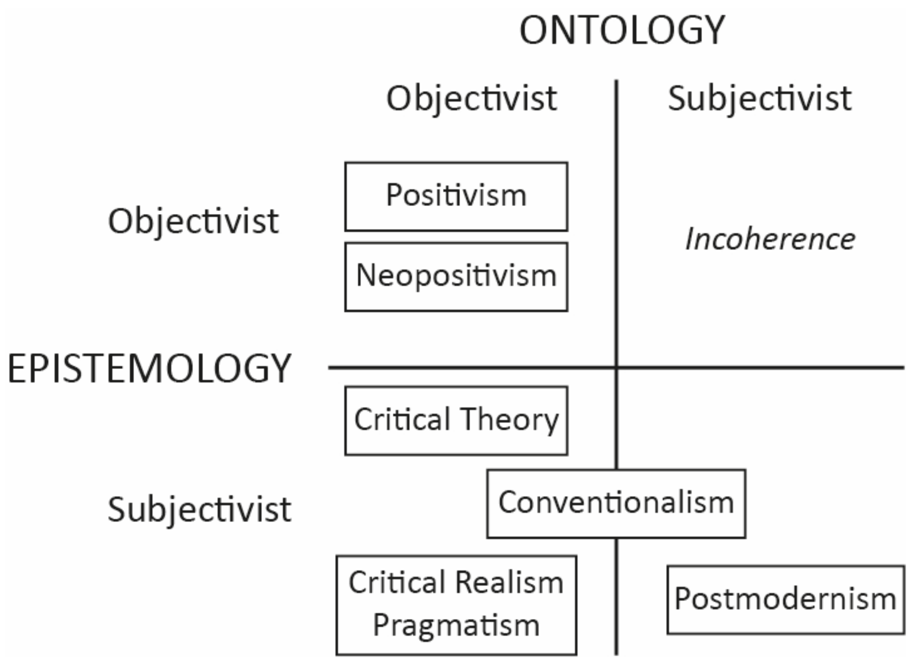
The general problem of epistemology is that it constitutes itself the basis for legitimate research and thereby creates an insurmountable circularity (Johnson & Duberley, 2000). How can the society justify a certain epistemological theory without having any certain approach in advance that could legitimate this action? In other words, it is scientifically not possible to define the basis of justified research without making certain assumptions on what scientific research truly means. Due to the epistemological paradox, there is no particular reason to favour one certain epistemological position over another (Johnson & Duberley, 2000; McAuley et al., 2014). From a philosophical perspective, all positions are equally justified, because there is no valid methodology that would enable scientists to verify or falsify a certain theory (Johnson & Duberley, 2000).

3.2.3 The Paradigm Matrix

The developed framework of Johnson and Duberley (2000) is a helpful tool to understand the various paradigms and to identify the most suitable for oneself (see Figure 4). The authors use the described scales of ontology and epistemology and their dualistic expressions to display the existing approaches and their relationship with each other. However, it is important to highlight that this framework provides merely an orientation of the different perspectives and is not able to represent all the subtle similarities and distinctions that exist between them.

Figure 4: Paradigm matrix

This figure provides an overview of the different research paradigms as suggested by Johnson and Duberley (2000). The different research paradigms are categorized according to their ontological and epistemological orientation. In this context, they can occupy a position between the two opposing dimensions of "objectivist" and "subjectivist". The two extreme positions are positivism (objectivist/objectivist) and postmodernism (subjectivist/subjectivist). The upper right corner of the matrix remains unoccupied since it is difficult to bring an objective epistemology in accordance with a subjective ontology. This assumption is simply incoherent.



To better understand the range of existing paradigms, it is often helpful to take a closer look at the two opposing positions of Positivism and Postmodernism. Positivism can be assigned to an objectivist ontology and epistemology (cp. Figure 4). It is one of the most dominant research approaches in Natural- and Social Sciences (Johnson & Duberley, 2000). According to Johnson and Duberlay (2000), more than twelve different subtypes of positivism can be found today. One of the most famous and dominant variations emerged in the beginning of the 20th century, referred to as *Logical Positivism* (Stadler et al., 2003). The main characteristics of Logical Positivism can be summarized in four commitments (cp. Johnson & Duberley 2000):

(1) The detailed observation of the empirical world is the only foundation for knowledge. At the same time, positivists implicate that such observation can be objective and value-free. (2) Everything that is not observable through the senses and thereby not empirically testable is metaphysical speculation and beyond the realm of real science. (3) Every branch of science should follow the same logical and empirical approach to uncover the truth (unification of research methodologies). (4) The prediction and control of social and natural events constitutes the main objective of science. Therefore, only information that fulfils this requirement represents real knowledge (Stadler et al., 2003).

In the course of time, almost each of the four commitments has been subject to certain forms of criticism. The criticism led to the emergence of new paradigms that define themselves through reference to or objections of the basic implications of Positivism.

One particularly interesting alternative is Postmodernism because it takes a fundamentally opposing position to Positivism (cp. Figure 4). Postmodernists assume a subjectivistic ontology and epistemology. Their central theme is the attempt to prove that objective knowledge does not exist (Johnson & Duberley, 2000). Moreover, postmodernists reject the positivistic assumption that truth is detached from the observer and thereby discoverable through empiric inquiry (Parker, 1992). To underpin their beliefs, postmodernists often refer to the ambiguity of language. They use concepts such as the "*linguistic turn"* (Gergen 1992) as well as Wittgenstein's "*language-games"* (Biletzki et al., 2019) to prove that language can never be neutral and, therefore, it is not capable to convey knowledge in an objective manner. [21] According to postmodernists, the main problem is that scientists often forget their own authorship

[21] Gergen (1992) provides an example for the concept of the linguistic turn by analysing the meaning of the word "logical". The author argues that the understanding of the word can only derive from other words or "*signifiers"* such as "rational", "coherent" or "right-thinking". The reference to other words is required, because there is no neutral vantage point outside the language that would allow a precise and objective understanding McAuley et al. (2014). The central problem is that the search for the true meaning will result in an endless process since every single signifier gains its meaning from other ambiguous signifiers.

of the world and take ideas, concepts and theories for granted that were invented by themselves (Berger & Luckmann, 1967; Chia, 1995).

While some scholars refer to Postmodernism as a “parasitic” approach to knowledge that lives from the criticism of other paradigms and does not contribute to progress (Alvesson, 1995), others consider Postmodernism as an important component of scientific development. Rosenau (1992), for example, argues that postmodernists examine what scientists often take for granted, overlook or understate. By doing this, postmodernists give voice to suppressed ideas and break the “tunnel vision” of traditional approaches. In other words, the paradigm encourages scientists to be more suspicious about the exactness of findings and, thereby, promotes a humbler approach to knowledge claims.

Both presented paradigms (Positivism and Postmodernism) have a decisive influence on this thesis. On the one hand, this thesis strives to pursue an objective research approach, which is oriented towards the common practices of Positivism. On the other hand, influences from Postmodernism become apparent through the fact that traditional approaches (i.e., statistical significance testing of model results) are repeatedly questioned.

3.2.4 Philosophical Position of this Thesis

The awareness of the philosophical position is crucial because it discloses the assumptions and potential biases of the researcher. According to Babbie (2013), two major benefits occur when researchers recognize their philosophical position. First, they tend to better understand the views and actions of others who are operating from a different standpoint. Second, they can benefit from stepping outside their own paradigm at times. By doing this, they can discover new ways of seeing and explaining things. Overall, the realization that different viewpoints or paradigms exist promotes a humbler approach to science and a fairer treatment of each other.

As the underlying philosophical position, this thesis identifies most strongly with *Pragmatism*. On the ontological level, a pragmatist does not see a conflict in accepting both that there is a “real world” and that individuals generate subjective interpretations of this world (Morgan, 2007). Rather than arguing about the subjectivist or objectivist position, a pragmatist regards “truth” as something that provides practical utility, facilitates the lives of individuals or contributes to the progress of society. Rorty (1982) argues that scientists should stop thinking in Greek terms of a distinction between appearance and reality or between mind and body or between intellect and sense. Instead, the author suggests that scientists should take a pragmatic viewpoint and think of themselves as “clever animals” who continuously search for cleverer ways of dealing with the world, but never penetrate beyond the world of senses.

Pragmatists demonstrate a comparably open position towards epistemology (Johnson & Duberley, 2000). While many paradigms tend to favour one specific approach on acquiring new knowledge, pragmatists focus on “what works” (Frey, 2018). Therefore, any approach is legitimate if it produces new knowledge that helps to predict or explain events satisfactorily. Pragmatists highlight that knowledge must be open for revision and adjustments through better ideas or theories, independent of the researcher’s philosophical position (Johnson & Duberley, 2000).

In accordance with Pragmatism, this thesis holds the view that the research on Home Bias in crowdfunding can be handled by different methodologies and is not bound to one specific paradigm. From a philosophical perspective, no valid argument exists that would allow to favour one specific epistemological approach over another (cp. *epistemological circularity* in section 3.2.2). Each methodological approach has claim to validity if it produces knowledge that helps people to make better decisions in practice. For example, an entrepreneur might not be concerned about the origin of the knowledge on Home Bias or how it was acquired as long as this knowledge can be used effectively to develop effective marketing strategies (i.e., identify the most promising backing countries) or predict the potential shipping costs. The attained knowledge should help individuals to anticipate the consequences of manipulating things in the world. For example, in the context of crowdfunding, this could be understanding the effect on the fundraising performance of changing the campaign language or default currency. This thesis holds the view that the purpose of science is not to argue about the different forms of realities or the "right" methodological approach to study Home Bias but rather to derive knowledge that might guide human practice and purposes. In the context of this thesis, the aim of the research is to answer the question whether the Home Bias phenomenon can be found in crowdfunding on Kickstarter and whether entrepreneurs can expect most of the support from geographically proximate backers in practice.

3.3 Research Methodology

Research in social science is multifaceted. One possibility to distinguish the different research methodologies is to classify them among three dualistic expressions: (1) idiographic vs. nomothetic explanation, (2) inductive vs. deductive theory and (3) qualitative vs. quantitative data (Babbie, 2013).

The idiographic and nomothetic explanations address the scope of research. The term "idio" means unique, separate, distinct or peculiar. Therefore, idiographic research focuses on specific examples and seeks to understand the causes of what happened in a particular instance (Babbie, 2013). It deals with subjective phenomena in detail and aims to provide a holistic explanation. In social sciences, idiographic research is common because humans demonstrate particularly complex and distinctive research subjects. For example, their behaviour can be influenced by unique life history or specific character traits (Neuman, 2011).

Nomothetic research, on the other hand, is a more general approach. It covers a wider range of observations and seeks to explain a class of situations or events rather than single one. While idiographic research seeks towards a full explanation for a phenomenon, nomothetic research usually focuses on few explanatory variables and settles for a partial explanation (Babbie, 2013).

This thesis classifies as nomothetic because the focus lies on Home Bias in crowdfunding, more specifically, the influence of distance (between backers and entrepreneurs) on the count of backers that an entrepreneur can expect from a given country. The analysis is conducted across a multitude of independent crowdfunding projects. This focus requires a wide range of observations and a certain degree of abstraction, which can be associated with the nomothetic explanation approach.

In terms of reasoning, research is commonly distinguished in deductive and inductive modes of inquiry (Babbie, 2013). The deductive approach focuses on developing questions based on existing theory (Sondhi, 2011). It analyses whether the predicted behaviour or pattern occurs under specific circumstances. Moreover, it deals with causal relationships between concepts and variables. Deductive research is also referred to as the reasoning from the general to the particular. The inductive method is exactly the opposite research approach (Babbie, 2013). Scientists deal with the observation or data first and seek to find patterns in or develop theories from them. The inductive approach implies the reasoning from the particular to the general (Saunders et al., 2012). Both research approaches are valid in social sciences and complement each other (Babbie, 2013).

This thesis, however, uses a deductive approach. It aims to verify whether the assumptions of the Home Bias theory are appropriate to predict the behaviour of backers in the context of reward-based crowdfunding.

Another common distinction in research methodologies concerns the type of data and how it is analysed. Data can be quantitative or qualitative (Neuman, 2011). As an example, if a scientist says that a person is “intelligent”, they make a qualitative assertion of a person. The key problem of qualitative data is its subjective nature (Punch, 2014). The word “intelligent” provides room for interpretation and its conception can differ among people. For this reason, scientists sometimes attempt to quantify such qualitative assessments. For example, the level of intelligence can also be expressed in the form of the intelligence quotient (IQ) of a person. Converting observations into numerical representations can be useful, as it makes findings more explicit and provides possibilities to aggregate, compare or summarize data. Moreover, quantification of information allows statistical analyses, ranging from simple averages to complex mathematical models (Babbie, 2013).

However, quantifying data also introduces disadvantages. One example is the potential loss in richness of meaning. The process of quantification cannot occur without a certain degree of abstraction (Neuman, 2011). Valuable information can be lost, such as the individual motivation, the driving cognitive processes or conceptions of the acting person.

Qualitative and quantitative data demonstrate unique characteristics. Both are common in social sciences and are equally valid (Babbie, 2013). In general, qualitative research is often considered appropriate for the discovery of new patterns or theories (Neuman, 2011). Therefore, it is frequently associated with inductive research. Quantitative research, on the other hand, is often used to test specific hypotheses or evaluate the influence of certain variables. Quantitative data is therefore more common in deductive research (Punch, 2014). Accordingly, quantitative research fits particularly well with nomothetic explanations, whereas qualitative research seems to align more with idiographic explanations. Although these associations are frequent in practice, they are not absolute. Different combinations of these characterizations are possible to a certain extent (Babbie, 2013).

This thesis uses quantitative data to inspect the influence of distance on crowdfunding backing decisions. Quantitative data is a necessary requirement for econometric or statistical modelling, which demonstrates the main method of inquiry within this thesis.

Econometric models are quantitative methods that have proven to be useful to reveal or explain potential causal relationships among variables that exist in the real world (see Hornuf and Schwienbacher 2018 or Burtch et al. 2014). The common approach in econometric modelling is first to define a specific variable of interest, the dependent or target variable. The dependent variable can take different forms (binomial, discrete or continuous values) and follows a certain type of probability distribution (such as binomial, normal or Poisson distribution). The characteristics of the dependent variable often provide the general conditions for the construction of the model (Beaujean & Grant, 2016). The second step is to choose a set of explanatory or independent variables that are likely to influence the dependent variable. In the third step, scientists commonly use statistical software (such as SPSS or R) to estimate a regression model that best reflects the pattern of the data. In this context, the magnitude and direction (e.g., positive or negative) of each variable's coefficient provides valuable information about the potential relationship. In other words, the individual coefficients reveal how the change of the independent variable will affect the values of the dependent variable. Econometric models help to study the causal relationship between dependent and independent variables and derive valuable insights that are relevant for theory and practice (Katzner, 2017).

While crowdfunding research is still in an initial stage, several publications use econometric modelling as a tool to enhance our understanding of the crowdfunding industry (Burtch et al., 2014; Campennì & Cecconi, 2019; Mollick, 2014). A possible explanation for the tendency to use econometric models is the availability of historic data on crowdfunding projects that is free and accessible on the web. The large amount of data makes statistical modelling particularly interesting for scientists because it provides extensive possibilities to test concrete hypotheses and assumptions.

In accordance with preceded research, this thesis employs an econometric model to study and quantify the influence of distance on investment decisions in reward-based crowdfunding.

3.4 Model Development

This section describes the chronological process of model development. First, different theoretically possible levels of analysis are explained, and their respective advantages and disadvantages are discussed. This discussion provides the justification for the adopted level of analysis. Second, the well-established Gravity Model is introduced as a possible approach to study the influence of distance on investment decisions in crowdfunding. Different arguments are provided as to why the Gravity Model, in its original form, is insufficient for the research purpose and why it requires further adaptations. Based on adapting the Gravity Model, and under consideration of the available data, and their characteristics and with proper justifications, a Negative Binomial regression model is developed demonstrating the most appropriate approach to study the influence of Home Bias on investment decisions on Kickstarter. Since the interpretation of the results in Negative Binomial regression is different from traditional OLS regression (because of the log link function), the results interpretation approach is explained in detail. The section concludes with a detailed description of the data and variables.

3.4.1 Level of Analysis

According to Burtch et al. (2014), crowdfunding research can focus on two possible levels of analysis. On the one hand, it can employ an analysis on the individual-level, which would examine individual backer decisions on which projects to support and how physical distance impacts those decisions. The benefit of this approach is that it allows the researcher to directly model the decisions of individual investors, and perhaps study how those decisions change with investment experience, age, or other characteristics of the backer. Although this is an interesting approach, Burtch et al. (2014) highlight potential issues in terms of measurement and execution that equally apply for this thesis.

Crowdfunding platforms are designed to protect the privacy of their members and therefore usually do not provide information about individual backers to the general public. For example, platforms do not publish exact location data of backers or their individual contribution to a project. The highly limited access to backer-specific information makes the individual-level analysis unfeasible (K. Kim & Hann, 2013; see also Marom & Sade, 2013).

On the other hand, researchers can pursue an aggregate estimation, examining the anonymised volumes of interactions between pairs of countries (see also Hortaçsu et al., 2009). Although this approach is less precise than the approach on the individual-level, it allows the researcher to derive several valuable insights. For example, it allows to study the count of backers from each country and to analyse whether this aggregate varies according to the distance between the entrepreneurs' and backers' home countries. Moreover, the aggregated approach aligns well with prior literature in economics on bilateral trade in both offline (Anderson & van Wincoop, 2004; Guiso et al., 2009; Helpman et al., 2008; Silva & Tenreyro, 2006) and online contexts (Blum & Goldfarb, 2006; Hortacsu, Martínez-Jerez, & Douglas, 2009).

This thesis pursues the second approach and examines the relationship between the aggregated count of backing actions from a specific country to a specific project and the corresponding physical distance between backers and project founders. The aim is to analyse whether a longer geographical distance between the two parties leads to fewer backers, as predicted by the Home Bias theory.

By studying the investment flows on an international level, this thesis makes a valuable contribution to knowledge on crowdfunding, as most of the prior research focuses on the analysis of individual countries (cp. Kim & Hann, 2013; Lin & Viswanathan, 2013; see Mollick, 2013, 2014). This thesis, so far, is the largest and most recent analysis of international investment flows in reward-based crowdfunding to date.

3.4.2 Regression Model

One frequently used model to evaluate the effect of distance in crowdfunding is a modified version of the gravity equation. The gravity equation was first introduced by Tinbergen (1962) as an adaptation of Newton's universal law of gravitation to describe the patterns of bilateral aggregate trade flows between two countries (see Equation 1). The equation is often used in international trade studies and has proven to be stable over time and across different samples of countries and methodologies (Chaney, 2018).

Equation 1: Gravity Model

$$PX_{ij} = \beta_0 Y_i^{\beta_1} Y_j^{\beta_2} D_{ij}^{\beta_3} A_{ij}^{\beta_4} u_{ij}$$

The gravity model predicts the aggregate volume of trade (PX_{ij}) from country *i* to country *j* through the economy volumes of both countries in terms of GDP (Y_i and Y_y), the distance (*D*) between the two countries and other additional factors (A) that facilitate or deter trade such as shared borders or common language, u_{ij} is the error term. This equation is often used to predict trade in global economics (Chaney, 2018) but has also found application in research on crowdfunding (Burtch et al., 2014; M. Lin & Viswanathan, 2016).

While the gravity model enjoys high value in the research on bilateral trade, a major concern is that, in its original form, it is insufficiently profound for the crowdfunding industry. Prior research has uncovered many peculiarities of crowdfunding markets that the gravity model might not account for (Mollick, 2014). For example, crowdfunding markets involve high information asymmetries, which makes them often more sensitive to emotional and social influencing factors (Mollick, 2014; Mollick & Robb, 2016). Moreover, crowdfunding platforms provide new mechanisms of interaction, evaluation and gratification, which are likely to have a strong influence on the overall engagement of backers (Zhang & Liu, 2012). Another important issue is that the required data for the gravity equation is not fully obtainable in the context of this research. For example, the examined crowdfunding platform (Kickstarter) does not allow the researcher to estimate the total volume of investment flows between two countries (i.e., variable PX_{ij} in Eq. 1).

Due to the reasons described above, the Gravity Equation is normally adapted in crowdfunding research to better fit the research context (cp. Burtch et al. 2014, Lin & Viswanathan, 2013). This thesis follows a similar approach by using the formula in Eq. 1 as a starting point and supplementing it with additional variables and transformations to better suit the conditions of this research.

Section 3.4.3 explains in detail the steps and justifications for developing a Negative Binomial regression model to investigate the effect of Home Bias in international reward-based crowdfunding.

3.4.3 Model Development and Justification

The relationship between the dependent and the independent variables can be expressed through different model architectures. One possibility could be to use a linear Ordinary Least Square regression (OLS), which is one of the most frequently used regression models (Montgomery et al., 2015). However, to make predictions and inferences with models based on OLS, certain conditions must be fulfilled (Brooks, 2008). First, OLS assumes a linear relationship between the mean response of the dependent variable (Y) and the explanatory variables (X). Second, the errors are assumed to be independent, meaning that there is no connection between how far any two points lie from the regression line. Third, the response variable is assumed to be normally distributed at each level of X and fourth, the variance or standard deviation of the responses is assumed to be consistently equal for all levels of X (Brooks, 2008; Montgomery et al., 2015). The assumption of equal variance is also often referred to as the assumption of "homoscedasticity" (Brooks, 2008). In practice, it is not unusual that especially the condition of homoscedasticity is not met (Cameron & Trivedi, 2010).

This thesis employs the count of backers from individual countries as the dependent variable. This type of variable is also referred to as *count variable* or *count data* (Cameron & Trivedi, 2010).

Count variables share three specific properties: (1) their values are always whole numbers (e.g. 1,2,3); (2) they have a lower bound at zero, meaning that the values do not become negative; and (3) count variables frequently demonstrate a shape that is right skewed, with most values being low and relatively few values being high (Beaujean & Grant, 2016; Cameron & Trivedi, 2010). The properties of count variables violate some of the described assumptions for typical OLS regression. One key problem is that neither the dependent variable nor the residuals are normally distributed, as assumed by OLS. Moreover, the residual variance often increases in analogy to the predictor variables, which violates the important assumption of homoscedasticity in OLS (Beaujean & Grant, 2016). Overall, using typical OLS regression with count variables can lead to wrong estimates of standard errors and confidence intervals, which diminishes the validity of the derived findings (Cameron & Trivedi, 2010).

Prior research commonly dealt with the described problems of OLS either by assuming that typical linear models are robust enough to handle any assumption violations caused by count variables, or by transforming variables to make them fit more traditional models (Beaujean & Grant, 2016). However, both approaches are problematic. Although in some cases OLS regression models can estimate unbiased regression coefficients for non-normally distributed data, they tend to produce inflated standard errors (Beaujean & Grant, 2016; see Box, 1953; Cochran, 1947; Lix et al., 1996). Additionally, OLS makes continuous predictions that also can become negative. Count data, however, is discrete and has a lower bound of zero. The consequence of this mismatch is that residuals tend to be heteroscedastic, which affects the validity of findings (Cameron & Trivedi, 2010).[22,23] A frequently used solution to account for heteroscedasticity in OLS is to log-transform the dependent or both the dependent and independent variables (Beaujean & Grant, 2016; Cameron & Trivedi, 2010). However, such transformation would require taking the log of zero values, which results in undefined values and the loss of valuable data. Moreover, the issue that the model continues to predict negative values remains unresolved. Another potential solutions to the heteroscedasticity problem is the *Box-Cox* power transformation (Sakia, 1992).[24] However, one important

---

[22] Heteroscedasticity is the opposite state of homoscedasticity, meaning that the variance of the errors is not constant.

[23] In addition to traditional OLS regression, Cameron and Trivedi (2010) suggest the use of the more robust weighted least squares (WLS) regression model. This type of model has been equally considered in this thesis but discarded in favour of a better solution that is specifically optimized for count data (see Eq. 9).

[24] The Box-Cox transformation is a mathematical tool used in regression analysis and time series analysis to achieve stabilisation of variance.

issue when dealing with nonlinear transformations is that they usually come at a cost of having a more difficult model to interpret, which is the reason why using OLS has been ruled out for this thesis. [25]

Instead of modifying the data to fit the traditional OLS regression, this thesis adapts the regression model to the specific requirements of count data. The Generalized Linear Model (GLM) is a regression framework specifically designed to handle non-normally distributed variable types (McCullagh & Nelder, 2018). GLMs consist of two components: (1) the specification of residuals' distribution; and (2) a function to link the model predictions and the linear combination of the independent variables (Beaujean & Grant, 2016; Cameron & Trivedi, 2010).[26] GLMs offer different types of models to deal with count variables. They mostly depend on the specific distribution of the dependent variable. The Poisson regression model is a GLM that is usually used to model count data (Beaujean & Grant, 2016; Cameron & Trivedi, 2010). It has two characteristics: first, it assumes that the dependent variable and the residuals follow a Poisson distribution, $y \sim Poisson(\lambda)$, where λ describes both the mean and the variance of the distribution (the Poisson probability density function is provided in Equation 2).[27] Second, the linear combination of the independent variables is linked to the dependent variable via a natural log transformation, which is similar to the approach in the common logistic regression (Beaujean & Grant, 2016; Coxe et al., 2009) as can be seen in Equation 3.

[25] All described solutions have been tried on the data sample. However, none of the approaches could resolve the problem of heteroscedasticity in a sufficient manner.

[26] A typical OLS regression can be achieved through GLM regression by using a normal distribution for the residuals and an identity link function for the dependent and independent variables (i.e. multiply the regression by one).

[27] The Poisson distribution is designed for non-negative integers. If λ is close to zero, the distribution is right skewed. For increasing λ, the distribution becomes less skewed and appears closer to a normal distribution.

Equation 2: Poisson Probability Density Function

$$P(y) = \frac{e^{-\lambda}\lambda^{y}}{y!}\ for\ y = 0,1,2,3 \ldots$$

Equation 3: Poisson Regression Model

$$\log(\lambda) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

One specific problem of the Poisson model, however, is that it assumes equidispersion (Coxe et al. 2009), which is the equality of the conditional variance and mean (Payne et al., 2018). In practice, this condition is often not met and the data is usually described as "overdispersed", meaning that the conditional variance exceeds the conditional mean (Hilbe, 2011). Allison (2009) refers to overdispersion as the condition in which there is more variation in the event counts than would be expected based on a Poisson distribution. Therefore, the application of a Poisson regression model on overdispersed data can result in incorrect small standard errors and p-values of the model coefficients (W. Greene, 2008; Payne et al., 2018). This in turn can lead to erroneous and overly optimistic conclusions of statistical significance of regressors (Payne et al., 2018).

To evaluate the suitability of Poisson regression for this thesis, the degree of overdispersion was calculated for the sample data. The presence of over- or underdispersion can be estimated by dividing the Pearson Chi^2 dispersion statistic by the residuals' degrees of freedom (McCullagh & Nelder, 2018; Payne et al., 2018).[28] This approach is illustrated in Equation 4.

[28] Underdispersion is the opposite case in which the variance is smaller than the mean. However, this is rarely the case in practice and overdispersion is considerably more common.

Equation 4: Formula for the Estimation of Dispersion

$$\frac{\chi^2}{Degree\ of\ freedom\ of\ residuals} = Estimation\ of\ Dispersion$$

Equation 5: Estimation of Dispersion for the data sample

$$\frac{1.10\mathrm{e}^7}{1{,}118{,}624} = 9.83$$

A ratio close or equal to 1, indicates that the data is drawn from the Poisson distribution, for which the variance is equal to the mean. However, a value smaller than one implies underdispersion, while a value significantly greater than 1 indicates overdispersion.[29] By calculating the model as a Poisson, the degree of dispersion is estimated. The obtained value of 9.83 indicates the presence of overdispersion in the data sample and suggests that the Poisson regression model is not an adequate choice for the data sample.

One way to incorporate overdispersion in count variables into GLM regression is to use the Negative Binomial distribution for the residuals (Cameron & Trivedi, 2013; W. Greene, 2008; Hilbe, 2011; Payne et al., 2018). The Negative Binomial regression model is a generalization of the Poisson distribution that does not impose equidispersion. Two different forms of the Negative Binomial regression can be found in the literature, but the most common implementation is the so called "NB2" model (Cameron & Trivedi, 2013; Date, 2019). The NB2 model uses an additional dispersion parameter α, that specifies the level that the distribution's variance exceeds its mean (λ).[30] This thesis,

[29] It is important to note that this approach is an approximation and that there is no fixed threshold for an affirmative statistical intervention (Hilbe, 2011).
[30] As the dispersion parameter goes to zero, the variance becomes equal to the mean. In this case, the Negative Binomial regression estimates equal the Poisson estimates (see Cameron and Trivedi (2013)).

therefore, follows the common practice and uses the NB2 model for estimations. Equation 6 provides the variance function.[31]

Equation 6: Dispersion Parameter

$$E(y) = \lambda$$

$$Var(y) = \lambda + \alpha\lambda^2$$

The NB2 model requires the definition of the parameter α, which is used to express the variance in terms of the mean. Cameron and Trivedi (2013) suggest a technique that they call "auxiliary OLS regression". This method can be used to estimate the parameter for α. However, some scientific software (i.e., *R*'s *Mass* library) calculates the dispersion parameter automatically and does not require any additional effort from the researcher.

The Negative Binomial distribution can arise as a gamma mixture of Poisson distributions. One common parametrization of its probability density function is provided in Equation 7, with the mean or expected value of the distribution λ, the shape parameter θ and the gamma function Γ(…).

Equation 7: Negative Binomial Probability Density Function

$$P(y) = \frac{\Gamma(\mathrm{y} + \theta)}{\Gamma(\theta) * y!} * \frac{\lambda^y * \theta^\theta}{(\lambda + \theta)^{\theta + y}}$$

[31] The other Negative Binomial regression model that can be found in the literature ("NB1") differs only in one parameter. Instead of using the squared mean, NB1 uses the simple mean: $\lambda + \alpha\lambda$.

Different versions of this formula can be found in related literature (cp. Cameron & Trivedi, 2013 with ; UCLA, 2019b), which may lead to confusion. In this context, it is important to understand that the relationship between α and θ is α = 1/θ.[32] For a detailed discussion on the Negative Binomial regression and its varieties see W. Greene (2008), Hilbe (2011) and Cameron and Trivedi (2013).

The log-link function allows the linear combination in parameters. In order to model the expected value of y, where $E(y) = \lambda$, a Negative Binomial regression model can be constructed of the form presented in Equation 8.

Equation 8: Regression Function

$$log(\lambda) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

Where X_p describes the explanatory variables and β are the fixed, unknown parameters that can be estimated via maximum likelihood estimation (MLE); β_0 denotes the intercept and β_p the different slopes for the explanatory variables.

Based on the explanations provided above, a Negative Binomial regression model is employed in this thesis to examine the influence of distance on the count of backers in crowdfunding projects from a given country as described in Equation 9.

Equation 9: Final Model Equation

$$Log\left(NumB_{ij}\right) = \beta_0 + \beta_1 D_{ij} + \beta_2 GDP_i + \beta_3 GDP_j + \beta_4 C_k + \ \beta_5 PWL + \ \beta_6 LP + \ \beta_6 P$$

[32] This information is important for conducting estimations because the input parameters may differ within different statistical software programs (i.e., Python's *Statsmodels* library and R's *MASS* package).

$Log(NumB_{ij})$ is the aggregated count of backing actions from backers in country *i* to entrepreneurs in country *j*. Distance (D_{ij}) is the key variable of interest and denotes the estimated distance between the backer's and entrepreneur's home country capitals of *i* and *j* respectively. In addition to distance, several control variables are included that are likely to impact the count of backers.

First, the model controls for the respective wealth of both the backer's and entrepreneur's home countries. For this purpose, this thesis uses the GDP per capita of both countries (GDP_i *and* GDP_j). To control for possible variations across different project types (i.e., "technology", "games"), the model includes dummy coded project categories (C_k) to which each project has been assigned. The dummy variable "PWL" is used to specify whether a project has been endorsed by the crowdfunding platform (taking values of "true" or "false" respectively). To estimate the potential effect of herding behaviour, the model uses a dummy variable ("LP") that specifies if a project attracted a particularly large count of backers. To study the effect of the Covid-19 pandemic, the model uses a dummy variable (P) that signals whether a project was launched during the pandemic or not. Finally, the model controls for unobserved heterogeneity in time by including the project's launch year (dummy coded). Table 3 provides a brief overview of all variables, their definitions, and sources. The detailed descriptions and justifications for the chosen variables are presented in section 3.4.4.

Table 3: List and definitions of all constructed variables

This table lists and defines all variables in Equation 9 including the type and unit of measure.

Definition	Description
Count of Backers (*NumB*)	The dependent variable is the observed aggregated count of backing actions from one country to a specific project. *Source:* Kickstarter website.
Distance (*D*)	Distance in 1000 km between the entrepreneur's country capital and the backers' country capital (key variable of interest). *Source*: Constructed variable that is based on the geographical information on backers' and entrepreneurs' location obtained from Kickstarter website.
Project Category (*C*)	Dummy variable for 15 different project categories: Technology, games, design, photography, fashion, comics, journalism, crafts, art, publishing, film & video, dance, music, food and theatre. *Source:* Kickstarter website.
"Projects We Love" (*PWL*)	Dummy variable for all projects that carry the badge "Projects We Love". *Source* Kickstarter Website.
GDP per capita Backer (GDP_i)	The GDP per capita value in USD for the backer's home country. *Source:* World Bank WDI Database.
GDP per capita Entrepreneur (GDP_j)	The GDP per capita value in USD for the entrepreneur's home country. *Source:* World Bank WDI Database.
Large Projects (*LP*)	Constructed dummy variable that specifies if a project attracted a particularly large number of total backers (>107). *Source:* Constructed variable based on the project data extracted from Kickstarter.
Covid-19 Pandemic (*P*)	Dummy variable for projects that were launched during the Covid-19 pandemic. *Source*: Constructed variable based on the project launch dates extracted from Kickstarter.

3.4.4 Model Variables

3.4.4.1 Count of Backers (Dependent Variable)

The dependent variable is the observed aggregated count of backing actions between pairs of countries over the period from April 2009 to June 2020. The construction of the dependent variable is similar to that of Burtch et al. (2014), who use the aggregated count of lending actions between two countries to study the influence of distance in social lending. This thesis obtains the aggregated count data from Kickstarter, one of the largest business-related crowdfunding platforms. This is, so far, the first study to use the aggregated count of backing actions between country pairs in reward-based crowdfunding as a dependent variable to study the effect of Home Bias.[33]

The variable is used in this thesis for two reasons. First, the estimated count of backers is an important guideline for entrepreneurs who need to make campaign set-up decisions (i.e., necessary translations or the currency of the project). Knowing where most backers may come from allows entrepreneurs to focus on the relevant countries, design more efficient marketing activities and better estimate the resulting shipping costs. The assessment of backers' origin is considered one of the biggest challenges for entrepreneurs engaging in reward-based crowdfunding (cp. Kickstarter, 2017). By choosing the aggregated count of backers as the dependent variable, this thesis can provide important practical guidelines for entrepreneurs on how to improve the efficiency of their campaigns. Second, the aggregated count of backers is also used as the dependent variable because Kickstarter does not allow to reproduce the exact funding volume that a project received from a specific country, which is the common approach in the gravity equation model (Chaney, 2018; Burtch et al., 2014b; Lin & Viswanathan, 2013). However, the aggregated count of backers provides a good

[33] Prior studies used pledge results (Guo et al., 2018), probability of transaction (Lin & Viswanathan, 2013, Agrawal et al., 2011), or funding success (Mollick, 2013) as the dependent variable.

alternative because it similarly assesses the degree of interrelation between two countries. [34]

Choosing the aggregated count of backers as the dependent variable introduces one potential bias. The literature review in section 2.4 (p. 64) discussed the important role of family and friends for the success of crowdfunding projects (see e.g. Agrawal et al., 2013; Kuppuswamy & Bayus, 2013). The problem is that family and friends typically do not make their decisions objectively but are usually influenced by the personal relationship to the entrepreneur. Moreover, it can be assumed that most friends and family members share the same country of residency as the entrepreneur. The result is that crowdfunding projects might be inclined to receive comparably more backers from the same country of the project founder. This tendency might distort the results of the model in favour of the existence of Home Bias.

One approach to handle this potential bias is to subtract the count of backers that are related to the entrepreneur from the overall count of backers that support the project from the entrepreneur’s home country. Unfortunately, the limited data from Kickstarter does not allow to identify these specific “home backers” with a sufficient accuracy. To take the influence of related backers into consideration nevertheless, this thesis uses a threshold of 150 backers. This means that the number of home backers of a specific project must exceed 150 or is otherwise treated as an observation of zero. Accordingly, an observation of 160 backers is treated as 10 backers (160-150 = 10). This threshold is based on Dunbar’s (1992; 1993) findings that individuals are usually unable to maintain a stable social relationship with more than 150 different people at the same time. The author suggests that this limit is imposed by neocortical processing capacity and is, therefore, genetically determined. Dunbar’s (1992; 1993) theory has been confirmed in

[34] No justification could be found why the aggregated funding volume should demonstrate a better approach than the aggregated count of interactions, that is used in this thesis. Therefore, both approaches are considered of equal value.

multiple studies and appears to persist despite the rise of online social networks (Arnaboldi et al., 2015; Zhao et al., 2014).

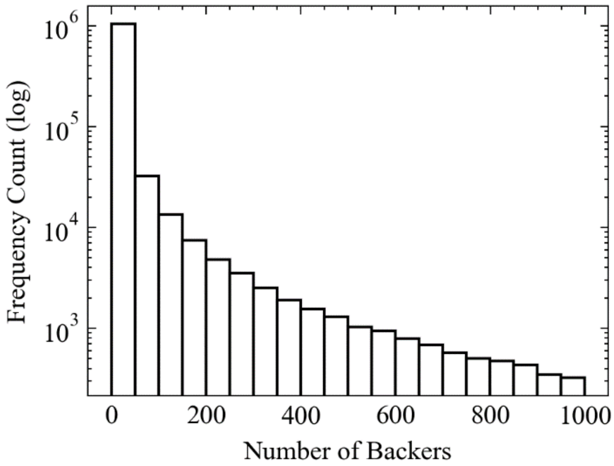
This thesis uses the comparably high threshold of 150 for the number of home backers to ensure that the results are not distorted by the direct social links of the entrepreneur. This conservative position means that if a significant relationship can be observed between the count of backers and their distance form entrepreneurs, despite the threshold, the results are more likely to be true. This thesis is the first to use a threshold approach to counterbalance the influence of friends and family members on crowdfunding performance.[35]

The estimated count of backers is a typical count variable because it only takes discrete values, has a lower bound of zero and no upper bound. Figure 5 shows a section of the overall distribution of the dependent variable (for illustration purposes the diagram is limited to 1000 backers). The observable right-skewed shape is another typical characteristic of count variables.

[35] Some of the prior research estimated the influence of friends and family members by studying the social networks of the founders (cp., Agrawal et al., 2013, Kuppuswamy & Bayus, 2013). However, this approach is error-prone because it bases on many assumptions and can no longer be implemented due to the privacy policy of platform operators that forbids the web-scraping of personal data (cp. Kickstarter 2018).

Figure 5: Distribution of the Dependent Variable

This figure shows a section of the overall distribution of the dependent variable in the form of a histogram. The dependent variable is the aggregated count of backing actions from one country that invested into a specific project. The x-axis shows the different discrete values of the count variable, while the y-axis shows its frequency in the data sample on logarithmic scale. The right skewed distribution is a common characteristic of count variables.



3.4.4.2 Distance

The aim of this thesis is to study the influence of geographical distance on the count of backing activities and whether the phenomenon of Home Bias is present in reward-based crowdfunding. The distance between entrepreneurs and backers is the key explanatory variable of interest. Calculating such distance requires the availability of location data for both backers and entrepreneurs of crowdfunding projects. Unfortunately, this information is difficult to obtain as crowdfunding platforms typically do not publish exact location data of backers because of privacy reasons. Kickstarter, however, is the only reward-based crowdfunding platform that provides information on the entrepreneur's and backer's country of origin in the project description section. This information is accordingly used as an approximation for their location. In analogy to Burtch et al. (2014), the distance is therefore calculated using the geographical coordinates of the country capitals for each country-pair. Pursuing this approach, the first step is to identify the latitude and longitude coordinates for each country's capital. This can be achieved through the Google Maps API, which is a digital service offered by

Google that allows researchers to make requests via the Hypertext Transfer Protocol (HTTP) to determine the specific coordinates of cities. [36] Google's service for geographical data enjoys wide popularity and has been used by multiple researchers in the past (cp. Guenther et al., 2018; Guo et al., 2018). This thesis uses the Google Maps API to obtain the coordinates of each country's capital that is present in the data sample. The full list of city coordinates can be inspected in Appendix A.

In a second step, the geodesic distance between the country capitals is calculated. The geodesic distance is the shortest distance on the surface of an ellipsoidal model of the earth (Karney, 2013). A drawback of using country capitals as a proxy for geographical location is that in the scenario where backers are from the same country as the project creator, the calculated distance is zero. In other words, the reported distance tends to understate the actual distance. This characteristic needs to be considered in the interpretation of the results because it favours the Home Bias phenomenon. Although the approach taken in this thesis has some weaknesses, Burtch et al. (2014) suggest it as the most suitable method to study Home Bias in crowdfunding, given the data limitations. The difficulty and laboriousness of collecting distance data in crowdfunding research is highlighted by many scientists in the literature and is considered a major reason why few scientific studies focus on this subject (cp. Guo et al., 2018; Burtch et al., 2014).

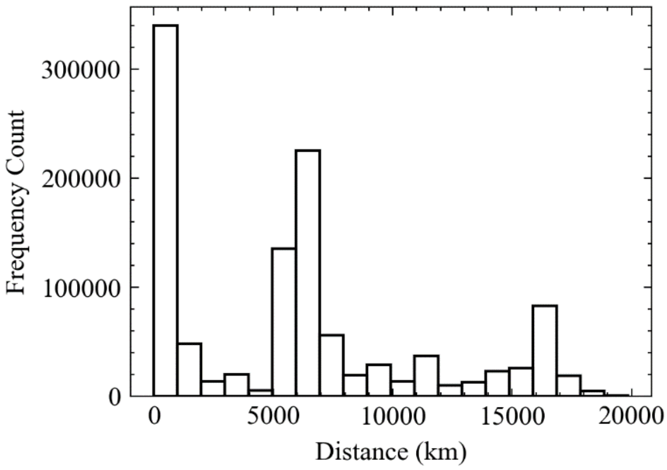
Figure 6 shows the distribution of the distance variable in the form of a histogram for the obtained data sample over the period of April 2009 to June 2020. A striking pattern is that most observations can be found at the zero level. This might be an early indicator that backers do prefer to invest in their home countries.

---

[36] The documentation for the geocoding API can be found on: https://developers.google.com/maps/documentation/geocoding/

Figure 6: Distribution of the Distance Variable in the data sample

This figure shows a histogram for the distance variable. The x-axis shows the distance between backers and entrepreneurs in km whereas the y-axis denotes the respective frequency in the data sample.



3.4.4.3 GDP per Capita

Burtch et al. (2014) find a significant effect of GDP on backer behaviour in their analysis of crowdfunding lending decisions. Also, other research in crowdfunding identifies GDP as a relevant explanatory variable (M. Lin & Viswanathan, 2016). Consequently, this thesis controls for the GDP per capita of the backers' and entrepreneurs' home countries. The GDP per capita is different from the GDP as it is the total output of a given country divided by its population (UNICEF, 2019). The GDP per capita is chosen over the traditional GDP because it has a slightly different focus. Instead of focusing on the total production output of a country, the GDP per capita sets the value of all produced goods and services of a country in relation to its population and thereby provides a better estimate of individual wealth. The GDP per capita is commonly used as a broad measure of average living standards or economic well-being (OECD, 2009). This thesis is the first to use the GDP per capita as an explanatory variable to predict crowdfunding performance.

The general assumption of this thesis is that backers in countries with high GDP per capita are more likely to participate in crowdfunding due to their surplus in financial resources. Simultaneously, projects in countries with low GDP per capita might tend to attract more funds. Indications for this behaviour are explained by Gerber and Hui (2013) who show that backers are often more likely to back a project if they feel that the funds are going to be used for a good cause. Also, Burtch et al. (2014) find that lenders on the donation-based crowdfunding platform "Kiva" are more inclined to provide funds to borrowers who are in countries that are comparably less wealthy. Entrepreneurs from developing countries or people in need seem to enjoy an advantage in the attraction of funds (Burtch et al., 2014). In this context, different scholars speak of the increasing tendency of investors to fund projects or companies not only for the financial return but also for the public good (Caseau & Grolleau, 2020; Kish & Fairbairn, 2018). This trend is referred to as "Impact Investing" (Kish & Fairbairn, 2018).

According to the Global Impact Investing Network (GIIN), Impact Investing describes investments that are made with the intention to generate positive, measurable social and environmental impact alongside a financial return (GIIN, 2019). The Impact Investing movement has grown exponentially in recent years (GIIN, 2019; Kish & Fairbairn, 2018) and is, therefore, likely to also affect the crowdfunding industry. One possible pattern could be that, similar to Kiva.org, funding on Kickstarter flows from high individual-wealth to low individual-wealth countries.

To estimate the effect of GDP per capita on crowdfunding behaviour, the dataset is extended by the GDP per capita information for both entrepreneurs' and backers' home countries. The most recent GDP per capita values are extracted from the World Bank's *World Development Indicators* (WDI) database. The database is compiled from officially recognized international sources and presents the most current and accurate global development data available. It is the typical source for country specific estimates on economic performance and has been used by numerous researchers in the past (Burtch et al., 2014; Mendes-Da-Silva et al., 2016).

Figure 7 provides a histogram for the GDP per capita distribution of the backer countries in the data sample (GDP-B). The figure shows that most backers are from countries that have a GDP per capita value of around US$40,000-60,000. The high imbalance of backer countries introduces a potential bias. The results of the model estimation might disproportionally reflect the behaviour from these few countries and superimpose the behaviour of the underrepresented backer countries. To control for this bias, an additional robustness test is implemented which estimates the same model in Equation 9 on a sample that excludes the most frequent countries (see section 4.4).

Figure 7: Histogram of the GDP per capita Distribution for Backers

The figure shows the distribution of the GDP per capita for the backers' countries. The x-axis shows the different GDP per capita values, whereas the y-axis shows the frequency of these values.

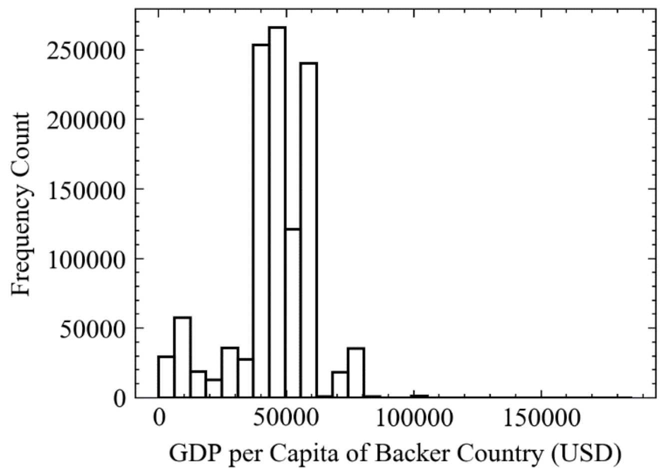
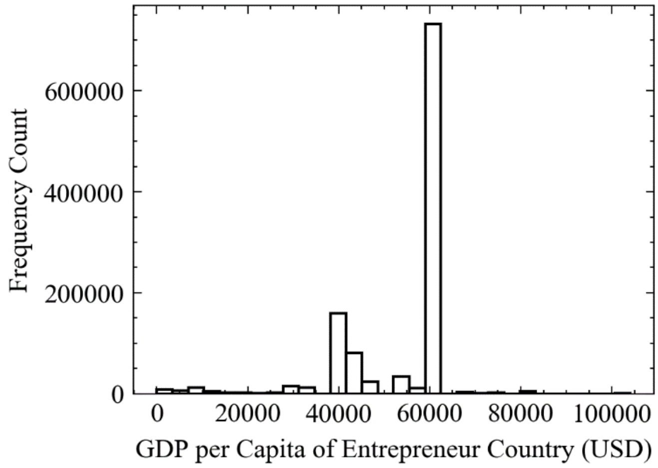


Figure 8 provides a histogram for the GDP per capita distribution of the entrepreneur countries (GDP-E). This figure shows a similar imbalance in the data sample. The high count of GDP-E at the level of US$ 60,000 can be attributed to the high number of US based projects. The observation that most of the projects in the sample are based in the US is not surprising as Kickstarter is a US-based platform. Moreover, the US demonstrates the oldest market for reward-based crowdfunding. Nevertheless, this characteristic introduces a potential bias and must be considered in the estimations. To control for this bias, a robustness test is implemented which estimates the same model in Equation 9 on a sample that excludes the most frequent countries (see section 4.4).

Figure 8: Histogram of the GDP per capita Distribution for Entrepreneurs

The figure shows the distribution of the GDP per capita for the entrepreneurs' countries. The x-axis shows the different GDP per capita values, whereas the y-axis shows on a logarithmic scale the frequency of these values. The figure shows that most entrepreneurs are from countries that have a GDP per capita value of around US$60.000. A possible explanation for this characteristic is the high number of US entrepreneurs in the dataset. The figure also suggests that countries with low GDP per capita tend to use crowdfunding more frequently than countries with high GDP per capita.



3.4.4.4 Project Category

The literature review in section 2.4 revealed that crowdfunding performance can vary significantly across different types of projects (K. Kim & Hann, 2013; Mollick, 2014). For example, K. Kim and Hann (2013) find that technology-related ventures tend to perform exceptionally well in crowdfunding. To account for the respective differences, project category is included as a dummy variable in the model. Dummy variables can only take the two values of *true* or *false*.

The data for the variable is obtained from Kickstarter which distinguishes between 15 different project categories: *Technology, Games, Design, Photography, Fashion, Comics, Journalism, Crafts, Art, Publishing, Film & Video, Dance, Music, Food* and *Theatre*. The variable is directly extracted from the Kickstarter platform and does not require any further modification than the encoding into a dummy variable.

The overall distribution of projects categories can be inspected in Table 4. The category of *Games* is the most frequent category in the sample, accounting for 19% of all projects. Projects in the *Journalism* category are the least common, accounting for less than 1% of all projects.

Table 4: Observation Frequencies for Product Categories

This table shows the distribution of the 1,118,654 observations (generated from 211,695 crowdfunding projects) by category in a descending order for the entire data sample. The category of "Games" demonstrates the most frequent type of projects, whereas "Journalism" is the least common type of projects.

Category	Observations	% of total
Games	215,744	19%
Film & video	130,960	12%
Design	125,520	11%
Technology	108,452	10%
Music	104,559	9%
Publishing	100,344	9%
Art	77,449	7%
Comics	77,227	7%
Fashion	66,497	6%
Food	37,041	3%
Photography	24,587	2%
Theatre	1,7639	2%
Crafts	15,513	1%
Dance	8,742	1%
Journalism	8,380	1%
Total	1,118,654	100.00%

3.4.4.5 PWL Badge (Third-Party Endorsements)

The “Projects We Love” (PWL) badge is rewarded by Kickstarter employees for interesting and creative projects and has established itself as a signal of quality (Kickstarter, 2019a). Moreover, PWL projects are normally promoted in the weekly newsletters and search categories of the website, increasing the likelihood of being seen. According to Kickstarter (2019a), projects that strive to receive the badge need to provide a detailed project description that includes a thorough plan for completing the endeavour, captivating images or videos, an excited community, and a high degree of creativity.[37]

The important role of this kind of third-party endorsements has been documented by prior literature in both online and offline markets (Mollick, 2013; Mollick & Robb, 2016; Spence, 1973; Stiglitz, 1974). Some research confirms the high relevance of the PWL badge on crowdfunding performance (Mollick, 2014; Qiu, 2013). However, this research focuses exclusively on the US market and does not consider the influence in an international context. Therefore, it remains an open question whether the badge is sufficient to reduce the information asymmetry and help entrepreneurs to attract distant investors.

To study the influence of the PWL badge, a dummy variable is used in Equation 9 to indicate the projects that demonstrate this characteristic. Table 5 provides an overview on the overall number of PWL projects in the data sample. The table shows that approximately 25% of all observations carry the PWL badge. The low number is comprehensible and necessary because it guarantees a priority status. An excessive use of the badge might diminish its perceived value as a sign of quality.

[37] Additional information for the decision criteria that Kickstarter uses to award the “Projects We Love” badge is described on their website at: https://help.kickstarter.com/hc/en-us/articles/115005135214-How-does-my-project-become-a-Project-We-Love-or-get-featured-on-the-homepage-.

Table 5: Observation Frequencies for PWL Projects

This table shows the share of PWL projects in the data sample. The badge is rewarded by Kickstarter employees for interesting and creative projects. The table suggests that less than 25% of all observations (from 211,695 projects) receive this signal of quality.

Variable	Observations	% of total
PWL	275,843	25%
Non-PWL	824,811	75%
Total	1,118,654	100.00%

3.4.4.6 Large Projects (Herding Behaviour)

Preceding research on crowdfunding emphasizes the important role of herding behaviour on investment decisions (cp. section 2.3.2.1). Studies have found that visible accumulations of backers encourage additional participation by others and can, therefore, rapidly lead to a self-enforcing mechanism that favours already successful projects (Colombo et al., 2015; Zhang & Liu, 2012). Herding behaviour is considered an effective quality signal that can help market participants to lower the information asymmetry problem and promote the integration of markets (Thierer et al., 2015).

Due to the high relevance of herding behaviour in crowdfunding, this thesis attempts to implement this influencing factor in the model. Unfortunately, since no chronological data is available that would make it possible to trace the development of the count of backers over time, this thesis must rely on a proxy variable instead.[38] The assumption is that projects that display an exceptionally large count of total backers, are likely to have been affected by herding behaviour. Therefore, this thesis uses a dummy variable to distinguish between small and large projects. The used threshold is 107 backers, which

[38] A proxy variable is a variable that is not in itself directly relevant, but that serves in place of an unobservable or immeasurable variable.

is the median for the total count of backers across the sample. The median is a better measure than the average count of backers because in count variables the average is typically distorted by few exceptionally large observations. Overall, the data sample contains 562,163 observations (49%) that are classified as large projects and are likely to have been affected by herding behaviour.[39]

This thesis is aware that the used approach to control for herding behaviour exhibits certain weaknesses. The final count of backers is not always an indicator for herding behaviour as it only provides a retrospective measurement. Similarly, the choice of the threshold might be questioned for its accuracy. However, the used approach currently demonstrates the best solution for this thesis to implement the influence of herding behaviour in the model, given the limited access to relevant data and lack of alternative measures in the literature.

One possible concern could be that the proxy variable correlates strongly with the dependent variable and might, therefore, affect the validity of results. However, *Large Projects (LP)* is distinctively different because it describes the total sum of backers (from all countries) that contributed to a project whereas the dependent variable describes only the count of backers from one specific country. The variables are not directly correlated because a project can also be considered "large" if most backers were received from only one country. An example for this scenario is provided by the project from the electric skateboard manufacturer "Arc Boards".[40] In total, the project attracted 180 backers and is, therefore, considered a large project (because the total count is larger than the defined threshold of 107 backers). However, the backer origin distribution (see Table 6) shows that most backers came only from one country and that comparatively few backers invested from other countries.

[39] Overall, the dataset consists of 1,118,654 observations from 211,695 projects.
[40] https://www.kickstarter.com/projects/arcboardsev/arcboardsev/community

Table 6: Observation Example for Comparison of Dependent- and LP variable

This table shows the backer origin distribution of a specific project from Kickstarter. It indicates that the count of backers is not necessarily affected by the total sum of backers, which is used to proxy herding behaviour.

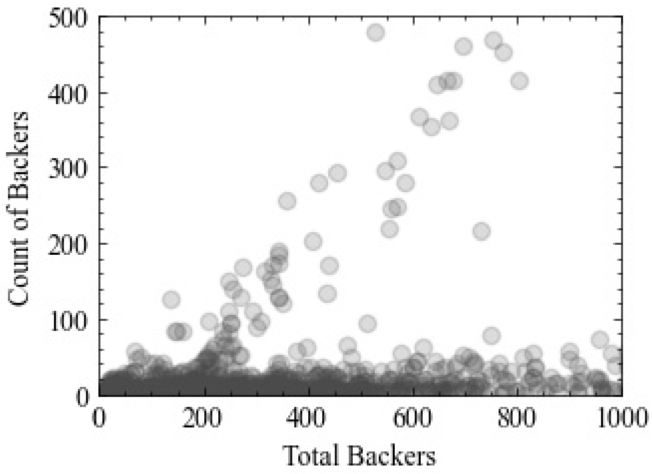
Count of Backers (Dependent Variable)	Total Sum of Backers (Large Project)	Backer's Country
119	180	Singapore
21	180	United States
4	180	Malaysia
3	180	Australia
3	180	Indonesia
2	180	Austria
2	180	Canada
2	180	France
2	180	Switzerland
1	180	China

The provided example shows that the dependent variable (count of backers) is not directly linked to the characteristic "Large Project" and, therefore, can take on disparate values. Unfortunately, no reliable method could be found in the literature to measure the true correlation between both variables. The difficulty lies in the fact that the dependent variable is a count variable that naturally does not follow a normal distribution. Most of the recommended approaches, however, presuppose normality (cp. Point-biserial correlation coefficient).

One approach to evaluate the possible correlation between the dependent- and LP variable is to conduct a graphical analysis. For this purpose, a scatterplot is created that shows the count of backers (dependent variable) on the Y-axis and the total sum of backers (which is the basis for the definition of the LP variable) on the X-axis (see Figure 9). The Figure shows that although in some cases the dependent variable increases in analogy to the total sum of backers, most observations are not directly related.

Figure 9: Correlation Analysis for the Dependent- and LP-Variable

This figure shows a scatterplot with the Count of Backers (dependent variable) on the Y-axis and the total sum of backers (basis for LP-variable) on the X-axis.



In addition to the graphical analysis, the results of the negative binomial regression model (Table 1, Appendix C) have been compared to the results of a model that excludes the LP variable (Table 2, Appendix C). Overall, no relevant differences could be found, neither in the influencing directions of the other independent variables, nor in their statistical significance in terms of p-values.

The decision to include the LP variable into the final model, however, is based on two arguments. First, herding behaviour is a highly relevant phenomenon in the crowdfunding literature and, therefore, including a possible proxy that measures the potential effect of herding behaviour is important for this research. The second argument in favour of including the LP variable is that the model estimates produce a lower Aikake Information Criterion (AIC) if LP is included. The AIC is a proven method to compare the relative quality of different models (McElreath, 2016).

3.4.4.7 Covid-19 Pandemic

In 2020 the Covid-19 pandemic significantly affected the global economy. For example in the US, the pandemic led to a 43% drop in the aggregate stock market between 19 February and 23 March 2020 (Thorbecke, 2020). Additional prevention measures introduced by governments (i.e., contact restrictions, travel bans, imposed closures) have slowed down the global economy and caused economic difficulties for many companies. The effects are visible across a wide range of sectors. Many “traditional” industries have suffered tremendous losses, most notably are companies related to aerospace, real estate, tourism, oil, brewers or retail (Thorbecke, 2020). However, the pandemic has also produced some winners. Especially internet or technology-related companies such as Amazon, Zoom, Pinterest or Netflix experienced a significant profit surge during the pandemic (De' et al., 2020). This is because the products of these companies can be consumed from the comfort of one’s own home and have not been affected by the earlier mentioned prevention measures. On the contrary, they offer a popular pastime or virtual alternatives to interact with other people and, therefore, attracted additional demand during the lockdowns. In this context, an interesting question is how the Covid-19 pandemic has affected the crowdfunding industry.

The assumption is that crowdfunding should count itself among the group of winners because it is a digital concept that takes place on the internet and does not require physical interaction. However, the statistics of crowdfunding platforms do not show a consistent pattern. While some (equity-based) crowdfunding platforms report new records in investment volumes (Bergman, 2020), others state exactly the opposite. According to Kickstarter (Leland, 2020), the number of live projects decreased by 25% in April 2020 compared to the previous year. In August 2020, live projects were still 7% below the prior year. Also, the news website “crowdfundinsider.de” (Alois, 2021) states that 2020 was the first year that the crowdfunding market has declined. One possible explanation for this observation is that global supply chains have been affected considerably in the weeks immediately following the virus outbreak. This affected crowdfunding projects in the respect that companies were unable to get essential supplies, get their products made, or send those products to backers (Hecht, 2020). The consequence is that multiple companies were forced to announce delays or suspend their crowdfunding campaigns. Similarly, the willingness of backers to invest into crowdfunding projects might have been considerably affected by the pandemic.

Personal economic uncertainty due to company closures and mass layoffs (International Labour Organization, 2020), might lead to backers being more cautious about their investments. Yet another possible scenario is that the effect of the crisis is not consistent but has changed over time. For example, it is possible that after the initial shock of the economic crisis the crowdfunding industry recovered quickly when governments started to provide support to both individuals and businesses in many countries around the world. Moreover, entrepreneurs and small business owners might have relied on crowdfunding in particular at a later stage of the pandemic to call for help after they realised that their financial reserves are insufficient to survive the crisis and they are not able to raise capital from traditional financing sources (e.g., banks).

An additional reason why the influence of the Covid-19 pandemic is interesting in the context of this thesis is because research shows that adverse economic shocks can lead to an increase in Home Bias in capital allocation decisions, also known as the "flight home effect" (Giannetti & Laeven, 2012). For example, Giannetti and Laeven (2012) find that the proportion of loans granted to domestic borrowers increases by approximately 20% (at the expense of foreign loans) if the country of origin of the bank experiences a crisis. Therefore, the Covid-19 pandemic is an important factor to consider in Home Bias research.

So far, no literature has been identified on the potential effect of the pandemic on the crowdfunding industry, apart from newspaper articles and a few reports from crowdfunding platforms and international organisations. The Covid-19 pandemic is still too recent to be subject of extensive scientific inquiry. Moreover, the (already) time-consuming data collection and analysis are additionally aggravated by the ongoing contact- and travel restrictions. In general, no scientific research exists that examines the influence of crises on the crowdfunding industry. Crowdfunding became popular only after the 2008 financial crisis and, therefore, has never been exposed to a global economic downturn.

This thesis is the first to study how the Covid-19 pandemic affected the performance of reward-based crowdfunding projects. In general, three different effect-scenarios are possible. One scenario is that the Covid-19 pandemic has a negative influence on the count of backers. For example, backers might be affected by the fear of losing their job or the general uncertainty about the future, as described above, and therefore reduce their engagement in crowdfunding activities. A second scenario is that the Covid-19 pandemic has a positive influence on crowdfunding performance. One possible explanation for this scenario would be the assumption that due to the increased home office regulations and additional free time, backers might have more time to engage with possible projects. Another explanation could also be that backers might actively want to support entrepreneurs and small-business owners who have been hit particularly hard by the pandemic (Leland, 2020). Finally, the third scenario is that the pandemic did not have any significant influence on the performance of reward-based crowdfunding projects.

To study the initial effect of Covid-19 on the count of backers in reward-based crowdfunding projects, a dummy variable is included to distinguish between projects that were launched before and during the global pandemic. The WHO declared Covid-19 the status of a global pandemic on March 11th 2020 (WHO, 2020). Therefore, this date is used as a threshold to distinguish between projects that have been affected by Covid-19. In total, the data sample contains 21,692 distinct observations (from 3,552 projects) that have been affected by the virus within the time frame of this thesis (April 2009 to June 2020).[41]

[41] Overall, the dataset consists of 1,118,654 observations from 211,695 projects.

3.5 Model Interpretation

The purpose of the Negative Binomial regression model is to examine how the count of backers in international reward-based crowdfunding (dependent variable) is affected by different influencing factors such as geographical distance, GDP per capita, PWL badge (third-party endorsements), project category, large projects (herding behaviour) and Covid-19 pandemic (cp. Eq. 9). However, to be able to interpret the results of the model estimation properly, it is necessary to first understand which key figures are important and which conclusions can be derived from them. The interpretation of the Negative Binomial regression results is different from traditional OLS because of the log link function which places the coefficients on the natural log scale and, therefore, requires additional explanation (Leeper, 2017). Another important factor that affects model interpretation and must be addressed is the exceptionally large data sample of this thesis. Current literature suggests that traditional model interpretation approaches might be unreliable when dealing with Big Data because p-values are demonstrably affected by the sample size. The purpose of this section, therefore, is to provide a general understanding of how to interpret results from Negative Binomial regression models and what additional attention must be paid to when dealing with Big Data models.

3.5.1 P-Value Problem in Big Data and Recommended Remedies

A general convention to interpret the results of econometric models is first to inspect the coefficient estimates and their respective significance in terms of probability values (see Leeper, 2017). The coefficient estimates suggest the effect magnitude and the influencing direction on the dependent variable for a one-unit change of a given independent variable. For example, a negative coefficient indicates that the increase of the independent variable contributes to a reduction of the dependent variable. Accordingly, a positive coefficient denotes a rise of the dependent variable as the independent variable increases. The intercept is the expected value of the dependent variable when all independent variables have a value of zero.

However, the observed relationship expressed through the coefficients alone is not meaningful on its own as the relationship can also occur due to random chance. Because of this reason, a common approach in quantitative research is to use the concept of null hypothesis significance testing (NHST) via statistical probability values (Singh Chawla, 2017). The fundamental ideas of NHST date back to Fisher's (1925) concept of significance testing and Neyman & Pearson's (1928) suggestions of acceptance based on critical rejection regions (Pernet, 2015). Despite of the concept's old age, NHST is still the most common statistical method of choice to evaluate the evidence of an effect in biological, biomedical and social sciences (Nuzzo, 2014; Pernet, 2015; Singh Chawla, 2017; Wasserstein & Lazar, 2016).

In NHST, the researcher defines the null hypothesis (H0), which is the assumption that no effect or no statistically significant relationship exists between the inspected variables. For example, in the context of this thesis, H0 could be the assumption that no statistically significant relationship exists between geographical distance and the count of backers from a given country funding a project in another country. In conjunction with the model estimation, probability values or *p-values* are calculated for the coefficients of each independent variable (Russell A. Poldrack, 2018). The p-value is the cumulative probability of observing a result at least as extreme as the test statistic, assuming that the null hypothesis of no effect is true (Harvey, 2014; Russell A. Poldrack, 2018). In other words, the smaller the p-value, the smaller the probability to observe the specific value in a population of estimates (Pernet, 2015).[42] To make an unbiased decision whether a result is statistically significant, researchers must specify a rejection threshold for H0 (Pernet, 2015). A p-value that lies below this predefined threshold suggests that H0 can be rejected in favour of an alternative hypothesis. In the beginning of the 20th century,

[42] If the data is not normally distributed, for example when dealing with count variables, more advanced techniques must be used for robust model estimations. One possible technique is to transform the respective variables (i.e., using the Log of the variable). Another is to use Generalized Linear Models such as Poisson regression or Negative Binomial regression (Cameron & Trivedi, 2010). The advantages and disadvantages of the different techniques are described in section 3.4.3.

Fisher (1925) suggested the threshold value of 0.05, which ever since has become a standard in scientific research (Mather, 1951; Pernet, 2015). However, it is also common practice to report multiple levels of significance at the same time (i.e., $p<0.05$, $p<0.01$ and $p<0.001$) (cp. Pernet, 2015).

In recent years, however, the widespread use of NHST has raised increased concern among scientists (Nuzzo, 2014; Wasserstein & Lazar, 2016). Experts emphasize different problems that can arise when researchers solely rely on p-values and coefficient signs to determine the significance of their findings (Benjamin et al., 2017; Harvey, 2014; Sidhu & Doyle, 2016). For example, one concern is that the threshold of 0.05 is not sufficiently low and can promote the generation of false positives (Singh Chawla, 2017). False positives or *type I errors* are findings that are wrongly believed to be true. In this case, H0 is rejected although no significant relationship between the variables exists in reality. The discussion about the validity of p-values has gained momentum after numerous studies failed to reproduce the results of important prior publications (Schooler, 2014). This is often referred to as "reproducibility crisis" in science (Schooler, 2014; Singh Chawla, 2017).

One potential solution to overcome this reproducibility crisis is to lower the p-value threshold from 0.05 to a value of 0.005 (Benjamin et al., 2017; Singh Chawla, 2017). However, this solution has equally received criticism. Researchers claim that by reducing the p-value threshold, studies run into the danger of producing false negatives or *type II* errors. This is when H0, the assumption of no effect, is accepted although a significant relationship between the variables exists in reality. Moreover, Singh Chawla (2017) argues that simply lowering the threshold for the p-value will not change the core of the problem. The major issue is that if an unlimited amount of different experiment designs is possible, from a statistical point of view, at least one of them will produce statistically significant results (Singh Chawla, 2017). In combination with the so-called "file-drawer" problem or "publication bias", which is the general tendency of journals to publish statistically significant or "positive" results, some individual (positive) studies might acquire disproportionate attention and outshine the large amount of "negative" results

(Singh Chawla, 2017). For this reason, scholars emphasize that continuous replication efforts are crucial to identify the validity of findings (Schooler, 2014). This suggestion is also in accordance with Fisher (1973), the inventor of the p-value approach, who consistently emphasizes that no single experiment, however significant in itself, can suffice for the experimental demonstration of any natural phenomenon.

The warning that new knowledge should not be dependent on individual studies, is particularly important for crowdfunding research. Crowdfunding is a comparatively new field of study and, as it is common for any emerging research field, the first publications have not yet been sufficiently scrutinised or replicated by other scientists to make clear or universally true statements about the industry.

Another frequently overlooked aspect in the current debate on NHST is that technological advancements have considerably changed the research context. The internet has significantly facilitated data collection, which enables researchers to construct unprecedently large data samples. Large data samples have positive and negative consequences. On the one hand, they provide scientists a better insight into the population and can help to develop theories that are closer to the "truth" (Elston, 2018).[43] On the other hand, researchers raise justified concerns whether the traditional approaches of statistical inference and hypothesis testing via p-values are still applicable and valid in research with "Big Data" (Demidenko, 2016; Harvey, 2014).

A concrete example for "the p-value problem" in Big Data is provided by M. Lin, Lucas, and Shmueli (2013) who show that p-values can quickly go down to zero with increasing sample sizes. Their findings suggest that if a sample is large enough, econometric models possess the power to identify marginally small, subtle, and complex patterns in the data

---

[43] M. Lin, Lucas, and Shmueli (2013) highlight different opportunities that large data samples provide for research. For example, the scientists argue that large samples allow the detection and quantification of small or complex effects. Moreover, large samples can be easier divided into subsamples of interest while maintaining enough power in each sample. Large samples also allow to incorporate many control variables without losing in estimation power.

that, however, have little or no practical value. The authors find that samples of more than 10,000 observations are already prone to the p-value problem. In this context, Chatfield (1995) warns that scientific research could miss the connection to practice. The author comments that scientists should not pursue statistically “significant” patterns, as in large samples they nearly always are, but rather search for patterns that are interesting and purposeful for practice.

The p-value problem in Big Data is not new and has been addressed by several scientists in the past and in more recent literature (Cohen, 1995; Demidenko, 2016; Gelman & Stern, 2006; Harvey, 2014; M. Lin, Lucas, & Shmueli, 2013; Pharoah, 2007). In crowdfunding research, the p-value problem is a particularly important influencing factor because most studies are conducted on comparably large data samples (cp. Table 2, p. 65). This is because crowdfunding platforms offer the possibility to access databases of thousands of projects. By using this information, researchers can construct highly representative data samples with high validity. However, due to the large sample sizes, crowdfunding research is particularly prone to type I errors. Meaning that H0 is rejected in favour of an alternative hypothesis, even if H0 is true. In other words, in crowdfunding research it is easier to produce statistically significant results.

Therefore, the major caveat of this thesis is to wrongly conclude that distance does have a significant influence on the count of backers although its potential effect size might be marginally small and irrelevant for practice.

The p-value problem is one major reason why this thesis extends the common approach of NHST and goes beyond the analysis of statistical significance. Lin et al. (2013) provide three recommendations to improve research with large samples. First, they suggest using *coefficient/p-value/sample-size* (CPS) charts to gain additional information about the variability of the results (cp., 3.5.2). CPS-charts show how p-values and coefficients change with increasing sample size and allow to reveal the p-value problem in the results. The CPS-charts can also be extended through a Monte Carlo simulation to better illustrate the range of possible p-value estimates (cp. Lin et al., 2013). Second, the authors recommend that researchers working with large samples should report confidence intervals. By reporting confidence intervals for a particular variable across different studies, it becomes easier for researchers to conduct meta-analysis, synthesize prior studies and help advance scientific knowledge. Third, Lin et al. suggest that researchers should report easily comprehensible measures for the effect size. This recommendation is supported by Cohen (1994) who states that although a "magic" alternative to the p-value problem might not exist, researchers can improve the validity of their research greatly by focusing on the practical effect size. Effect sizes can be reported in different ways. One approach is to measure the actual difference in the units of the dependent variable for a one-unit alteration of the independent variable of interest (i.e., marginal effect). Yet another approach is to measure the standardized mean difference (i.e., Cohen's d), which has the advantage that it allows a comparison between different studies that use variables with different units (Coxe, 2018).

3.5.2 CPS-Charts for the Data Sample

Lin et al. (2013) suggest using a meta-analysis via a *coefficient/p-value/sample-size* (CPS) charts to illustrate the influence of the sample size on the coefficients and p-values. CPS charts can be developed by repeatedly drawing samples of increasing sizes from the dataset, re-estimating the statistical model (as defined in Equation 9) and plotting the p-values on a chart to see how they are affected by the changing sample size. The detailed procedure is provided in Table 7.

Table 7: Algorithm for generating CPS charts

This table shows the algorithm suggested by Lin et al. (2013a) to create a coefficient/p-value/sample-size (CPS) chart.

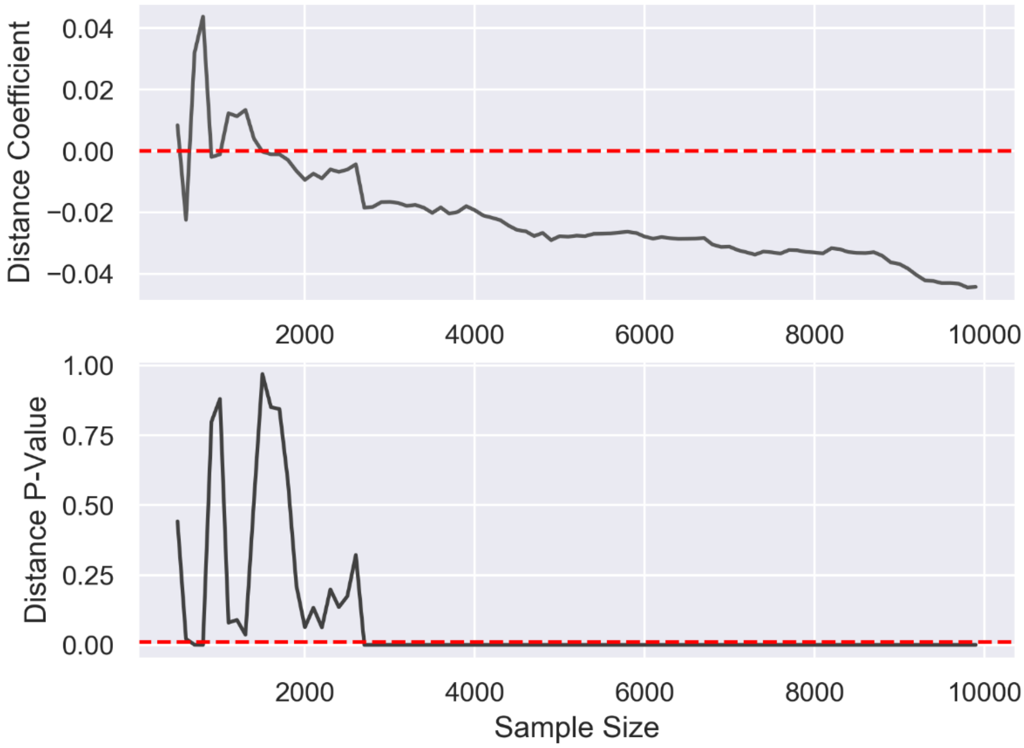
Procedure
1. Choose the minimum sample size that is reasonable for fitting the model;
2. Randomly draw a sample of determined size from the large data set;
3. Fit the model of interest to this sample, and retain the estimated coefficients, their standard errors, and the p-values;
4. Increase the last sample size by adding more observations, drawn randomly from the remaining data set;
5. Repeat steps (3) to (4) until the full original data set is used;
6. Finally, create a line plot of the coefficients vs. the sample size (on the x-axis), and in another panel the p-value(s) vs. the sample size.

Figure 9 shows the CPS chart for the Distance variable in Eq. 9. The CPS chart is in accordance with the findings of M. Lin, Lucas, and Shmueli (2013), showing that p-values tend to thrive towards zero with increasing sample size. The p-value for distance, for example, drops to near zero at a sample size of 2,300 observations and remains at this level for all subsequent estimations.[44]

[44] Although CPS charts are only reported for Distance, the observation that p-values drop to near zero with increasing sample size is consistent for all variables used in this thesis.

Figure 10: Coefficients and P-Values in Relation to Sample Size

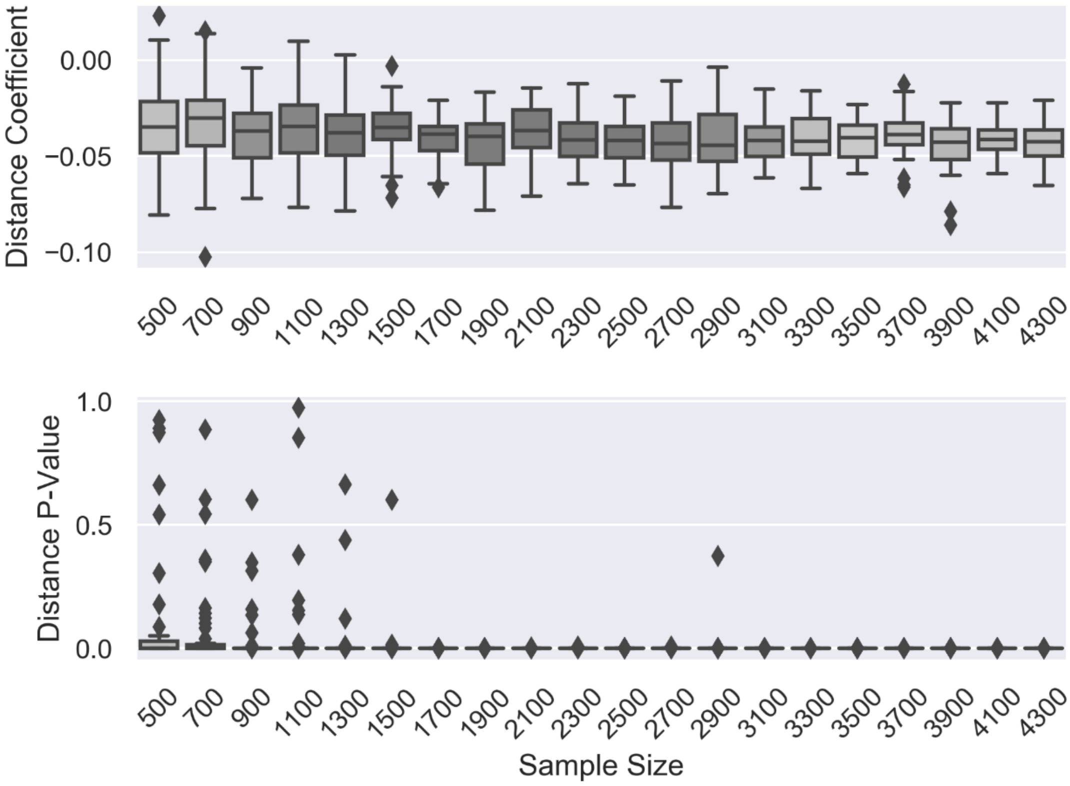
This figure shows a CPS chart as suggested by Lin et al. (2013). The upper figure shows the development of the coefficient for distance in relation to the sample size. The lower figure shows the development of the p-value for distance in relation to the sample size.



The CPS chart displayed in Figure 9 is based on one random sample at each sample size. This means that the diagram is directly affected by extreme values that might occur due to chance. Because of this reason, M. Lin, Lucas, and Shmueli (2013) suggest additionally to use Monte Carlo simulations to better illustrate the range of possible p-value estimates. A Monte Carlo simulation can be employed by drawing various random samples at each sample size. This approach extends the CPS charts because it reduces the risk of attaining extreme results at any given sample size. The list of estimates, generated through the Negative Binomial regression model (as specified in Eq. 9), is used to create boxplots, which display the different distributions of p-values in a compact manner.[45] This thesis uses a Monte Carlo simulation with 50 repetitions.[46] The Monte Carlo simulation for distance is presented in Figure 10.

Figure 11: Monte Carlo CPS Chart

This figure shows a Monte Carlo CPS chart, which is the coefficient and p-value as a function of sample size. The upper figure shows the development of the distance's coefficient in relation to the sample size. The lower figure shows the development of the distance's p-value. In total, 50 random samples were used at each sample size.



[45] The box plot is a diagram used to graphically represent the distribution of a characteristic. It combines various robust dispersion and location measures in one plot. A box plot is intended to give a quick impression of the area in which the data are located and how they are distributed over this area. Therefore, all values of the so-called five-point summary, i.e. the median, the two quartiles and the two extreme values, are displayed.

[46] M. Lin, Lucas, and Shmueli (2013) use in their example 400 random samples at each sample size. Due to computational constraints it is not possible to use the same number of samples. However, 50 random samples are sufficient to illustrate the effect of increasing sample size.

Figure 10 shows that the median coefficient value for distance is stable across different sample sizes, and its variability decreases with increasing samples. Moreover, for samples below *n* = 1,500 the distribution covers the value zero, yielding statistical insignificance at traditional significance levels (cp. Lin et al., 2013). The plots show decreasing noise in the coefficient estimation, reflecting the power of an increasing sample size. A similar observation can be made for the p-values. The plots reveal that both the levels of p-values as well as their variability in the distribution decrease rapidly with increasing sample sizes.

Overall, the observations are in accordance with the findings of M. Lin, Lucas, and Shmueli (2013) and show that the statistical significance of a result is a function of the sample size (Lykken, 1968). The CPS analysis demonstrates that p-values on their own represent an unreliable measure in Big Data and can lead to wrong assumptions. Different authors have provided possible explanations for this phenomenon. Meehl (1990) suggests that, ultimately, everything correlates to some extent with each other. The author refers to this observation as the "Crud Factor" (see also Cohen, 1994). Lykken (1968) refers to the same phenomenon as "the ambient correlation noise". The author states that "statistical significance, perhaps the least important attribute of a good experiment, is never a sufficient condition for claiming that (1) a theory has been usefully corroborated, (2) a meaningful empirical fact has been established, or (3) an experimental report ought to be published" (Lykken, 1968). To guarantee the relevance and reliability of findings, supportive assessment measures are required that go beyond reporting p-values and coefficient signs. In this context, and as a potential solution to the p-value problem, many authors in the current literature emphasise the great importance of determining the magnitude or the actual effect size of the discovered relationship patterns between variables.

3.6 Evaluating Practical Relevance of Results

Research has often been criticized for focusing too much attention on statistical significance, while often neglecting the practical usefulness of findings in the real-world (Carver, 1978; M. Lin, Lucas, & Shmueli, 2013; Sawyer & Peter, 1983). Kirk (1996) refers to the usefulness of findings also as the ”practical significance”, a term that enjoys increasing popularity in the literature (M. Lin, Lucas, & Shmueli, 2013; Mohajeri et al., 2020; B. Thompson, 2002). Mohajeri et al. (2020) introduce an additional distinction between “practical significance” and “practical relevance” of findings. According to the authors, practical significance describes only the “research impressiveness”, whereas the term “practical relevance” refers to the real-world usefulness.

To minimize the confusion of terms, this section explains the different levels of analysis that researchers can conduct to evaluate the real-world usefulness of their findings. In this context, this thesis adopts some of the recent ideas introduced by Mohajeri et al. (2020). However, additional explanations are provided how this thesis distinguishes the terms statistical significance, practical significance, and practical relevance. A summary can be found in Table 8.

Table 8: Evaluation Criteria for Findings

This table describes the different levels of analysis that researchers can conduct to evaluate the relevance and usefulness of their research findings.

	Statistical Significance	**Practical Significance**	**Practical Relevance**
Purpose	Measure statistical discernibility	Measure impressiveness of a relationship or effect	Measure usefulness of research and its results
Primary Adjudicator	Researcher / Academic Audience	Researcher / Academic Audience	Practitioners / Non-academic Audience
Approach	NHST with predefined p-value thresholds	Estimating effect-sizes and confidence intervals	Translating research results into more comprehensible scale and interpret them on their practical influence in a current context
Weakness	In large datasets p-values thrive towards zero. Does not allow to infer practical usefulness of results	Effect sizes are often misleading or difficult to understand for practitioners	No unambiguous mapping from effect size to a value of practical importance

3.6.1 Statistical Significance

In many research disciplines, researchers have been used to considering statistical significance, in terms of NHST and predefined p-value thresholds, a cornerstone of “scientific” inference (Mohajeri et al., 2020). However, with increasing popularity of Big

Data, where the described “p-value problem” has drawn much attention (Lin, Lucas, & Shmueli, 2013), more and more scientists begin to question the adequacy of statistical significance to warrant the scientific merits of research (e.g., Chatfield 1995; Cohen 1977; Daniel 1977; Fisher 1925; Kerlinger and Pedhazur 1973; Kirk 1996; Nickerson 2000; Pearson 1900; Selvin 1957).

The main motivation for going beyond traditional statistical analysis is that results can sometimes be statistically significant (especially in large samples) but provide little added value for practitioners or scientists (Hryniewicz, 2018; M. Lin, Lucas, & Shmueli, 2013; Singh Chawla, 2017). Poldrack (2018) describes the potential divergence of statistical and practical significance of findings with the example of a study that analyses the effect of a particular diet on body weight. Although researchers might find a statistically significant effect at $p < 0.05$, the finding does not describe how much weight was factually lost. The author claims that the effect of a diet that is statistically significant but merely leads to a weight loss of ten ounces (i.e., the weight of a bag of potato chips) would be described by many people as largely irrelevant in practice.

For this reason, many research fields encourage scientists to go beyond statistical analysis and evaluate the practical significance of their findings (Kirk, 1996; Noordzij et al., 2017; B. Thompson, 2002; Vacha-Haase & Thompson, 2004).

3.6.2 Practical Significance

According to Mohajeri et al. (2020), the term practical significance describes the research *impressiveness* of results. In other words, an analysis of practical significance is typically conducted to ensure that the magnitude of the reported results is sufficiently impressive so that the hypothesized relationship or effect does make a difference in the real world.

In many research fields, the effort to encourage researchers to address the practical significance of findings, alongside the statistical significance, is mainly focused on promoting the practice of reporting and interpreting effect sizes (Mohajeri et al., 2020).

For example, Lin et al. (2013) suggest that researchers should, as much as possible, be objective and clear in helping readers to understand the true meaning of the coefficient estimates within the study context. Some authors refer to this shift of focus also as the "effect size movement" (Robinson et al. 2003, p. 51).

Interpreting effect sizes in Negative Binomial regression models is more difficult than in traditional OLS because of the log link function, which places the coefficients on the natural log scale (Leeper, 2017). Therefore, scientists often prefer to analyse the incidence rate ratios (IRRs) (Beaujean & Grant, 2016). IRRs are the relative measures of the coefficients that can conveniently be interpreted as the multiplicative effect or *semi-elasticity* (Hornuf & Schwienbacher, 2018). They describe the relative influence of the independent variable on the prediction (i.e., count of backers). In this context, all estimates <1 indicate a negative effect or a decrease of the dependent variable, while estimates >1 reveal a positive relationship.

The IRRs can be obtained by exponentiating the coefficient estimates that result from the regression model (as specified in Equation 9), as described in Equation 10. The interpretation of the IRRs, rather than the coefficient estimates, is particularly common for Poisson and Negative Binomial regression models (see Hornuf & Schwienbacher, 2018; UCLA, 2019a).

Equation 10: Transformation of Eq. 9 to Attain the Incidence Rate Ratios

$$log(\lambda) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

$$\lambda = \exp\,(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \ \beta_p x_p)$$

$$\lambda = \exp(\beta_0) * \exp(\beta_1 x_1) * exp(\beta_2 x_2) * \ldots * \exp\,(\beta_p x_p)$$

IRRs are useful because they describe the *marginal effect size* of a variable. The marginal effect size is the estimated influence that an independent variable (i.e., distance) has on the dependent variable for a one-unit change (Mohajeri et al., 2020).

Although reporting IRRs as a measure of effect size is common in count regression models, it can lead to potential confusion and misconceptions that have been described by different authors in the literature (Higgins et al., 2019; Kampenes et al., 2007; Noordzij et al., 2017) and should, therefore, be addressed in the context of this thesis. For example, Mohajeri et al. (2020) claim that despite what seems to be a major turn toward practical significance, it appears that what is referred to as practical significance is vastly conflated with the notion of relevance. The authors highlight the problem that a tendency exists in the literature to equate the research impressiveness of quantitative results with the real-world usefulness of findings, and hence not to make a full distinction between practical significance and relevance (see also Kelley & Preacher, 2012; Kirk, 1996). One possible explanation for this tendency, as described by Mohajeri et al. (2020), is that the relevance of the conceptual aspects (e.g., research models, hypotheses, variables, etc.) of statistically conducted research is often a taken-for-granted quality among statisticians. On this basis, statisticians might presume that the only remaining condition to achieve relevance for the entire research is to obtain impressive results, which according to the authors is questionable.

This thesis supports the distinction between practical significance and relevance as explained by Mohajeri et al. (2020) and extends their arguments by drawing attention to the fact that different possibilities exist to report effect sizes, which sometimes can be misleading. The following section aims to explain why even results with apparently high relative effect sizes (and practical significance) can sometimes be considered marginally relevant in terms of practical usefulness (practical relevance).

3.6.3 The Problem of Relative Effect Sizes

Effect size measures can broadly be divided into ratio measures and difference measures, or relative and absolute measures, respectively (Higgins et al., 2019). Absolute measures describe the factual difference (i.e., between two groups) in the units of a specific variable (e.g., euros, kg or count of backers). Relative measures, such as the IRR, describe the effect of a variable on the outcome of interest (dependent variable) compared to a baseline estimate. For example, relative effect size measures are common in medical studies and are used to describe whether the count of a certain incident (e.g., lung cancer) is higher or lower among one specific group of individuals (e.g., group of smokers) compared to another (e.g., group of non-smokers) (cp. Higgins et al., 2019). Relative effect size measures range from 0 to infinity and are free of unit.

Relative effect size measures are frequently used in medical studies because of their appealing feature of summarizing two numbers (i.e., the risk of getting a disease in one group and another) into one measure (Noordzij et al., 2017). However, some scientists emphasize that relative measures alone can sometimes obscure the real magnitude of an effect, conveying that a discovered influence is more important than it actually is. Noordzij et al. (2017) elaborate on this problem by analysing a study of the World Health Organization (WHO) on the risk of thyroid cancer. In 2013, newspapers reported that according to the WHO's study, the risk of thyroid cancer is 70% higher among females that have been exposed as infants to the Fukushima Daiichi nuclear disaster in Japan in 2011 (Nebehay, 2013). This relative effect size measure is estimated from statistics showing that about 1.25 out of every 100 girls (1.25%) in the area developed thyroid cancer due to the radiation exposure, instead of the natural rate of about 0.75% (1.25/0.75 = 1.7) (Nebehay, 2013). Noordzij et al. (2017) highlight that although the relative effect size of 70% sounds alarming, it must also be emphasized that due to the low baseline rates of thyroid cancer overall, even a large relative increase in the risk probability represents only a small absolute increase in the risk of obtaining the disease. In other words, females that grew up in the area of the Fukushima Daiichi nuclear disaster still have a comparably low risk probability of obtaining the disease in practice (1.25%). Accordingly, growing up in the affected region only increased the probability of attaining thyroid cancer by 0.5%, in absolute terms (1.25/100 – 0.75/100 = 0.005) (cp. Noordzij et al., 2017).

The presented example illustrates the core problem of relative effect size measures. They can become over-proportionally large when the baseline values are comparatively small and, thereby, lead to wrong impressions of the practical relevance of a discovered effect. In the example above, a relative effect size measure of 70% must be seen in relation to the increased risk probability of 0.5%. Noordzij et al. (2017) also describe a scenario in which scientists might claim that a certain medical treatment reduces mortality by 50%, when the intervention reduces death rates of a specific disease merely from 0.002% to 0.001%. According to the scientists, the practical relevance of such improvement can be questioned (especially if other side effects of the treatment are involved).

Understanding the difference between relative and absolute effect size measures is important in the context of this thesis, because a similar misconception of the true relevance of a discovered relationship can occur. For example, if the mean count of backers (dependent variable), estimated through the negative binomial regression model, is mostly small (i.e., between 2 to 5 backers), large IRRs might describe comparatively small absolute differences, if measured in the units of the dependent variable (count of backers). In other words, a notable IRR for distance (e.g., 50%) might represent a comparatively marginal reduction in the factual count of backers. This is because 50% of 2 backers is still only a decrease by 1 backer in absolute terms, which might be marginally relevant in a practical context.

Due to the possible scenarios described above, different authors suggest that both relative- and absolute effect size measures should be reported, as neither the relative nor the absolute measures alone provide a complete picture of the effect and its implications (Higgins et al., 2019; Kampenes et al., 2007; Noordzij et al., 2017).

3.6.4 Practical Relevance

In contrast to statistical significance and practical significance, which are mostly focused on the quantitative results of a particular form of research (i.e., statistically conducted research), the concept of practical relevance is concerned with the actual usefulness of findings in a given context, whether that research is statistically conducted or not (Mohajeri et al., 2020). The concept of practical relevance connects to Kootnz’s (1961) demand that new discoveries in the field of management research should always be considered against the background of what impact they have on practitioners. Similarly, Hambrick (1994) emphasizes that the main objective of management research should be to make a significant contribution to the solution of major problems facing our society and its value-creating enterprises.

One general problem of evaluating practical relevance of findings is that there is no unambiguous mapping from an effect size to a value of practical importance or usefulness (Breaugh, 2003; Hill & Thompson, 2005; Kirk, 1996; Rosenthal et al., 2011; Smart, 2005; Trusty et al., 2004; Vacha-Haase & Thompson, 2004). This is because practical relevance is highly context dependent and different people might have a different perceptions of the practical value of a result (Hill & Thompson, 2005). For example, the relevance of a result can be time dependent. H. M. Cooper (1981) provides an example in which gasoline savings resulting from regularly checking the tire pressure of vehicles might have been considered marginally important in 1970 (at a gasoline price of 30 cents per gallon) but might enjoy a significantly higher relevance in 1980 (at a gasoline price of 1,20 per gallon) or today.

One approach to translate statistical model estimates into more comprehensible measures for practical relevance of a studied relationship is introduced by Coxe (2018). The author recommends using the standardized mean difference (SMD) as an additional effect size measure to IRRs for nonlinear count regression models. SMD removes the unit of a variable in the effect and describes its influence in terms of standard deviations. Using SMD has different advantages. For example, it can help researchers to evaluate the size of an effect when the units of measurement or the scope of variables is not intuitive. Moreover, it allows the comparison of effect sizes across studies that use different units for variables. One common SMD effect size measure for group differences is *Cohen's d* (Cohen, 1988). *Cohen’s d* is determined by calculating the mean difference between two groups, and then dividing the result by the pooled standard deviation

(Coxe, 2018).[47] Another advantage of Cohen's approach is that it provides easily comprehensible interpretation suggestions for the *d*-values that are widely recognized in the literature (Coxe, 2018; Poldrack, 2018; Lakens, 2013). This avoids the problem of choosing arbitrary thresholds to determine if an observed effect size is relevant in practice. According to Cohen's (1977) suggestions an effect size of $d < 0.2$ can be considered as "small", $d > 0.5$ and < 0.8 as "medium" and $d > 0.8$ as "large" (Lakens, 2013). The descriptors have been expanded by Sawilowsky (2009) for $d < 0.01$ and $d > 1.2$ (see Table 9).

Table 9: Descriptors for magnitudes of Cohen's d

This table shows the interpretations suggestions for different magnitudes of the d-value.

Effect size	D value
Very small	0.01
Small	0.2
Medium	0.5
Large	0.8
Very Large	1.20

Another approach to better evaluate the practical relevance of findings is to translate the estimated relative effect size of a given relationship (i.e., the IRR) into a more comprehensible scale that allows also non-academic readers to evaluate the magnitude of its influence. This approach is in accordance with Lin et al. (2013), who suggest that effect sizes should preferably be translated into scales that are relevant for practitioners. For example, this means that rather than stating that a one-standard deviation increase

[47] Coxe (2018) offers an internet-based tool to estimate Cohen's d value for nonlinear count regression models at: https://stefany.shinyapps.io/RcountD/.

in distance will decrease the count of backers by a certain percentage, researchers could explain that a 1,000 km increase in distance will result, on average, in *X* fewer backers. This information is probably easier to understand for entrepreneurs and enables them to evaluate the practical relevance of geographic distance on their crowdfunding campaign more effortlessly.

In Negative Binomial regression, estimating the absolute effect size of a variable in the units of the dependent variable (count of backers), requires setting all remaining independent variables at specific values. This is because the dependent variable is the product of the exponentiated independent variables (cp. Eq. 10). The values of the variables can either be set to the mean (if continuous) or the reference value (if categorical). This approach is also frequently referred to as “Marginal Effect at the Means” or MEM (see Leeper 2017). The MEM is the difference that an independent variable has on the dependent variable when it is changed by one unit and all other independent variables are kept constant at their means (ceteris paribus). According to Lin et al. (2013) using MEM is a more robust approach in nonlinear models to interpret effect size than inspecting the p-value or magnitude of the IRR.

In addition to the described approaches, some authors demand that researchers should be much more strongly encouraged to make their own subjective assessment of practical relevance of their findings. For example, Kirk (1996) argues that although “an element of subjectivity is introduced into the decision process when researchers make this kind of judgement (…) no one is in a better position than the researcher who collected and analysed the data to decide whether or not the results are trivial.” Further, Kirk emphasizes that “it is a curious anomaly that researchers are trusted to make a variety of complex decisions in the design and execution of an experiment, but in the name objectivity, they are not expected or even encouraged to decide whether data are practically significant.”

This thesis offers a broad mixture of different strategies to assess the practical relevance of its findings. In addition to traditional significance testing via p-values, this thesis applies both measures of effect size: MEM as well as *Cohen’s d* to examine the effect of geographical distance on the count of backers in international reward-based crowdfunding. Moreover, confidence intervals are provided for all IRRs. Confidence intervals provide an estimated range of values which is likely to include the real population parameter. Most commonly, the 95% confidence interval is used (Zar, 1999). When calculating a 95% confidence interval, the range in between the reported interval limits includes the true parameter in 95% of cases. To avoid type I errors (cp. section 3.5.1), scientists should preferably take a conservative position by focusing on the lower boundary of the confidence intervals (Lin et al. 2013a). By reporting p-values, confidence intervals and effect sizes, the "Big Three" (Hatcher, 2013), as well as translating the effect sizes into more comprehensible scales (count of backers), this thesis provides a detailed analysis of the statistical significance, practical significance and practical relevance of findings.

3.7 Data

Multiple researchers have addressed the difficulty of collecting location data in crowdfunding research (K. Kim & Hann, 2013; Marom et al., 2014). Typically, researchers have two possibilities to obtain data. One possible approach is to collect data from first-hand experience, also referred to as *primary data* collection (Currie, 2005; N. Salkind, 2010). For example, primary data research can be implemented through experimental design, similar to that of M. Lin and Viswanathan (2016) where the researcher collects data in an "artificial" test scenario (Neuman, 2011). Typically, the researcher has high control over the different variables and can thereby measure the effect size of each variable on the outcome of interest (see Campbell & Stanley, 1967). Other primary data sources include qualitative interviews, as conducted by Gerber and Hui (2013) where participants are directly questioned by researchers about their personal experiences and motivations. The common feature of primary data collection is that the researcher generates new data to answer the question of interest that result from the direct interaction or observation of the research subject (Neuman, 2011; Punch, 2014). Primary data collection has many advantages. For example, it allows focusing directly on the specific research question and provides higher level of control over the collected information. However, primary data collection also involves many limitations. For example, it is often more time and effort consuming to collect data in a sufficient amount. Another common problem, particularly in social science, is that the validity of the findings under “real” conditions is not always provided (Currie, 2005; N. Salkind, 2010). Research in social science deals with human beings who may alter their behaviour depending on the situation. Biased feedback is not uncommon as the research subjects can consciously or subconsciously alter their behaviour and their opinions to fit the socially acceptable norms or to cover up the reality (Neuman, 2011). A famous example is the so-called “Hawthorne-effect”, also referred to as the “observer effect”. This is the change in behaviour of individuals in response to the awareness of being observed (Last & Porta, 2018). The deviation of the observed behaviour from actual behaviour demonstrates a potential threat to the validity of the findings.

Another common challenge that can be associated with autonomous data collection concerns the sample size. Quantitative research requires a sufficiently large and representative sample to guarantee validity of the findings (Neuman, 2011). Crowdfunding is a comparably new industry and still relatively unknown. The identification of potential participants with crowdfunding experience proofs to be difficult, especially if the aim is to conduct a representative, international analysis. Crowdfunding platforms are usually not willing to share user data and their privacy policies forbid the collection of user data from their platforms (see Kickstarter, 2018).[48] Studies that adopt primary data collection are therefore often based on comparably small sample sizes, which might call the validity of the results into question (see Gerber and Hui 2013, Moritz et al., 2014). Another problem of primary data collection is that the researchers, often unintentionally, choose participants that share certain types of characteristics in terms of demographics or cultural background. For example, Gerber

[48] Kickstarter and Indiegogo refuse to share additional information other than the information that is already publicly available on their websites.

and Hui (2013) interview in their crowdfunding research 83 participants that are exclusively from the US. Similarly, Kim & Kim (2017) focus in their cross-border study only on German speaking participants. This narrow focus does not do justice to the versatility of crowdfunding participants and can lead to biased findings.

A different approach to obtain data for research purposes is to rely on *secondary data*. Secondary data is data that has been collected by individuals or institutions other than the researcher. It provides information on past changes or developments and allows time- and cost-efficient research that would otherwise be unfeasible for any individual researcher (MacInnes, 2017). The data demonstrates high validity because it results from real-world behaviour and often provides large enough sample sizes for hypothesis testing (Neuman, 2011). The challenge in secondary research is to rearrange or extend the data in a way that new insights can be gained (Patzer, 1995).

A special characteristic of the crowdfunding industry is that crowdfunding projects usually remain on the internet after termination and are therefore accessible to the general public. Crowdfunding platforms provide valuable secondary data sources, which allow entirely new possibilities for quantitative research. Consequently, many studies rely on secondary data in crowdfunding research (Belleflamme et al., 2013; see Kuppuswamy & Bayus, 2013; Mollick, 2013; Robertson & Wooster, 2015).

3.7.1 Data Sources

Although much information can be directly extracted from crowdfunding platforms (e.g., pledged amount, total count of backers, project goals), location data is significantly more difficult to obtain (Guo et al., 2018; Marom et al. 2014; Kim & Hann 2013). Platform operators who possess the location data of their users are difficult to reach and are often reserved to share any internal information. This might be explained by the fear of competition and the intention to protect the privacy of their users. The lack of cooperation as well as the difficulty to collect location data are possible explanations why only few studies focus on the influence of distance in crowdfunding to date (Guo et al., 2018; Marom et al. 2014; Kim & Hann 2013).

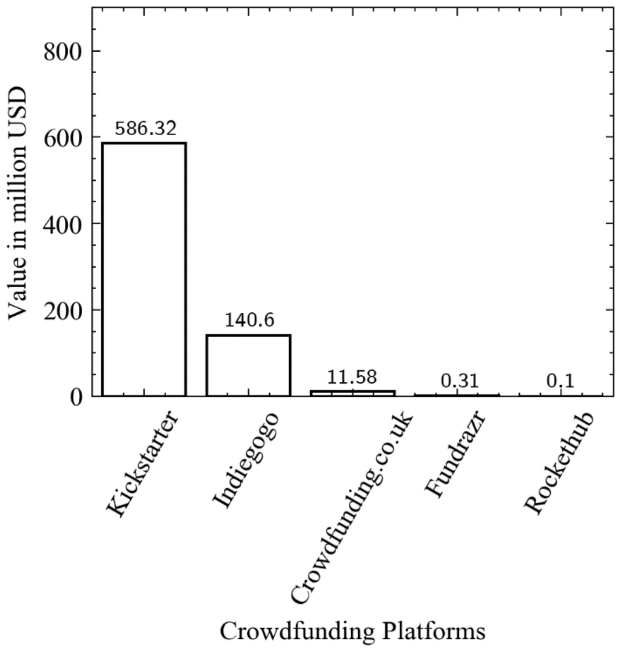
Although the platform operators of Kickstarter show a similar reservation in terms of communication, they do provide certain statistics to the general public. One example is a general statistics page which illustrates the overall number of funded projects by different categories as well as their success rates.[49] Kickstarter also provides data on the individual project pages such as the pledged amount or the number of supporting backers. Although the platform does not publish individual location data of backers, it does provide a general overview on the aggregated count of backers per project from individual countries within the project description.[50] The illustration of this information is a unique feature of the Kickstarter platform and one of the reasons why Kickstarter is chosen as the primary data source for this thesis.

Starting its business in early 2009, Kickstarter is one of the oldest and largest reward-based crowdfunding platforms in the world (Mollick 2014, Frydrych et al. 2014). It has a global community of more than 18 million members and has raised over US$5 billion in funds (Kickstarter, 2020). The platform supports crowdfunding projects in a wide range of areas such as technology, art, gaming and film. However, Kickstarter deliberately excludes charity projects from participation, making it an entirely business oriented crowdfunding platform (Kickstarter, 2019b).

[49] The statistics page can be found on: www.kickstarter.com/help/stats
[50] The country data can be found in the community section of the project description.

Figure 12: Leading Crowdfunding Platforms in 2016

This figure shows the leading crowdfunding platforms worldwide in 2016 by the value of funds raised in million USD. The comparison shows that Kickstarter is more than four times as large as the second largest competitor.[51] *Source: Statista, 2018*



Kickstarter is chosen as the primary data source due to its global scope and dominance in the industry (see Figure 11). It is one of the most famous platforms and is frequently used as the showcase example for crowdfunding (Guo et al., 2018; Mollick & Robb, 2016). Due to the so called "network effect" in which the value of a platform increases with greater number of users, it is likely that Kickstarter will further extend its supremacy in reward-based crowdfunding in the future.[52] Kickstarter's dominant role is also widely recognized by scientists, which explains why much of the existing research builds on data extracted from this platform (Agrawal et al., 2013; Colombo et al., 2015; Frydrych et al., 2014; K. Kim & Hann, 2013; Y. Lin et al., 2014; Mollick, 2014).

[51] Unfortunately, no recent market data could be found on the largest crowdfunding platforms. However, due to its large distance to the competitors and the so called "network effect", it can be assumed that Kickstarter has remained the largest reward-based crowdfunding platform in 2020.

[52] The "network effect" refers to the self-enforcing mechanism of large online networks (see Shapiro and Varian (2008)).

In the context of this thesis, Kickstarter is used to collect data on the aggregate count of backers from the different countries that supported a specific crowdfunding project. This data is used as the dependent variable "*Count of Backers*" (see Eq. 9). The collected data from Kickstarter on the backers' and entrepreneurs' countries of origin is the basis for the construction of the "*Distance*" variable as well as the "*GDP-B*" and "*GDP-E*" variables. The "Projects We Love" badge and the project classification on the platform allows to create the dummy variables "*PWL*" (for third-party endorsements) and "*Project Category*". The collected information on the total count of backers that supported a project (across all countries) is used to construct the dummy variable "*Large Project*" (which accounts for Herding Behaviour) whereas information on the project launch dates is used to create the dummy variable "*Covid-19 Pandemic*".

Another important data source in the context of this thesis is the *World Bank's World Development Indicators* (WDI) database. The database is compiled from officially recognized international sources and presents the most current and accurate global development data available. It is the typical source for country specific estimates on economic performance and has been used by numerous researchers in the past (Burtch et al., 2014; Mendes-Da-Silva et al., 2016). The WDI database is used to extract the GDP per capita information for both entrepreneurs' and backers' home countries (GDP-B and GDP-E variable).

Finally, this thesis uses data from the Google Maps API, which is a digital service offered by Google that allows researchers to determine the specific coordinates of cities. [53] Google's service for geographical data enjoys wide popularity and has been used by multiple researchers in the past (cp. Guenther et al., 2018; Guo et al., 2018). This thesis uses the Google Maps API to obtain the coordinates of each country's capital that is present in the data sample. The city coordinates, in turn, are used to estimate the

[53] The documentation for the geocoding API can be found on: https://developers.google.com/maps/documentation/geocoding/

geographical distance between backers and entrepreneurs and construct the key variable of interest.

3.7.2 Data Collection Process

To collect sufficient data from individual project pages, a web-crawler is used to extract the relevant information. A web-crawler is a computer program that can access the internet via the "Hypertext Transfer Protocol" (HTTP) and copy specific data of interest from the website into a local database or spreadsheet. Web-crawling is a common approach in research on crowdfunding because it facilitates the collection of comparably large data samples with high levels of data validity (Frydrych et al., 2014; K. Kim & Hann, 2013; Mollick, 2014). Because no web-crawler exists that suits the highly specific demands of this thesis, an own crawler needed to be developed. The development required the acquisition of specific programming skills within different computer languages which proved to be elaborate and time-consuming in the process of completing this research. The crawler is written using Python programming language and targets specific website sections of interest. To approach these sections, a fundamental understanding of the Hypertext Markup Language (HTML) is necessary.

The information on crowdfunding projects is extracted from the Kickstarter website and stored in a local database. As of the sampling technique, the crawler uses simple random probability sampling. The most critical requirement of random probability sampling is that each collected observation in the population has an equal chance of getting selected (A. Stuart, 1976). This condition is met because the crawler collects all publicly available links from the platform.

3.7.3 Consideration of Ethics for Data Collection

Ethical guidelines are moral principles that govern or influence conduct. These guidelines are important to ensure that conducted research does not harm or violate social responsibility, human rights, animal welfare, compliance with the law as well as public health and safety. Moreover, ethical guidelines promote honesty, openness, and unbiasedness, which can also lead to a higher credibility and acceptance of findings.

This research is conducted in accordance with the general standards of good practice as described in the guidelines of the Research Ethics Committee of the Sheffield Hallam University (SHU, 2017). At the core, this means that this research orients itself at "beneficence" or the need to "do good". To follow this principle, research must minimize the risk of harm to participants, avoid deceptive practices and protect the anonymity and confidentiality.

This thesis uses the internet to collect data. Internet-mediated research (IMR) is a comparably novel approach of data collection in social sciences. The British Psychological Society (2017) defines IMR as any type of research that involves the remote acquisition of data from or about human participants using the internet and its associated technologies. In general, IMR can take place in a range of online settings. Researchers can use new internet-based tools to collect their own data, for example via online questionnaires and experimental design or rely on existing online databases that accumulate extensive information. Social networks such as Twitter or Facebook offer a rich data source and enable researchers to study social interactions, groupings, and developments over an extended period. A key aspect of IMR is that it usually involves data acquisition from or about individuals without the need of the researcher to be physically close to the research subject (British Psychological Society, 2017). In addition to the great scope and speed of data collection, IMR offers unprecedented possibilities for social sciences.

The new possibilities of data collection, however, also raise several potential ethical issues that need to be considered in detail (Schneble et al., 2018; Woodfield, 2018). For example, the great scope and speed of the internet make it more difficult to stay in control of the research set-up. In addition, the absence of face-to-face co-presence restricts the researcher's capacity to monitor, support or even terminate the study if adverse reactions become apparent (British Psychological Society, 2017). Another important aspect is that in many cases extensive research can be conducted without the individuals' awareness of being studied (Woodfield, 2018). Data on individuals can be collected unobtrusively from blogs, discussion forums or other online spaces (Schneble et al., 2018). The new technologies enable researchers to analyse search engine histories or other digital traces of "likes", content sharing or mobile app usage without users being aware of it. In this context, IMR falls into the current debate on what constitutes "privacy" in the online environment.

The public-private domain distinction online is also important for this thesis, as the data is collected from a public website. According to the British Psychological Society (2017), the answer to what constitutes private information depends much on the aspect whether the users of a particular online service can reasonably expect their data to be public or not. For example, it would be reasonable to expect that posted messages via Twitter are public facing and, therefore, observable by strangers. Conversely, members of a closed forum or Facebook group may expect a certain degree of privacy, since the membership of the group presupposes a specific application process, and the content is mostly invisible for outsiders.

This study relies on the extensive data from the "Kickstarter" crowdfunding platform and analyses the influence of geographical distance on backing decisions. The implementation of this study requires the collection of data on the countries of origin of all crowdfunding participants, as well as data on the project specific characteristics (i.e., project category, count of backers, PWL badge, and launch date). For this reason, the question of potential privacy violation is highly relevant for this thesis. Overall, potential privacy or intellectual property violations must be considered for three groups of people: The backers, the entrepreneurs, and the crowdfunding platform operators.

3.7.3.1 Backers

One important aspect of the Kickstarter community is that public disclosure of the real name, country of origin or any other personal data is entirely voluntary. Backers can decide which information should remain within the closed boundaries of the community and which information will be displayed to everyone on the platform ("Privacy Policy — Kickstarter," 2021). Since Kickstarter does not allow underage persons to participate on their website, who are often less aware of the scope of their actions, this study assumes that all available information was consciously provided and conceded by the backers themselves. The data on backers that is collected for this thesis is open to the general public and does not require any kind of membership or application process to access it. Another important characteristic of this research is that it does not collect the names of the individual backers or any other information that would allow to identify people retrospectively. Backers are held anonymous and only the information on the country of origin is used. Given the large community size of Kickstarter, it is unlikely to retrace the individual backers from this information. Due to the maintained anonymity, this thesis does not compromise the privacy of backers.

3.7.3.2 Entrepreneurs

The second group that is affected by this research are the entrepreneurs. However, it is also reasonable to argue that there are no privacy or intellectual property violations in the context of this thesis. Entrepreneurs are well informed that the disclosed projects on Kickstarter are visible to the public. Therefore, all content of the individual crowdfunding projects is published consciously by the creators and does not reveal any trade secrets. In crowdfunding, the success of a campaign is dependent on the ability to generate attention. Entrepreneurs usually try to actively broadcast their projects via different forms of media (news articles, social networks, private websites, blogs, and forums) and facilitate the possibilities for sharing. Projects can be viewed by the public without any membership or application process. Due to the applied openness of crowdfunding projects, this study does not violate any privacy or intellectual property regulations.

3.7.3.3 Platform Operators

The third group that is affected by this research is the Kickstarter crowdfunding platform itself. The first approach was to contact the platform operators directly and ask them for permission to use the data. Unfortunately, this proved to be rather difficult. Having a community size of more than 12 million users, Kickstarter must process many requests every day. The communication channels are limited and standardized. Most of the requests via the public communication channels are handled by service centres, which often use predefined answers. Moreover, it is difficult to keep a continuous conversation with an employee since every time a different person is responsible for the reply. After several unsuccessful attempts, it was possible to receive an answer from Kickstarter, which however was vague and did not address all asked questions. Kickstarter recommended to use the general statistics page that is provided to the public.[54] Unfortunately, this page does not illustrate any information on the geography of backers and entrepreneurs. Instead, the information must be retrieved directly from the community project pages. The answers from Kickstarter, however, did not provide any information as to whether these data may be collected on a large scale.

Unsatisfied with the reply, a general research was conducted on the privacy regulations for websites. While there seems to be no uniform regulation that would cover this aspect, many websites use the “Robots Exclusion Protocol” as an approach to restrict the collection of information from their websites (Huang & Hoeber, 2010). The “Robots Exclusion Protocol” or “robots.txt” provides website operators the possibility to inform web-robots about which areas of the website should not be processed. Web-robots, also called web-crawlers or web-scrapers, can be described as software that browses through websites and searches for specific content or key words. Web robots are used, for example, by the major search engines (e.g., Google) to categorize websites. Website operators can place the protocol in the root of the website hierarchy and thereby

[54] The statistics page can be found on: www.kickstarter.com/help/stats

instruct robots to ignore specific files or directories, that may contain sensitive information.[55] The robots.txt for Kickstarter is illustrated in Figure 12.

Figure 13: Robots.txt for Kickstarter

This figure shows the robots.txt file for Kickstarter. The robots.txt specifies areas of the website that the website creators do not want to be searched by common search engines. The limitations in Kickstarter are mostly focused on backer related personal data (e.g., profiles, individual contributions).

```
User-Agent: *
Disallow: /projects/*/backers
Disallow: /projects/*/pledge/
Disallow: /projects/*/widget/
Disallow: /projects/*/widget

Disallow: /profile/
Disallow: /profiles/

Disallow: /v1/
```

While the Robots Exclusion Protocol is still mostly advisory, this thesis does respect the guidelines for the usage of web-crawlers. In this context, Kickstarter mainly limits the scraping of profile data. Kickstarter aims to protect the private information of its community members. This data, however, is not relevant for this thesis. All the necessary information can be found within the crowdfunding project description. The project description is not stated in Kickstarter's Robots Exclusion Protocol, nor the community section from which the country of origin is extracted. A possible explanation for this decision is that Kickstarter does not want to put on any limitations on the distribution of crowdfunding projects, as it is important for the success of campaigns to

[55] In July 2019 Google announced its endeavour to turn the Robots Exclusion Protocol (REP) into an official internet standard. The company also made its own robots.txt parser open source to increase its usage (cp. Google Webmaster Blog (2019))

be found and shared. Due to the described reasons, it is assumed that this thesis does not violate any privacy or intellectual property regulations.

3.7.4 Sample Description

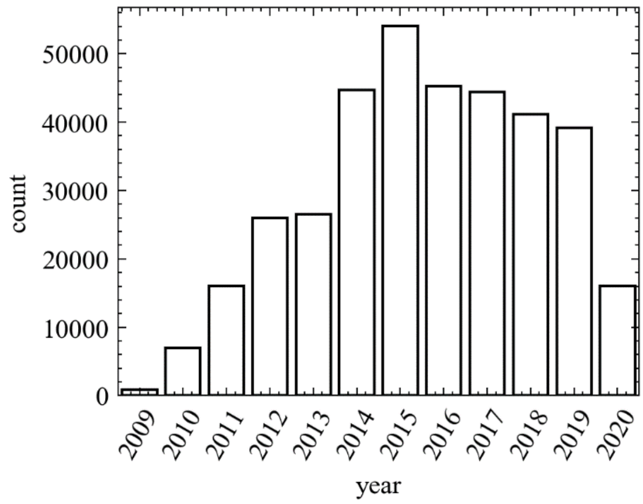
In total, the crawler extracted raw data on 361.278 random projects from April 2009 until June 2020. This is 74% of all published projects that were launched on Kickstarter since its start of operation to date (cp. Kickstarter, 2020). This is a satisfying number because not all projects remain available on the web after termination. Kickstarter regularly removes some of the published projects because of policy violations or legal reasons. The assumption is that the crawler collected all publicly available projects on the Kickstarter platform and extracted the relevant information as specified in section 3.4.4.

The project-launching dates range from April 2009 to June 2020 as seen in Figure 13.[56] The highest number of projects is from 2015.[57] From the obtained information, a pooled cross-sectional dataset is constructed that includes random projects of the Kickstarter project population at different points in time over the period of study. The observations are independent from each other. However, it cannot be assumed that they are identically distributed across the years in terms of project type or country of origin, since the overall population is subject to constant change. For example, the number of projects or participants can vary across different years, categories and countries.

[56] The oldest project is from April 2009, because this was the launching month of the Kickstarter platform.
[57] One possible explanation for the observed peak in this year is the crowdfunding campaign of the "Pebble Smartwatch". The collected record sum of US$20 million attracted wide attention and led to high media presence, which in turn attracted many new participants.

Figure 14: Number of Projects by Year

This figure illustrates the cumulative number of projects by year that was collected via the web-crawler. The highest number is in 2015, the lowest number is in 2009 as this is the launching date of the Kickstarter crowdfunding platform.



This thesis focuses on projects that carry the status of "successful" or "failed". Successful projects are crowdfunding endeavours that managed to reach the self-set funding target. A project that is labelled as "failed" does not necessarily imply that the project did not manage to raise any funds. Instead, it means that a project failed to reach the pre-set funding target and that, due to Kickstarter's "all-or-nothing" policy, the collected funds are entirely returned to the backers.

In practice, it is therefore possible that a project which is labelled as "failed" has raised a significant amount of funding. An extreme example is the project by "CENTR Camera" that raised more than 600,000€ from more than 2,000 backers but was labelled as "failed" because it was not able to reach the target value of 800,000€.[58] Despite the misleading label, "failed" projects provide valuable data on backers and are therefore

[58] CENTR Camera Project can be found on: https://www.kickstarter.com/projects/1307511016/centr-interactive-panoramic-video-in-the-palm-of-y/description

included in the dataset. In contrast to “successful” and “failed” projects, projects with the status “live”, “cancelled” or “suspended” do not provide added value to this analysis and are therefore eliminated from the dataset. “Live” projects only demonstrate a snapshot of the fundraising process and do not capture the result. Similarly, “cancelled” and “suspended” projects are discontinued projects that were either stopped by the owner or the platform operators prior to the deadline due to errors or rule violations.

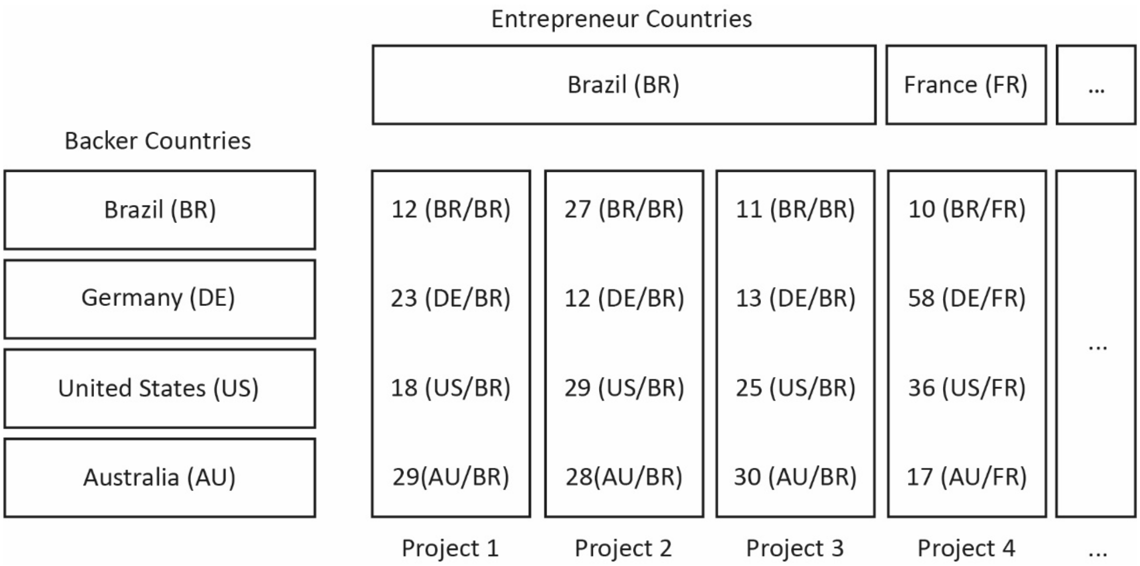
From the initial raw dataset of 361.278 projects, 149,583 projects (41%) do not contain the relevant statistics on the backers’ countries of origin. These are mostly old projects that were launched before Kickstarter started to collect geographical data, or projects that did not manage to raise any funds. Consequently, these projects are deleted from the dataset, leaving a usable data sample of 211,695 “successful” and “failed” projects. This is 59% of the initial dataset and 44% of the published Kickstarter data.

Although the term "Big Data" is not clearly defined in the literature, the dataset of this thesis qualifies as being designated as such. Overall, the 211,695 projects from 176 countries provide 44,653,657 distinct backing actions from investors across 215 different countries. Despite the described exclusions, this thesis uses the largest, most international and most recent data sample in crowdfunding research to date. Comparable studies rarely exceed samples of more than 100,000 crowdfunding projects (cp. Table 2, p. 65). The term "Big Data" is often used for data sets where the computational and analytical effort exceeds the capacities of ordinary computers and software (Govindaraju et al., 2015). During the model calculation, this thesis repeatedly faced the performance constraints of commercially available computers and was limited by them.

The data sample is transformed according to the level of analysis described in section 3.4.1. The individual backing actions are aggregated to the level of project-specific country-country tuples, obtaining 1,118,654 distinct observations. In other words, each observation describes the characteristics of a specific project and the aggregated count of backers that the project received from a specific country (dependent variable). Figure 14 visualizes the combination structure.

Figure 15: Country Tuples

This Figure explains the composition of the project-specific country-country tuples. For each of the 211,695 projects from 176 different entrepreneur countries, the unique backer/entrepreneur-country combination is used as a data record. For example, 23 backers from Germany invested in "Project 1" from Brazil. In total, the data sample contains 1,118,654 distinct project-specific country-country tuples.



It is important to note that Kickstarter only provides information on the ten biggest backing countries per project. Therefore, the collected data does not reflect all occurred transactions. Another important characteristic of the data sample is that it only incorporates constellations where at least one backer invested into a project. It does not, therefore, consider all possible country-country constellations that could have occurred. For example, if German backers actively decide not to invest into a Brazilian crowdfunding project, because of its distance or other reasons, this scenario is not captured by the data sample as it only incorporates positive observations. This characteristic of the data sample introduces a potential bias and can distort the results. One possible solution to deal with this problem is to use a *zero-inflated* data sample instead. The idea behind the zero-inflated data sample is to consider potential backing actions that did not occur between two countries. This, for example, can be achieved by identifying all possible pair-wise combinations between entrepreneur and backer countries and include them into the dataset with a value of zero (M. Lin & Viswanathan, 2016). Extending the example above, the fact that no German backer invested into a Brazilian project would be considered by adding an additional observation with the value of zero as the dependent variable.

However, using a zero-inflated data sample introduces several problems. The first major problem is that Kickstarter does not provide any information on the composition of its community. Therefore, it is difficult to estimate which countries should be included in the generation of country-pairs. Another major problem is that it is not possible to distinguish between instances where backers consciously decided not to fund a project and instances where projects were simply not seen. Kickstarter is one of the biggest crowdfunding communities and hosts simultaneously thousands of projects. The disadvantage of the large community is that it cannot provide equal visibility to all projects.

Moreover, the incorporation of non-observed transactions would produce a plethora of “excess” zeros that would not only limit the model’s performance but also not adequately represent the reality. In many cases, the model would wrongly assume that backers from a specific country decided not to back a project although the reason for the zero value might be the lack of visibility. Another arising problem that should not be underestimated is the computational feasibility. A simulation showed that, by including all possible country combinations, the dataset would inflate to more than 50 million different observations (on the project level). The required computational power to calculate the model on such data exceeds the power of most commercially available computers.

Moreover, to deal with the problem of excess zeros in zero-inflated data samples, some studies suggest to use a zero-inflated Negative Binomial regression or a hurdle model (Agresti, 2002, 2007; Long, 1997). These models calculate two phased equations, one for the excess zeros and one for the count model, making the overall model increasingly complex.

After careful consideration of different possibilities, this thesis adopts the non-zero-inflated approach because the considerable additional effort and complexity required for the zero-inflated model is not in proportion to its benefits. Moreover, the approach of using non-zero-inflated data is in accordance with most of the prior research on crowdfunding (Burtch et al., 2013; M. Lin, Prabhala, & Viswanathan, 2013; M. Lin & Viswanathan, 2016).

3.7.5 Sample Descriptive Statistics

The correct interpretation of the results requires a profound understanding of the underlying data (Datar & Garg, 2019; Jeffers, 1994; N. J. Salkind, 2010). Descriptive statistics and graphical analysis can help scientists to simplify large amounts of data and facilitate the understanding of the research context. This, in turn, can help them to disclose potential biases (Keren & Lewis, 1993; Tukey, 1977). Table 10 provides the summary statistics for the metrical variables in the data sample.

Table 10: Summary Statistics of the Data Sample

This table shows the summary statistics for the metrical variables of the data sample. In total, the sample consists of 1,118,654 distinct observations. For illustration purposes, and due to limited informative value, dummy variables are excluded from the table.

Variable	Mean	w. Mean	Median	Std. Dev.	Min.	Max.
Count of backers	29	-	2	368	0	152,836
Distance	5,964	3,460	5,914	5,239	0	19,854
GDP-B	44,941	53,683	45,032	16,003	1,563	185,829
GDP-E	52,812	53,352	59,532	12,127	5,420	104,103

The summary statistics for the dependent variable (count of backers) show that the mean and median are distinctively closer to the minimum than to the maximum value. This pattern indicates a right-skewed distribution of the dependent variable where a large proportion of observations has a count of less than 29 backers and few observations with particularly high counts. This type of distribution is common for count variables (Cameron & Trivedi, 2010).

The mean distance across all country-country tuples is 5,964 km. Another interesting statistic is the weighted mean for distance because it considers the count of backers per observation and thereby allows an easier comparison with the average distance estimates of other studies (Agrawal et al., 2013, see also; Guo et al., 2018). The weighted mean can be calculated in two steps. First, the count of backers (per project and country) is multiplied by distance (to the entrepreneur). Second, the sum of the products is divided by the sum of all backers. The result is a weighted mean of 3,460 km. This number is higher than the average distance of 2,877 km reported by Guo et al. (2018) for data from Kickstarter platform.[59] It is also considerably higher than the reported value by Sorenson and Stuart (2005) who show that traditional VC investors typically invest into ventures within a radius of 70 miles (113 km) (see section 1.2). The high average distance suggests that Home Bias might be less pronounced on Kickstarter than in traditional (offline) financial markets. One possible explanation for this observation is that Kickstarter, due to its digital nature, significantly reduces different forms of transaction costs and, thereby, facilitates cross-border investments. Another possible explanation might also be that Kickstarter introduces new mechanisms to assess the quality of business projects from a distance, such as new forms of reputation signalling, crowd-based trust mechanism and new trustworthy intermediaries (see section 2.2.2).

[59] The difference might be due to the fact that this thesis employs a larger and more recent data sample for the estimation.

3.7.6 Exploratory Data Analysis of Backer Origin Distribution

Exploratory Data Analysis (EDA) is an approach of exploring and visualizing data that often goes beyond formal modelling or hypothesis testing (Tukey, 1977).[60] By organizing, plotting and summarizing information, EDA helps to make sense of data, to identify potential relationships and form initial assumptions about causalities. This step is an important foundation for further analysis as it can reveal potential biases and problems in model formulation (Datar & Garg, 2019; Roth, 1998). According to Tukey (1977), EDA is not a set of concrete techniques but rather an approach or philosophy that resembles "detective" work looking for patterns, anomalies and new insights (Keren & Lewis, 1993; Tukey, 1977). Tukey states that "it forces us to notice what we never expected to see" (Tukey 1977, v-viii, 1-3).

To explore potential preliminary patterns of the effect of distance on backers' actions within the data sample, individual countries are inspected for their overall backer origin composition. For this purpose, the data sample is aggregated to the country level. In total, the sample contains projects from 176 countries that received funding from 215 different backing countries. The count of backers from each backing country is calculated as the percentage share of the total count of backers that invested into a given project country.

Since the comparison of all 176 countries would be extensive, this EDA analysis focuses only on three sub-samples. First, the distribution of backing countries is compared for the five largest crowdfunding nations on Kickstarter (United States, United Kingdom, Canada, Australia and Germany). This comparison is interesting because these countries represent 87% of all projects created on Kickstarter. Second, the distribution is compared for five randomly selected countries from different continents (Mexico, Netherlands, Singapore, New Zealand, Japan). This comparison is important because it

[60] EDA is different from initial data analysis (IDA), which focuses rather on the suitability of the data required for model fitting, hypothesis testing and handling of missing values.

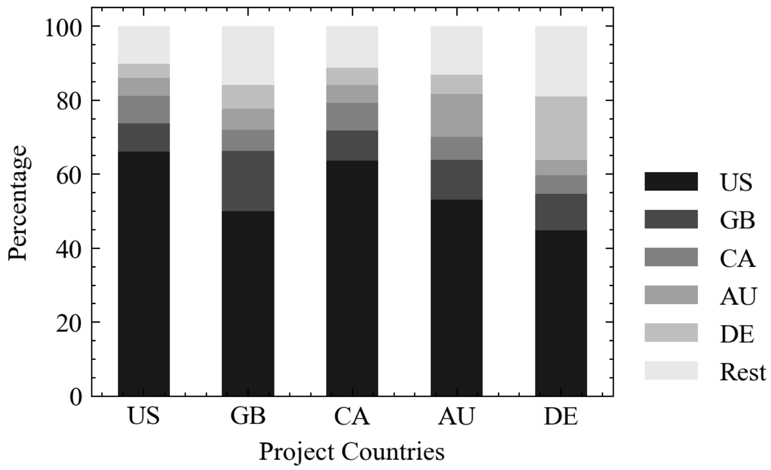
could reveal potential differences in the backer origin distribution across different geographical regions. Third, the distribution of backing countries is compared for the BRICS (Brazil, Russia, India, China, South Africa). The BRICS are commonly referred to as the countries with the highest economic potential and are frequently considered as one interesting group of analysis when comparing the performance of different countries (Stuenkel, 2013).

In general, this analysis might reveal three different potential outcomes for the geographical distribution patterns of backers. First, a distinct pattern might become apparent in which most backers are from the same or comparably close countries. This pattern would support the assumption of a strong Home Bias in crowdfunding. Second, a possible outcome might also be that most backers are always from the same group of countries, independent of the location of the entrepreneur. This outcome would speak in favour of a reduced role of Home Bias and highlight the importance of Kickstarter's community composition. Third, it might be that no distinct pattern is visible, and that crowdfunding follows entirely random, undiscovered dynamics. However, this would also speak in favour of a reduced role of Home Bias.

Figure 15 shows a stacked bar chart for the five largest crowdfunding nations. The x-axis shows the most frequent project countries, the US, the UK, Canada, Australia, and Germany (subsequently referred to as the "top-5"), while the y-axis shows the relative origin distribution of backers. A striking pattern is that all five countries demonstrate a similar distribution of backers. Approximately 80-90% of all backers are from the top-5 countries themselves.

Figure 16: Backer Origin Distribution for the Top-5 countries

This figure shows the backer origin distribution for the five most active crowdfunding countries. The x-axis shows the entrepreneur countries, whereas the bars summarize the origin composition of backers in percent. The figure shows that the largest count of backers is consistently received from the US.



Another striking pattern is that the US consistently account for 50-60% of all backers. But also the other countries (GB, CA, AU, DE) demonstrate a consistent share. Home backers, meaning backers from the same country as the entrepreneur, typically accounts for 10-15% of all backers (except for the US). Overall, the assumption of Home Bias that people tend to favour geographically close interaction partners is not supported by the chart. For example, Germany receives significantly more backers from the US than from the UK, although the distance to the US is significantly larger.[61]

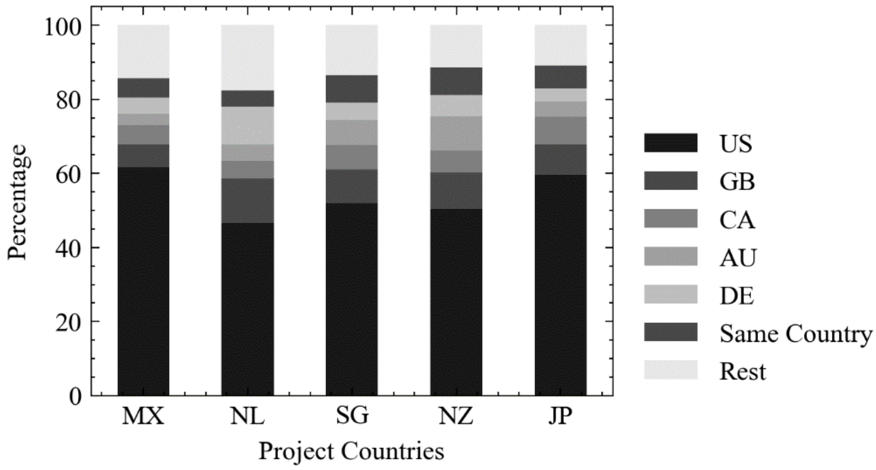
Figure 16 illustrates the backer origin distribution for five random countries from different continents. Overall, the figure shows a similar pattern to Figure 15. Approximately 80% of all backers are received from the top-5 countries. An interesting observation is that projects from New Zealand receive on average more backers from

[61] cp. Germany's distance to the US is 6,700 km, whereas Germany's distance to the UK is 930 km.

the US and the UK than from Australia, although Australia is considerably closer. The assumption of the Home Bias theory, again, is not supported by this graphical analysis.

Figure 17: Backer Origin Distribution for random countries

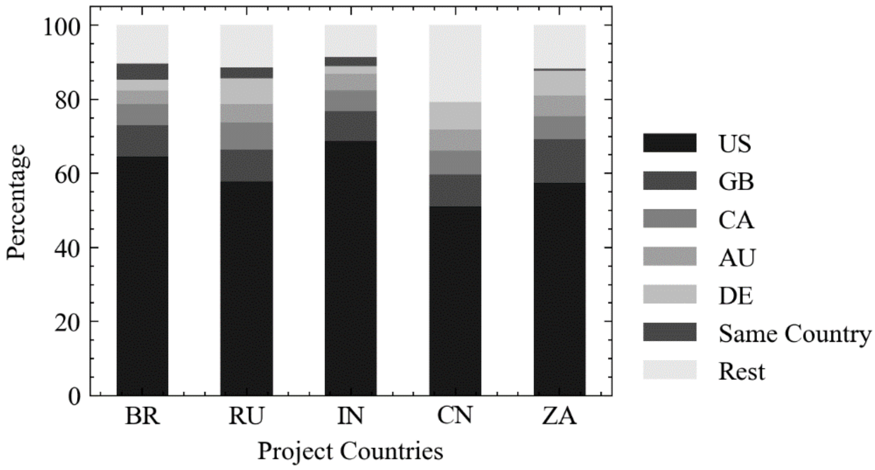
This figure shows the backer origin distribution for five randomly selected countries across the globe. The x-axis shows the entrepreneur countries, whereas the bars summarize the composition of backers in percent.



In Figure 17, the backer origin distribution is analysed for the BRICS countries. A striking pattern is that the share of US backers remains consistently high. Crowdfunding projects located in the BRICS countries receive on average more than 50% of the backers from the US. Overall, the top-5 countries account for approximately 80% of all backing actions. This pattern is consistent with the patterns found in Figure 15 and Figure 16.

Figure 18: Backer Origin Distribution for the BRICS countries

This figure shows the backer origin distribution for the BRICS countries. The x-axis shows the entrepreneur countries, whereas the bars summarize the composition of backers. The figure shows a reduced role of home backers in crowdfunding projects. Almost 80% of the pledges is received from the US, the UK, Canada, Australia and Germany.



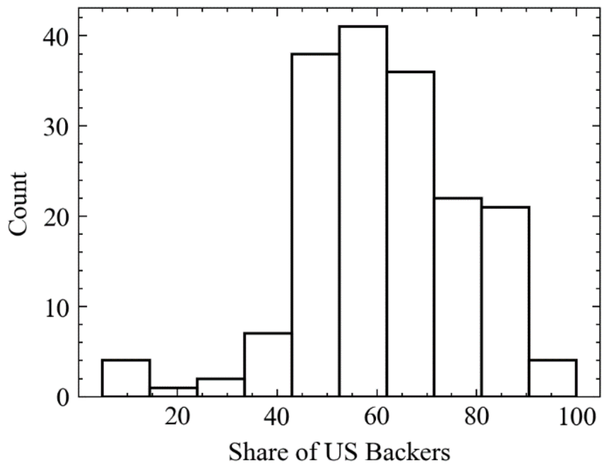
With the intention to further explore the role of the US, the average share of the US backers is calculated for all 176 project countries in the dataset as seen in the histogram in Figure 18. The x-axis shows the different share values, whereas the y-axis shows its frequency. The histogram suggests that the US account most frequently (in more than 40 countries) for 50-60% of all backing actions. The histogram is consistent with the previous figures (Figure 15, Figure 16 and Figure 17) and reveals the dominant role of the US.

Overall, the EDA reveals some preliminary patterns in the data that are not in line with the assumptions of the Home Bias theory. The Home Bias theory states that in business finance, most of the funds are received from geographically proximate investors (cp. section 2.2). The EDA, however, shows that crowdfunding projects seem to receive most of their funds from the same group of countries (US, GB, CA, AU, DE), independent of the entrepreneurs' country of origin. For example, projects from India receive on average as many backers from the US as projects from Canada, although the geographical distance between the countries is clearly different. This is an interesting preliminary pattern that needs to be confirmed by further in-depth analysis using the Negative Binomial regression (as presented in Eq. 9).

Another important pattern revealed by the EDA is the dominant role of US backers. This pattern might reveal a potential bias in which the investment behaviour of American backers overshadows the behaviour of backers from other countries (due to its high number). As a result, the discovered findings on Home Bias in international reward-based crowdfunding might apply for US backers but might lack validity for the behaviour of backers from other countries.[62] To address the potential bias that might occur due to the high dominance of the US in the data, this thesis implements an additional robustness test for the results of the main model estimation in which US backers and entrepreneurs are excluded from the sample.

Figure 19: Share of US Backers

This histogram describes the share of US backers for all 176 countries. It shows that most frequently the US account for 50-60% of all backing activities on Kickstarter.



[62] This bias is admitted by Guo et al. (2018) for their research on Home Bias in crowdfunding. However, the authors do not provide a robustness test.

3.8 Summary on Methodology and Data

Chapter Three explains the research approach of this thesis in detail. After discussing Pragmatism as the main paradigm adopted for this research. the chapter develops the Negative Binomial regression model to study the influence of geographical distance (and other control variables) on crowdfunding performance and justifies its application.

The chapter also dedicates particular attention to describing potential caveats of large datasets in quantitative analysis, revealing the “P-value Problem” as an important bias that is well-described in the statistics literature but vastly ignored in crowdfunding studies. In this context, this chapter describes the different remedies suggested by the literature to deal with Big Data that are implemented in the subsequent quantitative analysis.

Preliminary insights on Home Bias in international reward-based crowdfunding are provided via exploratory data analysis (EDA) conducted in this chapter. By visualizing the average backer origin distribution for individual countries, this preliminary analysis shows that crowdfunding projects launched on Kickstarter seem to receive most of their backers from the same five countries (US, UK, AU, CA and DE), independent of their country of origin. This pattern suggests that reward-based crowdfunding is less prone to Home Bias than traditional fundraising. However, the EDA considers only individual countries and is insufficient to derive generally valid conclusions. Therefore, the following chapter (Chapter Four) presents an extensive quantitative analysis of Home Bias in international reward-based crowdfunding via Negative Binomial regression.

Chapter Four: Findings and Analysis

4.1 Introduction

The aim of this thesis is to examine the existence of Home Bias in the emerging industry of international reward-based crowdfunding. For this purpose, a Negative Binomial regression model has been developed (as specified in Eq. 9) that examines how the count of backers in international reward-based crowdfunding (dependent variable) is affected by different influencing factors, such as geographical distance (D), GDP per capita (GDP), third-party endorsements (PWL), project category (C), herding behaviour (LP), and the Covid-19 pandemic (P).[63]

Equation 9: Final Model Equation

$$Log\left(NumB_{ij}\right) = \beta_0 + \beta_1 D_{ij} + \beta_2 GDP_i + \beta_3 GDP_j + \beta_4 C_k + \ \beta_5 PWL + \ \beta_6 LP + \ \beta_6 P$$

This chapter presents the findings of the quantitative analysis of the Negative Binomial regression. Besides examining the existence of Home Bias in international reward-based crowdfunding, the model also sheds more light on the potential influence of other factors on the success of crowdfunding campaigns.

The results of the quantitative analysis are critically discussed for each model variable, taking into account the relevant literature and potential implications of the findings. A further robustness test is conducted to validate the results and control for any potential biases that could have arisen from the high dominance of US participants in the dataset.

[63] A detailed overview of the different variables and their definitions can be found in Table 3, p. 99.

4.2 Interpretation Approach of Model Results

The results of the Negative Binomial regression are interpreted according to the approach described in section 3.5 as follows:

- First, the coefficients are transformed into Incidence Rate Ratios (IRR) through exponentiation (as described in Eq. 10). The reason for this transformation is that in Negative Binomial regression coefficients of variables are on the natural log scale, which makes them difficult to interpret. By transforming the coefficients into IRRs, they can conveniently be interpreted as the multiplicative effect. The IRR describes the relative change on the scale of the dependent variable (count of backers) for a one-unit increase of the respective independent variable.
- Second, the IRRs are inspected for their influencing directions and statistical significance in terms of p-values. The influencing direction of a variable on the count of backers can be determined by observing whether the IRR is larger or lower than 1. IRRs larger than 1 indicate a positive relationship in which an increase of the independent variable (e.g., distance) leads to an increase of the dependent variable (count of backers). Accordingly, IRRs lower than 1 indicate a negative or inverse relationship in which an increase in the independent variable leads to the reduction of the dependent variable (Hornuf & Schwienbacher, 2018). P-values indicate how likely it is that the observed relationship pattern between variables appeared by chance. This thesis reports the statistical significance of the different variables at the levels of <0.01, <0.001 and <0.0001 which is common in crowdfunding research (cp. Burtch et al., 2014; Guo et al., 2018; Singh Chawla, 2017) and the predefined standard of the statistical software programmes for model calculation used in this thesis (R's Mass package and Python's Statsmodels library) (cp. Zeileis et al., 2008). However, to account for the large sample size and with the intention to reduce the probability of attaining false positives, this thesis adopts the recommendations in the literature and primarily uses the lowest threshold of $p < 0.0001$ to determine whether an

observed relationship pattern is *statistically significant* (see also Singh Chawla, 2017).

- Third, the variables are inspected for their practical effect sizes. A recognized problem of Big Data models is that they have the statistical power to identify marginally small patterns in the data (Chatfield, 1995). These patterns might be statistically significant but marginally small to be relevant in practice (Lin et al., 2013). This thesis uses a sample of more than 1 million observations and is, thereby, particularly prone to the p-value problem. This means that some variables in Equation 9 might appear statistically significant (at $p < 0.0001$) due to the large sample size, although their practical relevance might not be significant. Therefore, to control for the effect of Big Data, this thesis uses the Marginal Effect at the Means (MEM) and Cohen's d to check the practical effect size of variables in Equation 9.

  The MEM is obtained by using Equation 9 to generate two predictions for the count of backers. This includes one reference prediction (where variables are set to mean or most common values) and one prediction where the variable of interest is changed by one unit while all other variables are held constant. For example, to obtain the MEM for distance, it is first necessary to calculate a reference prediction. For this purpose, this thesis sets the continuous variables (distance and GDP per capita) in Equation 9 to their respective mean values (as reported in Table 10, p. 155), dummy variables (PWL badge, Large Project and Covid-19) to 0 (i.e., characteristic is not present) and categorical variables (Project Category) to the most frequent values (i.e., "Games" as shown in Table 4, p. 109). To obtain the MEM for distance, the distance variable is then changed by one unit (1,000 km) while all other independent variables are held constant at the described values. The estimated difference in the dependent variable (count of backers) between the two predictions is the marginal effect size for distance.

  Cohen's d values are estimated by calculating the mean difference between two groups, then dividing the result by the standard deviation (explained in detail in section 3.5.3). For categorical or dummy variables (i.e.., PWL badge, Large Project, Covid-19 and Project Category), the difference in predictions is measured for the different characteristics (e.g., the average change in the count of backers for projects that show the PWA badge compared to project that do not show the PWA badge) and then divided by the standard deviation (of projects that do not show the PWA badge) (cp. Coxe, 2018).

For continuous variables (e.g., distance and GDP per Capita), the difference in predictions is used for a one-unit change in the predictor (e.g., the change in the count of backers for a 1,000 km increase in distance) and then divided by the standard deviation.[64]

4.3 Negative Binomial Regression Results

The model estimation results (Table 11) and the corresponding effect size analyses (Table 12) for the different variables are presented in a consecutive manner because it allows the simultaneous assessment of both statistical and practical significance. It is discussed in Chapter 3 that statistical significance alone (measured by p-values) is insufficient in Big Data models to determine that a discovered relationship pattern is also meaningful in practice (cp. section 3.5.1). This is because Big Data models possess the power to identify marginally small patterns in the data that might be irrelevant in practice. By considering both the statistical significance and the different effect size measures, a much better assessment of the true relevance of a discovered relationship pattern can be made. This approach is maintained throughout this chapter and each variable is discussed with reference to Tables 11 and 12 in a consecutive manner.

Table 11 shows the model estimation results. It lists all the independent variables as described in Equation 9 (Column 1) and shows their coefficients (Column 2) and IRR

[64] Coxe (2018) offers an internet-based tool to estimate Cohen's d value for nonlinear count regression models at: https://stefany.shinyapps.io/RcountD/.

values, including their statistical significance (Column 3). Moreover, the 95% confidence interval for each variable is reported in Column 4 and 5.

Table 12 presents the estimated effect sizes for the same variables. Column 2 restates the IRR (as presented in Table 11) whereas Column 3 shows the MEM for each independent variable. Column 4 presents the Cohen’s d values and Column 5 the respective interpretation suggestions that can be found in the literature (Coxe, 2018; Poldrack, 2018; Lakens, 2013).

Table 11: Results of Negative Binomial Regression

*This table shows the results of the Negative Binomial Regression which describes the effect of different variables (Column 1) on the count of backers (dependent variable) that supported a specific crowdfunding project from a specific country. Columns 2-5 show the coefficients, IRRs and 95% confidence intervals for the different variables. The IRRs are obtained from the coefficients through exponentiation (cp. Eq. 10). They can be interpreted as the multiplicative effect for a one-unit increase of the respective variable. In this context, all IRRs < 1 indicate a negative effect relationship between dependent and independent variables. Accordingly, IRRs > 1 reveal a positive relationship. The symbols *, **, and *** denote statistical significance of the respective variables at a p-value of 0.01, 0.001, and <0.0001, respectively (R software default settings). The fourth and fifth column present the lower (2.5%) and upper limit (97.5%) of the 95% confidence interval for the IRRs.*

Dependent Variable = Count of Backers				
Independent Variable	**Coef.**	**IRR**	**Conf. Limit 2.5%**	**Conf. Limit 97.5%**
Distance	-0.04	0.96***	0.96	0.96
GDP-B	0.03	1.03***	1.03	1.03
GDP-E	-0.01	0.99***	0.99	0.99
PWL	0.64	1.89***	1.88	1.91
Large Project	2.43	11.41***	11.32	11.47
Covid-19 Pandemic	0.21	1.24***	1.21	1.27
Category = Art	-0.87	0.42***	0.41	0.42
Category = Comics	-0.94	0.39***	0.38	0.39
Category = Crafts	-0.97	0.38***	0.37	0.39
Category = Dance	-1.24	0.29***	0.28	0.30
Category = Design	-0.01	0.99	0.98	1.00
Category = Fashion	-0.65	0.52***	0.51	0.53
Category = Film & Video	-0.83	0.44***	0.43	0.44
Category = Food	-1.11	0.33***	0.32	0.34
Category = Journalism	-0.89	0.41***	0.39	0.42
Category = Music	-1.17	0.31***	0.31	0.32
Category = Photography	-0.95	0.38***	0.38	0.39
Category = Publishing	-1.00	0.37***	0.36	0.37
Category = Technology	-0.14	0.87***	0.86	0.88
Category = Theatre	-1.22	0.29***	0.29	0.30

Table 12: Effect Size Measures for the Independent Variables

This table is an extension to table 11, investigating the practical effect sizes of the independent variables by using different measures such as IRR, MEM and Cohen's d. Column five provides the respective interpretation suggestion for Cohen's d values as described in the specific literature (cp. Table 9).

Dependent Variable = Count of Backers				
Independent Variable	**IRR**	**MEM**	**D**	**Effect Size**
Distance	0.96	-0.08	-0.02	Very Small
GDP-B	1.02	0.07	0.02	Very Small
GDP-E	0.99	-0.03	-0.01	Very Small
PWL	1.89	1.84	0.54	Medium
Large Project	11.40	21.40	6.34	Very Large
Covid-19 Pandemic	1.24	0.49	0.15	Very Small
Category = Art	0.42	-1.20	-0.35	Small
Category = Comics	0.39	-1.26	-0.37	Small
Category = Crafts	0.38	-1.28	-0.38	Small
Category = Dance	0.29	-1.46	-0.43	Small
Category = Design	0.99	-0.02	-0.01	Very Small
Category = Fashion	0.52	-0.99	-0.29	Small
Category = Film & Video	0.44	-1.16	-0.34	Small
Category = Food	0.33	-1.38	-0.41	Small
Category = Journalism	0.41	-1.22	-0.36	Small
Category = Music	0.31	-1.42	-0.42	Small
Category = Photography	0.38	-1.27	-0.37	Small
Category = Publishing	0.37	-1.30	-0.39	Small
Category = Technology	0.87	-0.28	-0.08	Very Small
Category = Theatre	0.29	-1.45	-0.43	Small

Figure 19 provides a summary of the main findings from the quantitative analysis. The consecutive sections in this chapter discuss the model results for each variable and their implications and justify the core findings in Figure 19.

Figure 20: Overview of Core Findings of This Thesis

Research Question

Does geographical distance between backers and entrepreneurs affect the overall count of project supporters in international reward-based crowdfunding?

Core Findings

In crowdfunding research, more attention must be paid to distinguishing between statistical significance and practical relevance of a discovered relationship pattern (due to the p-value problem in Big Data).

●

Although a statistically significant negative influence of geographical distance on the count of backers can be identified, its effect size is rather small and, therefore, likely to be negligible in practice. This suggests a small relevance of Home Bias in international reward-based crowdfunding.

●

Although a statistically significant influence of Project Category and GDP per capita (of backers and entrepreneurs home countries) on the count of backers can be identified, the effect size is rather small and likely to be negligible in practice. This finding questions the results of previous studies.

●

Herding behaviour and third-party endorsements are found to have a statistically significant and practically relevant impact on the count of backers. Therefore, they demonstrate effective quality signals that can positively impact the outcome of crowdfunding projects.

●

Although a statistically significant positive influence of the Covid-19 pandemic on the count of backers can be identified, the effect size is rather small and likely to be negligible in practice. This finding suggests that crowdfunding could demonstrate a reliable source for capital during global economic crises.

●

The exclusion of US participants from the data sample does not seem to affect the results of the quantitative analysis. This suggests that US participants do not bias the findings of this thesis.

4.3.1 Distance

Distance is the key variable of interest in this thesis. By examining the potential effect of geographical distance between backers and entrepreneurs on the count of backing activities, this thesis aims to scrutinize whether the Home Bias phenomenon persists in online investment decisions in international reward-based crowdfunding.

The model estimation results in Table 11 show that *Distance* is statistically significant at $p < 0.0001$. The IRR of 0.96 suggests that geographical distance has a negative influence on the count of backers in international reward-based crowdfunding campaigns. A one-unite increase in distance (1,000 km) leads to a decrease of the count of backers by approximately 4% (1 – 0.96 = 0.04). This finding is supported by the considerably narrow 95% confidence interval of 0.96 to 0.96.

Overall, this result suggests that Home Bias exists in international reward-based crowdfunding as projects tend to receive less backers from distant countries. This parallels the findings of Burtch et al. (2014) who show that distance has a statistically significant negative effect on the count of international lending actions on Kiva.org, a social-lending crowdfunding platform. It is also in accordance with some of the previous research on Home Bias in other crowdfunding models (see Agrawal et al., 2013; Mollick, 2013; Lin & Viswanathan, 2016; Cumming et al., 2019; Guo et al., 2018).

However, because of the Big Data problem that previous crowdfunding literature has largely ignored, this finding must be interpreted with caution because the low p-value could most likely be an artefact of the large sample size (cp. section 3.5.1). Therefore, to assess the practical relevance of the influence of distance on the count of backers, the MEM and Cohen’s d are estimated as seen in Table 12.

According to MEM results in Column 3, holding all other independent variables constant at their means and increasing distance by one-unit (1,000 km) results in 0.08 fewer backers. This result suggests a comparably small influence of distance on the count of backers and relativises the practical significance of this variable. A further examination of the practical effect size of distance using Cohen’s d reveals a d value of 0.02 (cp. Table

12, Column 4). According to the interpretation approach suggested in the literature (Lakens, 2013), it can be concluded that distance has a "very small" effect on the count of backers. Hereby, both the MEM and Cohen's d value suggest that geographical distance has a small practical effect size on the count of backers.

This finding implies that although some pattern of Home Bias can be detected in international reward-based crowdfunding on Kickstarter (as seen in Table 11), its practical effect size is rather small and might be negligible in practice (cp. Table 12). A change in geographical distance between entrepreneurs and backers seems to influence the count of backers only marginally. This finding is in line with the initial observations made by the exploratory data analysis (EDA) conducted in Chapter 3 (cp. section 3.6.6).

The EDA supports the low practical relevance of distance by showing that crowdfunding projects tend to receive most of their backers from the same countries (most notably from US, UK, CA, AU and DE) independent of the entrepreneur's country of origin. For example, projects from India, South Africa and Japan demonstrate a similar backer origin distribution (cp. Figure 16), where the US account for most of the backing actions.

The presented findings are also in accordance with the literature on research associated with Big Data models. This thesis confirms the expressed concerns that p-values might be an unreliable measure for the significance of a variable in large samples (Cohen, 1995; Demidenko, 2016; Gelman & Stern, 2006; Harvey, 2014; Lin et al., 2013; Pharoah, 2007). This is in accordance with the CPS-charts (presented in section 3.5.2) which illustrate the problem that p-values tend to thrive towards zero with increasing sample size, making it easier for scientists to obtain statistically "significant" results according to the traditional interpretation approach of NHST.[65]

---

[65] The Negative Binomial regression model determines all variables to be statistically significant at $p < 1e^{-16}$, which is the lowest p-value threshold that is possible to measure on ordinary computers. This indicates that simply lowering the p-value threshold, as some authors suggest (cp. Lin et al., 2013), will not solve the p-value problem in large samples. Instead, scientists need to focus on different measures for effect sizes and evaluate whether the effect is large enough to be relevant for practice.

The finding that Home Bias plays a subordinate role in crowdfunding is supported by some prior studies (see Günther et al., 2018; Vulkan et al., 2016; Stevenson et al., 2019). For example, Mollick & Robb (2016) find that equity-based crowdfunding significantly relaxes geographic constraints, and that VC seed funding is considerably less concentrated. Agrawal et al. (2015) also conclude that Home Bias does not exist on SellaBand.com, a donation-based crowdfunding platform, when the influence of family and friends is controlled for. Stevenson et al. (2019) further show that reward-based crowdfunding in the US is geographically dispersed and non-discriminatory in terms of state or region.

However, these studies focus on individual countries (mostly US) or specific regions (such as the EU) and do not consider crowdfunding in a global context. The findings of this thesis are particularly interesting because they are based on the largest geographical scope and the most recent data that can be found in the literature. Furthermore, this thesis pays more attention to practical relevance (rather than statistical significance) by reporting two different measures for the effect size, namely MEM and Cohen's d.

Different explanations can be offered for the finding that Home Bias plays a subordinate role in international reward-based crowdfunding on Kickstarter. One possible explanation might be that crowdfunding provides more efficient mechanisms to reduce the information asymmetry problem between founders and backers (Ahlers et al 2015). The literature suggests three potential quality signals that might help backers to better deal with the information asymmetry problem (cp. section 2.3.2): (1) crowd-based trust mechanisms, (2) trustworthy intermediaries, and (3) reputation signalling through social networks. Thierer et al. (2015) suggest that especially crowd-based trust mechanisms provide a solution to the information asymmetry problem that many regulators consider insurmountable in offline markets. By displaying the engagement and comments of other backers in real-time for each project, crowdfunding platforms create more transparency and allow observational learning (Zhang & Liu, 2012).

Moreover, crowdfunding facilitates the incorporation of social networks. Backers can easily identify shared contacts and see which projects people in their social networks are backing (Kickstarter Support, 2020). The new possibilities provided by the internet contribute significantly to the reduction of the information asymmetry problem, which in turn, might give backers the confidence to invest into distant projects. This explanation parallels the findings of Mollick and Nanda (2016) who see crowd-based trust mechanisms as one possible explanation why backers are good at identifying fraudulent crowdfunding campaigns.

Another possible explanation for the reduced significance of distance found in this thesis is the standardization of projects. On Kickstarter, all projects share the same guidelines and appear homogeneous in their approach and layout. The entrepreneur's location is not particularly striking and can easily be overlooked in the project description. Moreover, Kickstarter limits the interaction possibilities for founders and backers in the sense that physical proximity should not provide any additional advantages. All interactions must be handled with the existing platform-based tools online. Thereby, crowdfunding platforms equalize transaction costs (H. Kim & Kim, 2017), which are considered a major driver for Home Bias in offline markets (Ahearne et al., 2004; Glassman & Riddick, 2001; Rowland, 1999).

The low relevance of geographical distance in this thesis might also be explained by the crowdfunding community culture itself. Prior research found evidence that crowdfunding participants are likely to have similar character traits such as open-mindedness, entrepreneurial spirit, and a high willingness to take risks (Gerber and Hui 2013). On Kickstarter, many backers are entrepreneurs themselves and have a fundamental interest in supporting entrepreneurship. This explanation aligns with the findings of Zvilichovsky et al. (2013) who show that reciprocal behaviour (meaning that former campaign creators become backers themselves) is an important component of the Kickstarter community culture (see also Kickstarter 2016b). Thereby, participants in crowdfunding might be generally less prone to Home Bias because they often act on both sides of the market themselves and know the difficulty of raising funds in the start-up phase.

Additionally, the low relevance of geographical distance could be because backers normally engage with comparably low amounts of funding, especially when compared to traditional fundraising (Florida & King, 2016). According to Kickstarter (2010), the average contribution of backers lies between US$50 and US$100, making the individual risk of each backer rather small.[66] The low funding volumes, that are typical for reward-based crowdfunding, might encourage backers to take higher risks and render the geographical distance to the project founders irrelevant.

Another possible explanation for the low relevance of geographical distance that applies particularly well to reward-based crowdfunding is provided by Steigenberger (2017). The author finds that many reward-based crowdfunding projects offer innovative and creative products that are unique in their design and that many backers derive their motivation for backing a certain project from the desire to pre-purchase these products that would otherwise be not available to them. Therefore, the lack of alternatives might lead backers to accepting higher risks and neglecting the influence of geographical distance.

The finding of this thesis that Home Bias plays a subordinate role in crowdfunding is substantial because it allows to derive important recommendations for different stakeholders. For example, the results presented in this thesis suggest that entrepreneurs in reward-based crowdfunding should expect backers from distant countries at least to the same extent as from nearby countries. This in turn has important consequences for the preparation, execution, and follow-up of crowdfunding campaigns.

[66] Unfortunately, a more recent estimate for the average contribution of backers could not be found. However, the average backer contribution that can be calculated from the dataset, by dividing all pledged by the total count of backing actions, provides a similar estimate of 98 USD.

While preparing and launching campaigns, entrepreneurs need to be aware that they are operating on a global stage. Products and supply chains, therefore, need to be designed appropriately to fit this international context. For example, it might be important for entrepreneurs to ensure that their intellectual property is protected outside of their home regions before engaging in crowdfunding. Moreover, before offering product rewards to international backers, entrepreneurs must consider the global shipping costs as well as potential import restrictions of other countries.[67]

The findings of this thesis also provide important recommendations for the design of crowdfunding campaigns. It is suggested that it can be a rewarding strategy to first inspect the backer composition of the respective crowdfunding platform (as conducted in section 3.6.6). For example, this thesis shows that on Kickstarter entrepreneurs should primarily optimise their campaigns for the US, independent of their country of origin. This is because US backers are much more strongly represented on Kickstarter than backers from other countries (cp. EDA in section 3.6.6).[68] Due to the identified low relevance of geographical distance, entrepreneurs can expect that most backers will come from the US. This insight contradicts some of the assumptions of the previous literature on Home Bias that most investors will come from geographically proximate regions (Hwang et al., 2019; Sorenson et al., 2016). Future research could explore how entrepreneurs might optimize their campaigns for US backers (e.g., the effect of choosing USD as the default currency, offering free shipping to the US, or using American magazines for testimonials).

For policymakers, the findings of this thesis are relevant because they suggest that crowdfunding offers new possibilities to deal with the Global Finance gap (cp. 1.1). Crowdfunding could help to reduce the supremacy of business hubs where capital is

[67] For example, some innovative food supplements developed in the US might not be allowed in the EU because of unauthorised ingredients (e.g., genetically modified organisms).
[68] The high number of US backers might result from the fact that Kickstarter was founded in the US and has been active on the market for significantly longer than on other markets (e.g., Germany).

often concentrated (cp. Florida & King, 2016). The new mechanisms of crowdfunding could increase the volume of financial transactions across geographic, linguistic, and cultural barriers (Ahlers et al. 2015, Mollick and Robb 2016). At the same time, governments could consider targeted support for different crowdfunding concepts (e.g., through legislation or tax reliefs) to promote the growth of the industry. For example, Klaes (2017) identifies several legislative barriers to cross-border transactions that prevent different crowdfunding platforms to operate across Europe. Eliminating some of these legislative barriers could promote a more equal distribution of capital, strengthen economic development, and contribute to higher global prosperity.

4.3.2 GDP per Capita

The *GDP-B* and *GDP-E* variables are used in this thesis to measure the potential effect of individual wealth (of backers and entrepreneurs) on the count of investment activities in crowdfunding projects. This analysis should help to examine whether reward-based crowdfunding contributes to a redistribution of financial resources from high-wealth to low-wealth countries.

Table 11 presents the results of the model estimation for the GDP-B and GDP-E variables. The IRR for GDP-B is 1.03. The IRR > 1 indicates a positive relationship that is statistically significant at a level of $p < 0.0001$. These initial results suggest that projects can except more backers from countries with high GDP per capita than from countries with low GDP per capita. However, the subsequent analysis of practical effect sizes questions the practical relevance of this finding. Both the MEM and Cohen's d value of 0.07 and 0.02, respectively (cp. Table 12), suggest that the influence of the GDP-B on the count of backers is very small and might be negligible in practice. For example, the MEM shows that for a one-unit increase in GDP-B (1,000 USD), the count of backers will increase by 0.07 (backers), which is a relatively small change. This finding suggests that although a statistically significant pattern can be detected, in which personal wealth has a positive influence on the likelihood of backers to support a project, the practical effect size is too small to be noticed in practice. Cohen (1994) states that differences in this magnitude ($d < 0.2$) are too subtle to be differentiated “by the naked eye”.

An inconsistency between statistical and practical significance can also be observed for the GDP-E variable, however, with a different influencing direction. The influence of GDP-E on the count of backers is statistically significant at the level of $p < 0.0001$, as can be seen in Table 11. The IRR < 1 suggests a negative relationship in which the increase of the GDP-E will result in a decrease of the dependent variable (count of backers). This suggests that entrepreneurs from poor countries should enjoy an advantage in raising funds through reward-based crowdfunding on Kickstarter as they tend to attract more backers.

Once again, the subsequent analysis of the practical effect sizes in terms of MEM and Cohen’s d of -0.03 and -0.01, respectively, suggests that the effect is rather small in practice (cp. Table 12). The MEM states that for a one-unit increase in GDP per capita (1,000 USD) the estimated count of backers will decrease by 0.03 (backers), a very small change. The effect size measured in terms of standard deviations (Cohen’s d) suggests that GDP-E has an even smaller influence on the dependent variable than the GDP-B variable.

The conclusion that can be derived from this analysis is that although the Negative Binomial regression model is able to detect a statistically significant negative influence of GDP-E and a statistically significant positive influence of GDP-B on the count of backers, the practical effect size of both variables is rather small and might be negligible in practice. This finding also confirms the results of the EDA conducted in section 3.6.6. The EDA (cp. Figure 15, 16 and 17, p. 159) reveals that most crowdfunding projects receive their support from predominantly five backing countries (US, UK, CA, AU and DE), independent of the geographical origin of the entrepreneur.

Consequently, the results of this thesis suggest that individual wealth of backers and entrepreneurs does not seem to have a noteworthy effect on the count of backers. Moreover, this finding questions the observation from previous studies that crowdfunding could promote investment flows from high-wealth individuals to projects of low-wealth individuals as found by Burtch et. al (2014) for Kiva.org, a social-lending crowdfunding platform, or described by the literature on *Impact Investing* (Kish & Fairbairn, 2018).

One possible explanation why the findings of this thesis differ from the findings of Burtch et al. (2014) is that the reward-based crowdfunding model, in contrast to social-lending, does not have its primary focus on prosocial activities. Davis et al. (2017) indicate that the performance of reward-based crowdfunding projects is often dependent on two factors: The perceived passion of the entrepreneur and the perceived product creativity, rather than the individual predicament of the entrepreneur.[69] Therefore, it might be irrelevant for backers whether the entrepreneur comes from a wealthy or a poor country, but what counts is the idea itself and the way the idea is communicated. This is an important implication for entrepreneurs because it suggests that every project has the same chances of raising funds – entrepreneurs from developing countries are as likely to receive funding as entrepreneurs from developed countries. Overall, this finding suggests that reward-based crowdfunding could help to distribute capital more fairly (based on actual quality of the project idea) and, potentially, alleviate the Global Finance gap.

Another possible explanation for the discovered low effect of GDP per capita on the count of backers might be that Kickstarter's community is largely composed of people from the middle- and upper classes of the societies of the respective countries and that the differences in individual-wealth are not as great as assumed by the GDP per capita estimates. Investing into crowdfunding projects requires a certain surplus in financial resources. Likewise, promoting a business idea on Kickstarter often presupposes some forms of initial investments (i.e., the development of a prototype of the product, creation of a company website or development of marketing materials). Therefore, the

[69] However, it cannot be ruled out that this criteria has changed during the Covid-19 pandemic, where many entrepreneurs used crowdfunding on Kickstarter specifically to ask for support during difficult times.

context of reward-based crowdfunding on Kickstarter could imply some form of natural selection in which only people who participate are those who have a certain level of income and do not require financial support to “survive” or improve their social status (as it is the case on Kiva.org, a prosocial crowdfunding platform). Accordingly, backers might not feel that they make a social impact when investing in business oriented crowdfunding projects in developing countries via Kickstarter.

4.3.3 PWL badge (Third-Party Endorsements)

The *PWL badge* is used to measure the potential influence of third-party endorsements on the count of backers in reward-based crowdfunding projects. This analysis should help scientists and practitioners to better understand if third-party endorsements demonstrate effective quality signals that can increase the success probability of crowdfunding projects and, therefore, should be pursued by entrepreneurs.

The model results show that the *PWL badge* has a statistically significant influence on the count of backers at $p < 0.0001$, as can be seen in Table 11. The multiplicative effect being > 1 indicates a positive influencing direction. This suggests that the third-party endorsements increase the expected count of backers by 89%.

The subsequent analysis of the practical effect sizes in terms of MEM and Cohen’s d supports this finding (as seen in Table 12). The MEM suggests that projects will receive 1.84 more backers from each country if they receive the PWL badge, holding all other independent variables constant at their means. Cohen’s d value of 0.53 suggests a “*medium*” effect size of the PWL badge on the count of backers. In direct comparison to other variables, the PWL badge has the second largest effect size after *Large Projects* (Herding Behaviour).

Overall, the results are conclusive and suggest that third-party endorsements do have a statistically significant and practically relevant effect on the count of backers, which is in accordance with previous literature. For example, Qiu (2013) argue that third-party endorsements are important quality signals for crowdfunding in the US and can lead to an increased confidence into a project.

Different possible explanations exist for the positive effect of third-party endorsements on the performance of crowdfunding projects. One explanation is that projects that demonstrate a certain form of third-party endorsements (e.g., display the PWL badge) are also simultaneously promoted more intensively by the platform operators. For example, these projects are more frequently displayed on the front page and featured in the weekly newsletters (Qiu, 2013) which increases their visibility to potential backers. Some platforms also dedicate special subsections on their websites to display their favourite projects (cp. Kickstarter's PWL subsection).[70] These actions increase the likelihood of a crowdfunding project to be seen and might, thereby, attract more backers.

Another possible explanation for the positive effect of third-party endorsements on the performance of crowdfunding projects might be associated with the described information asymmetry problem (cp. section 2.2.2). In a situation where information is limited about the project initiator, the PWL badge might serve as a valuable signal of quality and contribute to trust building (Spence, 1973). Backers might believe that projects which receive the PWL badge have been subject to more detailed examination and that Kickstarter would not promote any project that is not trustworthy. This explanation parallels the findings of Qiu (2013) who describes that being chosen by Kickstarter demonstrates a "seal of approval" (see also, Ackerberg, 2001).

However, in practice, the PWL badge must be interpreted with caution because crowdfunding platforms usually do not conduct any additional analysis to evaluate the trustworthiness of a project. On Kickstarter, the PWL badge is mostly awarded to projects that stand out in their creativity in campaign design or the product itself

[70] "Projects We Love" subsection can be found at https://www.kickstarter.com/discover/pwl

(Kickstarter, 2021). In some cases, Kickstarter awarded the PWL badge to projects that subsequently revealed themselves to be sham (cp. “Holus” project).

This thesis contributes to the advancement of knowledge by showing that third-party endorsements also have a significant influence on the performance of crowdfunding campaigns in an international context. Moreover, it is the first to describe the concrete change in the count of backers that an entrepreneur can expect by obtaining the PWL badge. These findings also have important implications for practice because they can help future entrepreneurs to better estimate whether it is a worthwhile strategy to invest effort and time in obtaining the PWL badge or other forms of third-party endorsements.

4.3.4 Large Project (Herding Behaviour)

The *Large Projects* dummy variable is constructed to approximate the potential influence of herding behaviour on the count of backers. In this context, each project that attracted more than 107 backers in total (the median for the total count of backers across the sample) is coded as 1 and otherwise as 0. The purpose of this variable is to examine whether projects that reached the threshold of 107 backers are more likely to attract additional support from other countries and are, therefore, subject to herding behaviour.

The *Large Projects* dummy variable is found to have a statistically significant effect on the count of backers at $p < 0.0001$ (cp. Table 11). The IRR of 11.41 (cp. Table 12) suggests that projects that reached the threshold of 107 backers attracted more than ten times as many additional backers from a given country than projects that did not reach this threshold.

The subsequent analysis of the practical effect sizes in terms of MEM and Cohen’s d supports this finding (cp. Table 12). The MEM analysis suggests that projects that reached the threshold can expect approximately 21 more backers, holding all other variables constant at their means. Cohen’s d value of 6.34 suggests that the effect of the variable is "*Very Large*". In direct comparison, the *Large Projects* variable has the greatest effect on the count of backers among all inspected variables in this thesis.

The finding that the *Large Projects* variable, which approximates the effect of herding behaviour, has a strong influence on the performance of crowdfunding projects parallels the findings of prior studies. Zhang and Liu (2012), for example, show that on Prosper.com, a major lending-based crowdfunding platform, backers often use lending decisions by other peers to infer the borrowers' creditworthiness. They find evidence that herding behaviour demonstrates an effective signal for quality, when additional information is scarce. Colombo et al. (2015) come to a similar conclusion in their US focused study on reward-based crowdfunding. The researchers find that early accumulation of backers encourages additional participation by others.

Different explanations can be provided as to why large projects tend to attract more backers. First, and most likely, the tendency of large projects to attract additional backers can be attributed to the reduction of the information asymmetry problem, as explained above. Backers engage in observational learning and use the behaviour of others to evaluate the quality of different projects (Zhang and Liu, 2012). This explanation also parallels the literature on the "wisdom of the crowds" (Clauss et al., 2018). Another possible explanation is that sorting algorithms on Kickstarter are responsible for higher visibility. Projects that gain certain momentum attracting backers are frequently featured on the front page and in the weekly newsletters on Kickstarter and might, eventually, attract even more backing actions.

This thesis makes an important contribution to knowledge because it is the first to introduce a potential threshold for the positive effect of herding behaviour. The findings of this thesis suggest that projects that reach a total number of 107 backers (across all countries) also significantly increase the likelihood of receiving additional backers from other countries. This is an important insight for future entrepreneurs engaging in crowdfunding because it provides them a clear threshold target if they seek to benefit from the positive effects of herding behaviour.

4.3.5 Covid-19 Pandemic

The purpose of the *Covid-19 Pandemic* dummy variable is to measure the potential impact of a global economic crisis on the crowdfunding industry. This analysis should help scientists, entrepreneurs, and policymakers to better understand whether reward-based crowdfunding is an effective alternative for business fundraising during economic grievances when access to traditional funding sources is limited (cp. Brunnermeier, 2009; Gorton, 2010).

The model estimation results show that Covid-19 has a statistically significant effect on the count of backers (at $p < 0.0001$), as can be seen in Table 11. The IRR of 1.24 indicates a positive relationship suggesting that projects which were launched during the Covid-19 outbreak received on average 24% more backers than projects that were launched before the global pandemic. This finding is interesting because it suggests that crowdfunding projects have benefited from the global crisis by attracting more backers on average than before the pandemic. However, the subsequent analysis of the practical effect sizes (cp. Table 12) suggests a comparably small difference in the count of backers for projects that were launched during and before the pandemic. Projects affected by the pandemic received only 0.49 more backers according to MEM results. This finding is confirmed by Cohen's d value of 0.15 which indicates a "Very Small" effect size. Overall, the results suggest that although a statistically significant positive effect of the Covid-19 pandemic on the count of backers can be detected, its practical influence is rather marginal and might not affect most entrepreneurs in their crowdfunding activities. This suggests that the low p-value of the variable is likely to be an artefact of the large sample size (cp. section 3.5.1).

The suggested statistical positive effect of the Covid-19 pandemic on backing actions, although having a small practical relevance, is particularly interesting because it does not align with some of the findings of previous research on the effect of global economic crises on funding allocation decisions. For example, Brunnermeier (2009) and Gorton (2010) suggest that adverse economic shocks usually lead to a more difficult access to capital because investors' risk tolerance decreases significantly. Consequently, the crowdfunding industry might be expected to have been negatively affected by the current economic crisis. Kickstarter, for example, provides evidence for this trend by reporting a significant decline in crowdfunding campaigns (by 25%) in April 2020 (Leland, 2020). This decline, however, appeared only in the first weeks after the official outbreak of the Covid-19 pandemic, perhaps because fear and uncertainty were at their highest (Hoffmann et al., 2013). After the initial shock, backers might have returned to crowdfunding and perhaps even given the industry new momentum.

One possible explanation for the rather positive influence of the pandemic on the count of backers (because IRR > 1) is provided by Grasso et al. (2021). The authors state that while, on the one hand, crises can exacerbate negative conditions, inequalities, and competition between groups for scarce resources, they can also often inspire individuals and social groups to engage in acts of solidarity, mutual support and help. Global economic crises can, therefore, promote cohesion. Kickstarter has actively supported this movement during the pandemic by offering free seven-day extensions to many crowdfunding campaigns and new tools to promote projects on social media (Hecht, 2020).

These findings are interesting for both theorists and practitioners because this thesis is the first to examine the potential effect of a global economic crisis on the crowdfunding industry. It helps entrepreneurs and policymakers to decide whether reward-based crowdfunding is an effective alternative for business fundraising during economic grievances. The presented evidence supports the assumption that reward-based crowdfunding might not be affected by economic crises in the same way as traditional models or fundraising channels (Brunnermeier, 2009; Gorton, 2010), perhaps because it emerged as a countermovement to reduce the dependence on established financiers which significantly limited the access to capital during the financial crisis in 2007/08. However, it must also be considered that the described findings can only be understood as “premature” findings because the underlying data only covers an early time frame of the pandemic (March - June 2020) which might explain the small practical relevance of this variable as measured by the effect size.

The effect of the pandemic might evolve differently at later stages, after financial markets and governments have absorbed the initial shock and found new ways to deal with the crisis. Hoffmann et al. (2013) find that investors' return expectations, risk tolerance and risk perception can change during different phases of economic crises. This might also affect the crowdfunding industry. Although further research is required at a later stage of the pandemic to fully understand its influence on the crowdfunding industry, this thesis provides a valuable first insight into the current developments.

4.3.6 Project Category

The purpose of the *Project Category* variable is to examine whether the different project types have a noteworthy effect on crowdfunding success (i.e., whether projects related to technology perform better or worse than projects related to photography). Overall, Kickstarter distinguishes between 14 different project categories (*Technology, Games, Design, Photography, Fashion, Comics, Journalism, Crafts, Art, Publishing, Film & Video, Dance, Music, Food* and *Theatre*).

The model estimates for the different project categories are in reference to the "Games" category, which is included in the intercept.[71] With the exception of "Design", all categories are statistically significant at $p < 0.0001$ and demonstrate a multiplicative effect of <1, which indicates a negative influence on the count of backers (as seen in Table 11). The "Design" category is not statistically different from the "Games" category, meaning that projects related to these two categories attract on average the highest count of backers. Projects related to "Dance" and "Theatre" attract on average the lowest count of backers. The IRR for "Dance" is 29%, this means that projects related to this category can expect 71% less backers on average than projects related to "Games".

[71] The "Games" category was chosen as the reference group because it is the most common project category in the dataset (cp. Table 4).

The subsequent analysis of the effect sizes in terms of MEM and Cohen's d suggests a rather small practical influence of the project category on the count of backers (cp. Table 12). The MEM values indicate that most projects receive on average 1 to 1.5 backers less than projects related to the "Games" category. The Cohen's d values also indicate a rather "small" effect for most of the project categories. This implies that although some performance differences between the project categories seem to exist, the actual magnitude of these differences might be too small to be relevant in practice.

Overall, the findings of this thesis partly parallel the findings of Guo et al. (2018) who show that projects related to "Games", "Technology" and "Design" receive on average the highest funding. Guo et al. (2018) argue that projects related to "Dance", "Food" or "Theatre" are often culture dependent and attract mostly backers that share a similar culture. The authors assert that projects related to "Games", "Technology" and "Design" have generally objective evaluation criteria which do not change with differences in the culture, education, social status, and geographical location of the investors. Another possible explanation for the comparably weak performance of "Dance", "Food" and "Theatre" might be their consumption context. For example, projects which are assigned to these categories often require the physical proximity of backers to "consume" the products.[72] Projects related to these groups often imply some sort of social activity that requires physical presence of the backers. Therefore, these types of projects address a significantly smaller target group. On the other hand, projects related to "Games", "Technology" and "Design" are often digital in nature and can be consumed at any location.

However, although statistically significant, these performance differences found between different project categories have a small practical effect size that is too small to be relevant in practice as measured by MEM and Cohen's d. To date, no prior research that examines the influence of the project category on crowdfunding performance has

[72] For example, a theatre performance might be less impressive if it must be viewed via the internet.

considered this potential discrepancy that is a potential result of the large sample size (cp. Guo et al., 2018; K. Kim and Hann, 2013; Mollick, 2014).

4.4 Robustness Test

One important factor that might distort the results of the regression analysis is the high number of US participants on Kickstarter and, therefore, their dominance of the dataset used in this thesis (as discussed in section 3.6). Kickstarter is a US-based company, therefore, the US demonstrates the oldest and most mature market for the crowdfunding platform. In the data sample, US participants account for 70% of the backers and more than 75% of the published projects. Prior studies have recognized the dominant role of the US on Kickstarter and highlighted it as a potential bias to their findings (see Guo et al., 2018). Despite the awareness of this problem, no prior study has examined Home Bias on Kickstarter without the influence of US participants.

By eliminating US backers and entrepreneurs from the sample, the number of observations in the sample reduces from 1,118,654 to 336,740. The Negative Binomial regression model (as specified in Equation 9, p. 97) is fitted on this reduced data sample and re-estimated for both statistical and practical significance. The results of the estimations can be inspected in Table 13 and 14 respectively.

The results of the robustness test show a similar pattern as the main model. All described influencing directions of the independent variables remain constant in the robustness test. For example, geographical distance (the main variable of interest) is recurrently found to have a statistical negative effect on the count of backers. However, by eliminating US participants from the data, this variable is not statistically significant at the threshold of $p < 0.0001$ as in the main model. The change in statistical significance might be explained by the considerably smaller sample size (1,118,654 vs. 336,740 observations). Therefore, the model might not have sufficient power to detect the particularly small effect of distance on the count of backers.

The MEM for distance of -0.002 backers (cp. Table 14) is also considerably lower in the robustness test than in the main model (-0.08 backers) as reported in Table 12. Overall, all results of the robustness test confirm the initial finding that geographical distance has a rather small relevance in practice, even when US participants are excluded from the sample. Therefore, confirming the finding that geographical distance between backers and entrepreneurs does not affect the overall count of project supporters in international reward-based crowdfunding.

The results of the robustness test (cp. Table 13 and 14), in terms of potential influencing directions and p-values for the inspected other variables (GDP per capita, PWL, Large Projects, Project Category and Covid-19), largely reflect the findings of the main model estimation (cp. Table 11 and 12). They confirm the overall findings of this thesis that individual wealth (of backers and entrepreneurs), project category and economic crises do not seem to have a practically relevant impact on crowdfunding performance. On the other hand, herding behaviour does seem to positively affect the outcome of crowdfunding projects. One noteworthy difference, however, is that the Cohen's d value for the PWL badge in the robustness test changes to "Small" (rather than "Medium" in the original sample). This could imply that non-US backers might be less influenced by third-party endorsements than US backers. One possible explanation might be that backers from other countries do not attribute Kickstarter (an US based enterprise) the same trustworthiness as do US backers (i.e., another form of Home Bias).

Table 13: Model Results for Robustness Test of Adapted Sample

*The table shows the results for the Negative Binomial regressions on a sample that excludes US backers and US entrepreneurs. The sample is considerably smaller than the original sample and consists of 336,740 observations. All statistics are reported as IRRs. The estimates can be interpreted as the multiplicative effect for a one-unit increase of the respective variable. In this context, all estimates <1 indicate a negative effect or a decrease of the dependent variable, while estimates >1 reveal a positive relationship. The symbols *, **, and *** denote significance of the variable at a 0.01, 0.001, and <0.0001 level, respectively. The third and fourth column present the lower (2.5%) and upper limit (97.5%) of the 95% confidence interval.*

Dependent Variable = Count of Backers				
Explanatory Variable	**Coef.**	**IRR**	**Conf. Limit 2.5%**	**Conf. Limit 97.5%**
Distance	-0.67	0.99*	0.99	0.99
GDP-B	0.01	1.00***	1.00	1.00
GDP-E	-0.01	0.99***	0.99	1.00
PWL	0.53	1.70***	1.68	1.72
Large Project	2.37	10.74***	10.63	10.85
Covid-19 Pandemic	0.08	1.09***	1.05	1.13
Category = Art	-0.71	0.49***	0.48	0.50
Category = Comics	-0.94	0.39***	0.38	0.40
Category = Crafts	-0.91	0.40***	0.38	0.42
Category = Dance	-0.82	0.44***	0.41	0.47
Category = Design	-0.08	0.92***	0.90	0.93
Category = Fashion	-0.62	0.54***	0.53	0.55
Category = Film & Video	-0.55	0.58***	0.57	0.59
Category = Food	-0.71	0.49***	0.48	0.51
Category = Journalism	-0.51	0.60***	0.57	0.63
Category = Music	-0.78	0.46***	0.45	0.47
Category = Photography	-0.72	0.48***	0.47	0.50
Category = Publishing	-0.77	0.46***	0.45	0.47
Category = Technology	-0.19	0.83***	0.81	0.84
Category = Theatre	-0.92	0.40***	0.38	0.42

Table 14: Effect Size Measures for Adapted Sample

This table shows the IRRs, Cohen's d values, and the respective interpretation suggestion for the d values in terms of effect size.

Dependent Variable = Count of Backers				
Explanatory Variable	**IRR**	**MEM**	**D**	**Effect Size**
Distance	0.99	-0.002	-0.001	Very Small
GDP-B	1.00	0.006	0.002	Very Small
GDP-E	0.99	-0.008	-0.003	Very Small
PWL	1.70	1.379	0.477	Small
Large Project	10.74	19.115	6.617	Very Large
Covid-19 Pandemic	1.09	0.171	0.059	Very Small
Category = Art	0.49	-0.997	-0.345	Small
Category = Comics	0.39	-1.198	-0.415	Small
Category = Crafts	0.40	-1.174	-0.406	Small
Category = Dance	0.44	-1.097	-0.380	Small
Category = Design	0.92	-0.159	-0.055	Very Small
Category = Fashion	0.54	-0.909	-0.315	Small
Category = Film & Video	0.58	-0.826	-0.286	Small
Category = Food	0.49	-0.998	-0.345	Small
Category = Journalism	0.60	-0.780	-0.270	Small
Category = Music	0.46	-1.067	-0.369	Small
Category = Photography	0.48	-1.009	-0.349	Small
Category = Publishing	0.46	-1.056	-0.366	Small
Category = Technology	0.83	-0.335	-0.116	Very Small
Category = Theatre	0.40	-1.178	-0.408	Small

The presented robustness test is an important contribution to knowledge because it is the first attempt in crowdfunding literature to account for the dominant role of the US on Kickstarter platform, which is a common data source for many studies (cp. Table 2, p. 65). Although some scientists are aware that the data imbalance might introduce a potential bias to their findings (e.g., Guo et al., 2018), no prior study has examined Home Bias on Kickstarter without the influence of US participants. This thesis is the first to test its findings for consistency and reliability by comparing the results of the main model that covers all international backing actions to the results of a regression model that excludes US participants from the data. Overall, the robustness test suggests that the presented findings of this thesis are consistent, reliable, and robust.

4.5 Summary on Findings and Analysis

Chapter 4 presents the results of the Negative Binomial regression and the additional assessment of practical effect sizes for the different variables. One key finding is that when dealing with Big Data models, researchers must be cautious with inferring practical relevance of their findings from observed statistically significant relationship patterns (via p-values). The analysis presented in this thesis shows that Big Data models have the power to detect particularly small patterns in the data that might be relevant from a statistical point of view, but negligible in practice.

For example, although increasing geographical distance has a negative statistical effect on the count of backers in crowdfunding campaigns, the actual change in the count of backers is too small to be relevant for entrepreneurs. A central implication of this finding is that some of the earlier studies might have overstated the practical relevance of Home Bias in crowdfunding. This is because most research in crowdfunding is conducted on comparatively large samples that are likely to be affected by the p-value problem.

The finding that statistically significant relationship patterns (as detected by large models) can be negligibly small in practical terms, is also reflected in some of the other inspected variables. For example, although statistically significant influences could be identified by the Negative Binomial regression model for the GDP per Capita (GDP), Project Categories (C) and Covid-19 pandemic (P) on the count of backers, the subsequent analysis of effect sizes shows that these effects are too small to be relevant in practice. Therefore, this thesis finds that these variables do not have a noteworthy effect on the performance of international reward-based crowdfunding projects.

This thesis shows that PWL (representing third-party endorsements) and Large Projects (representing Herding Behaviour) show both a statistically significant and practically relevant effect on the count of backers. This finding suggests that these variables can serve as effective quality signals that can positively affect the count of backers in crowdfunding projects.

Another important finding presented in this chapter is that the high number of US participants in the data does not seem to affect the overall model results. The results of the robustness test that excludes US participants (US backers and entrepreneurs) from the dataset is consistent with the main model results in terms of the variables' influencing directions, statistical significance (via p-values) and practical relevance (via Cohen's d and MEM). This is a particularly important finding because, so far, no prior research has investigated the potential bias resulting from the high number of US participants on Kickstarter platform.

Chapter Five: Conclusion

5.1 Introduction

This chapter provides a concise summary of this thesis and describes the research approach, most important findings, contributions, and limitations. In section 5.2, the research context is restated via discussing the insufficient capital supply to start-ups and its association with the Home Bias problem. Moreover, this section explains why crowdfunding bears the potential to alleviate the relevance of Home Bias in investment decisions. Section 5.3 recapitulates the identified research gaps in the relevant literature on Home Bias in crowdfunding, whereas section 5.4 describes the research approach of this thesis to close these gaps. Section 5.5 presents the findings, contributions, and limitations of this thesis as well as recommendations for future research. The chapter concludes with section 5.6 which addresses the confronted challenges within this research.

5.2 Context of This Thesis

The access to financial resources is considered one of the major growth obstacles for many businesses globally (Alibhai et al., 2017; Beck & Demirguc-Kunt, 2006; Carpenter & Petersen, 2002; World Bank, 2016a). It is particularly decisive for the creation and survival of start-ups which are important drivers for innovation, job creation and economic development (Ayyagari et al., 2016). Therefore, the development of new concepts to provide finance to young businesses has been one of the major priorities of governments, the World Bank Group and other development institutions around the globe (Dinh et al., 2010; Hwang et al., 2019; World Bank, 2012, 2013). However, only very few of the introduced measures have led to systematic change in the funding problem of young firms (Hwang et al., 2019). The current global credit gap is estimated to range between 3 to 5 trillion USD (Ferrando et al., 2019; International Finance

Corporation, 2017; World Bank, 2016a) and more than 65 million enterprises (World Bank, 2016) remain constrained by insufficient funding.

One important factor that contributes to the global finance gap is the Home Bias problem in investment decisions. Home Bias is the proven tendency of individuals to prefer geographically proximate interaction partners (Niemand et al., 2018). In business finance, Home Bias is considered an important problem because it promotes an uneven distribution of financial resources in terms of geography (Hwang et al., 2019; Sorenson et al., 2016). Individual business hubs concentrate large sums of venture capital that is mostly unavailable to firms that are not part of these geographical regions (e.g., the Silicon Valley region) (Florida & King, 2016). Thereby, Home Bias is to the detriment of both entrepreneurs and investors because entrepreneurs are unable to raise funds for their ventures if they are remote from established business hubs and investors miss potential business opportunities by focusing only on geographically proximate firms (Chen et al., 2009; Mollick, 2013; T. Stuart & Sorenson, 2003).

The rise of the internet has, however, introduced many influential changes. It has affected how people work, socialize, create and share information (Niemand et al., 2018). The financial sector has also experienced numerous internet-driven innovations (Niemand et al., 2018). One particularly important change has been introduced through the concept of “crowdfunding” that has fundamentally changed how savers and borrowers, and investors and investees could interact (Surowiecki 2004).

The term crowdfunding describes the attempt of collecting financial resources from a large and unaffiliated crowd of investors via specialized digital platforms (H. Kim & Kim, 2017). Crowdfunding is fundamentally different from traditional fundraising concepts because the entire process of capital procurement takes place in a virtual environment on the internet. Business founders use specialized crowdfunding platforms to describe their ideas and present it to a large and diversified crowd of potential investors (Mollick, 2014). Private and institutional investors, in turn, use crowdfunding platforms to search for new business ideas and support founder(s) with financial resources via the provided infrastructure of the specialized platforms (e.g., Kickstarter). In return, the investors or “backers” receive a reward in the form of an early version of the new product itself, interest on the investment, or equity of the firm. While different crowdfunding models exist (i.e., donation-, equity-, interest- and reward-based model), the reward-based model is most popular for international, business-oriented fundraising and, hence, is the focus of this thesis.

Multiple authors have highlighted the important impact of reward-based crowdfunding on the financial sector (Mollick & Robb, 2016, Cumming et al. 2019, Guo et al., 2018) and discussed how the emerging industry might help to overcome geographic, linguistic, and cultural barriers in business finance (Ahlers et al. 2015, Mollick and Robb 2016) by introducing new community-based trust mechanisms and eliminating some of the transaction-related costs. This thesis contributes to the prevailing discussion by investigating the existence of Home Bias in international reward-based crowdfunding and examining the practical relevance of geographical distance between backers and entrepreneurs in project backing decisions.

5.3 Addressed Research Gaps

Current research is inconsistent on whether internet-enabled crowdfunding can alleviate the Home Bias problem in investment decisions and contribute to a more efficient distribution of financial resources globally. On the one hand, some scholars argue that the digital nature of crowdfunding increases the scope of social networks and connects founders and investors more efficiently (Dekel et al., 2016; K. Kim & Hann, 2013; Mollick & Robb, 2016). Moreover, crowdfunding reduces transaction costs and introduces new approaches to deal with the information asymmetry problem, a major driver for Home Bias, which should render geographical distance irrelevant (Thierer et al., 2015). On the other hand, other scholars find evidence that Home Bias continues to matter in crowdfunding because the information asymmetry problem persists and different behavioural reasons, rather than rational, are responsible for the tendency of investors to choose geographically proximate target firms (Mollick, 2014; Kim & Kim, 2017; Niemand et al., 2018).

However, several gaps exist in the current literature on Home Bias in crowdfunding. One major weakness is that most research examines the effect of Home Bias in individual countries or regions but does not assess it in an international context (K. Kim & Hann, 2013; M. Lin & Viswanathan, 2016; see Mollick, 2013, 2014). Particularly striking is that most studies focus exclusively on the US market and deliberately exclude projects or backers from foreign countries. Moreover, the most cited studies only cover data from the early years of the industry (prior to 2013) when crowdfunding was highly fragmented and mostly a niche market (cp. Agrawal et al., 2013; 2015; Mollick, 2013; 2014; Lin & Viswanathan, 2016; Burtch et al., 2014). However, over time the crowdfunding industry has undergone considerable consolidation which has given rise to new global market leaders that integrate different areas of the world and provide unprecedented transnational markets for venture capital (Ruhnau, 2019). In recent years, very few authors have re-evaluated whether Home Bias is prevalent in this modern form of crowdfunding (cp. Guo et al., 2018). Moreover, no research has considered new potential influencing factors that might have changed the relevance of Home Bias in online fundraising (e.g., the global Covid-19 pandemic).

Furthermore, while many studies in crowdfunding use comparably large data samples (cp. Guo et al., 2018; Burtch et al., 2014), none of them has considered the potential caveats that are highlighted by the statistics literature on dealing with Big Data in quantitative analysis. For example, Lin et al. (2013) warn that low p-values can be an artefact of large sample size and that datasets with more than 10,000 observations are already prone to the so-called “p-value problem”. For comparison, many studies in crowdfunding exceed the sample size of 20,000 observations (cp. Agrawal et al., 2013; Breznitz & Noonan, 2020; Mollick, 2014; Stevenson et al, 2019). Therefore, ignoring the potential of Big Data models to find particularly small patterns and relationships in the data might lead researchers to finding statistically significant results that are of little or no practical value. Therefore, this thesis goes beyond the traditional analysis of statistical significance and devotes great attention to the assessment of different marginal effect sizes to identify the practical relevance of findings.

5.4 Research Approach of This Thesis

The aim of this thesis is to examine the existence of Home Bias in the emerging industry of international reward-based crowdfunding and to close some of the identified research gaps in the literature (as stated in the previous section). In this context, the central research question of this thesis is centred around whether geographical distance between backers and entrepreneurs affects the overall count of project supporters in international crowdfunding projects on Kickstarter, a reward-based crowdfunding platform.

The first objective of this thesis is to construct the largest and most recent crowdfunding dataset to date. Using a self-developed web-crawler, written in Python programming language, this thesis extracts data from Kickstarter on 211,695 individual crowdfunding projects from more than 200 different countries. Overall, the dataset describes more than 44 million individual backing actions that occurred in the time frame from April 2009 to June 2020, making it the largest and most recent data sample on crowdfunding in the literature to date.[73]

The second objective of this thesis is to construct a variable for geographical distance by calculating the distance between backers and entrepreneurs. For this purpose, this thesis uses the Google Maps API to estimate the latitude and longitude coordinates of backers and entrepreneurs.[74] These coordinates, in turn, are used to calculate the shortest distance for each backer-entrepreneur location combination in the dataset.[75]

---

[73] The dataset comprises information on the country of origin of backers and entrepreneurs, the project launch date, the count of backers from the individual countries, platform specific endorsements ("Projects We Love"), and the project category (i.e., "Technology").

[74] The thesis used the coordinates of country capitals to estimate the geographical location of backers and entrepreneurs. This is because Kickstarter's privacy regulations do not allow to collect data on the exact location of its community members (cp. Kickstarter, 2018).

[75] This thesis used the geodesic distance, which is the shortest distance on the surface of an ellipsoidal model of the earth Karney (2013).

The result is a consolidated dataset of 1,118,654 project-specific country-to-country investment observations that can be used for further quantitative analysis.

The third objective of this thesis is to construct a quantitative regression model that allows to examine the effect of geographical distance and other relevant control variables on crowdfunding performance in the obtained data from Kickstarter. After considering different types of possible regression models and their peculiarities, this thesis proposes a Negative Binomial regression model that describes the count of backers from a specific country to a specific crowdfunding project as a function of geographical distance to the entrepreneur. Moreover, the GDP per capita of both the entrepreneurs' and backers' home countries, third-party endorsements (PWL), project category, herding behaviour (Large Projects), and the potential effect of the Covid-19 pandemic are included as additional explanatory variables in the regression model.

The fourth objective of this thesis is to examine the potential problems of Big Data models in crowdfunding research. This objective is particularly important because the literature review of this thesis shows that although research in crowdfunding is frequently conducted on comparatively large samples, none of the existing studies adequately addresses potential caveats of dealing with Big Data models. Through graphical analysis of the dataset via CPS-charts and Monte Carlo simulations, this thesis highlights the existence of the p-value problem in Big Data models and describes potential remedies to address it that are employed in the interpretation of model results. This approach reveals that the traditional approach of statistical significance testing might be misleading in Big Data suggesting that some of the prior research on crowdfunding might have overstated the practical relevance of their findings, including the relevance of geographical distance on backing decisions.

The final objective of this thesis (Research Objective Five) is to test the firmness of the results on a dataset that excludes US participants from the data. This robustness test is essential because of two reasons: First, most of the prior research focuses exclusively on data from the US and deliberately excludes crowdfunding projects from other countries. This approach, however, is not sustainable because crowdfunding is becoming an increasingly international industry. Second, the exploratory data analysis of this thesis reveals that the collected data from Kickstarter is dominated by a high number of US backers and entrepreneurs, which might introduce a potential bias to model results. By eliminating US participants (backers and entrepreneurs) from the data, this thesis tests the validity of the findings for other countries and explicitly addresses a recognized bias of prior crowdfunding research (cp. Guo et al., 2018).

5.5 Contributions and Findings

Pragmatism, the research philosophy of this thesis, orients itself toward solving practical problems in the real world (Creswell and Clark 2011; Maxcy 2003; Rorty 2000). Consequently, this thesis must offer both explicit contributions to the advancement of academic knowledge and concrete action guidelines for practitioners on how to improve the success probability of their crowdfunding campaigns, to justify its existence. Therefore, this section describes the relevant findings and contributions of this thesis to both knowledge and practice.

5.5.1 Contributions to Knowledge

The literature on crowdfunding shows a considerable disagreement on the role of Home Bias in digital fundraising. For example, Mollick & Robb (2016) argue that the digital nature of crowdfunding increases the scope of business networks and connects founders and investors more efficiently across geographic, linguistic, and cultural barriers (see also Dekel et al., 2016, K. Kim & Hann, 2013). On the other hand, scientists such as Gallemore et al. (2019) provide evidence that spatial context continues to matter in online fundraising and that crowdfunding might not democratize the access to finance as some researchers hope (see also M. Lin & Viswanathan, 2016; Guo et al., 2018).

This thesis shows that although quantitative Big Data models can identify a statistically significant negative influence of distance on the count of backers, this influence is rather small and negligible if inspected for its practical effect size. This thesis contributes to academic knowledge by providing a possible explanation for the existing disagreement on the role of Home Bias in crowdfunding by showing that both statements “geographical distance has a statistically significant negative effect on the count of backers” and “geographical distance has a negligibly small practical relevance in crowdfunding” are not mutually exclusive.

This thesis makes an important contribution to academic knowledge by being the first to address potential caveats of large samples in crowdfunding research. This is important because crowdfunding research typically deals with comparatively large datasets (sample size > 20,000). Therefore, the thesis goes beyond the traditional analysis of statistical significance testing and devotes great attention to the assessment of different marginal effect sizes to identify the practical relevance of findings. This approach led to the important insight that statistically significant relationships between variables, detected through Big Data models (i.e., the influence of distance on the count of backers), do not necessarily imply practical relevance. Therefore, some of the earlier findings on Home Bias in crowdfunding should be handled with caution as their practical relevance might have been overstated due to the p-value problem in big samples. Overall, this thesis suggests that lowering the significance threshold (i.e., to $p < 0.0001$) will not solve the p-value problem in crowdfunding research and that instead scientists must pay more attention to distinguishing between statistical significance of detected relationship patterns (via p-values) and their practical relevance (via actual effect sizes).

This thesis complements the academic research on crowdfunding by providing the most extensive and most international analysis on Home Bias. The existing research on crowdfunding tends to focus on Home Bias in individual countries (mostly US) or specific regions (e.g., EU). This thesis provides a more realistic analysis of Home Bias in crowdfunding because it incorporates data on more than 211,695 individual crowdfunding projects from over 200 different countries. So far, this thesis is the largest and most recent analysis of crowdfunding that can be found in the literature. Therefore, the findings of this thesis are likely to offer higher validity for the reward-based crowdfunding industry.

The Negative Binomial regression model proposed in this thesis complements the existing quantitative research methods in crowdfunding. This thesis is the first to explain why generalized linear models, and specifically the Negative Binomial, are the best approach to model the count of backers via different explanatory variables (such as geographical distance). This thesis provides different arguments why the Negative Binomial regression model offers a more robust alternative to deal with count data (i.e., the count of backers) than, for example, the popular Linear Regression (for example, Negative Binomial regression does not require homoscedasticity). The analysis of different potential regression models, including their strength and weaknesses, provides valuable guidelines for future researchers.

This thesis re-evaluates the practical relevance of some of the variables that have been used by prior research. For example, it finds that Project Category, although statistically significant, seems to have a rather small practical effect on the count of backers. This is an important finding because it questions some of the earlier results on the effect of Project Category such as those presented by Guo et al. (2018) and K. Kim and Hann (2013). Moreover, this thesis finds that the actual difference in project category (measured by its effect on the count of backers) is rather small and should not discourage entrepreneurs from using crowdfunding for their ventures. This implies that crowdfunding can demonstrate an effective fundraising approach for very different types of products that might have been neglected by traditional financiers.

This thesis also adds valuable knowledge to the discussion on the relevance of individual wealth (of backers and entrepreneurs) in crowdfunding. It is, so far, the first thesis to examine the effect of GDP per capita of backers' and entrepreneurs' home countries on crowdfunding performance, which is considered a better estimate for individual wealth than GDP. The results suggest that individual wealth does not have a practically relevant effect on the count of backers. Consequently, international reward-based crowdfunding does not seem to promote investment flows from high-wealth individuals to projects from low-wealth individuals as found by Burtch et. al (2014) for Kiva.org, a social-lending crowdfunding platform. Instead, the findings of this thesis suggest that, on Kickstarter, the product idea itself might be more relevant than the alleged predicament of the entrepreneur (see also Davis et al., 2017).

Furthermore, the findings of this thesis advance the discussion on the information asymmetry problem in crowdfunding and, more specifically, on effective quality signals. This thesis finds that both herding behaviour and third-party endorsements have a statistically significant and practically relevant effect on the count of backers, although the effect of herding behaviour is considerably larger than the effect of third-party endorsements. Herding Behaviour, defined as projects that exceed the threshold of 107 backers in total, is found to have the highest impact on the count of backers among all examined influencing factors. This thesis is the first to use this threshold approach to gauge the effect of herding behaviour on the performance of crowdfunding projects. The central finding of this approach is that backers seem to consider backer accumulations as a valuable quality signal, as predicted by the Information Asymmetry Theory (cp. Akerlof, 1970). However, a possible explanation could also be that sorting algorithms on Kickstarter, which make projects with higher count of backers more visible, are responsible for the good performance of large projects.

This thesis is also the first to evaluate the effect of global economic crises on the crowdfunding industry by studying the effect of Covid-19 pandemic. In this context, it advances the academic discussion on the potential of crowdfunding to serve as a reliable source for capital for start-ups and small businesses during global economic crises. The findings of this thesis suggest that the Covid-19 pandemic does not have a practically relevant impact on the count of backers. Thereby, this thesis provides evidence that reward-based crowdfunding is not affected by economic crises in the same way other traditional financing channels might have and can serve as a reliable alternative for entrepreneurs when access to capital via traditional financiers is difficult. One possible explanation for the small influence of economic crises is that crowdfunding is a digital fundraising method that mostly relies on private investors. Grasso et al. (2021) highlight that economic crises can sometimes inspire private individuals and social groups to engage in acts of solidarity, mutual support and help which in turn might compensate the initial economic shock. However, it must also be admitted that the findings presented in this thesis can only be understood as preliminary because the underlying data covers an early time frame of the Covid-19 pandemic (March – June 2020). The observed influence might have changed during the course of the pandemic and, therefore, requires further analysis.

Another important contribution to research on Home Bias in crowdfunding is that this thesis is the first to conduct a robustness test that controls for the potential bias caused by the dominance of US participants (backers and entrepreneurs) in the dataset. The existing crowdfunding literature is either US-focused (cp. Mollick, 2014; Kim & Hann, 2015; Gallemore et al., 2019) or uses data that is dominated by a high number of US participants (as in this thesis) (cp. Guo et al., 2018). By excluding US participants from the data, this thesis is the first to verify the firmness of the results for other countries and to address a recognized bias of prior crowdfunding research.

5.5.2 Contributions to Practice

As far as professional contributions are concerned, this thesis provides several findings that might be relevant for future entrepreneurs that plan to engage with crowdfunding. The most important finding is that geographical distance plays a subordinate role in international reward-based crowdfunding when inspected for the practical effect size. Consequently, entrepreneurs should be aware that, when using crowdfunding, they are operating on a global stage. Products and supply chains, therefore, need to be designed appropriately to fit this international context. For example, entrepreneurs should ensure that their intellectual property is protected outside of their home regions before launching a crowdfunding campaign. Moreover, before offering product rewards to international backers, entrepreneurs must consider the global shipping costs as well as potential import restrictions of other countries.

Another important aspect that entrepreneurs need to reconsider is the focus of their marketing activities. The findings presented in this thesis on the low relevance of geographical distance suggest that entrepreneurs in reward-based crowdfunding should expect backers from distant countries at least to the same extent as from nearby countries. This, in turn, has important consequences for the preparation, execution, and follow-up of crowdfunding campaigns. A particularly important insight for Kickstarter is provided by the Exploratory Data Analysis presented in section 3.6.6 (p. 148) of this thesis. The backer origin analysis for different countries shows that on Kickstarter most backers typically come from only five countries (US, CA, AU, UK, DE). Consequently, focusing the marketing activities on these countries (i.e., adapting language and currency) might be a more rewarding strategy.

Further practice-relevant findings can be drawn from the analysis of effective quality signals in international reward-based crowdfunding. The findings of this thesis suggest that obtaining the PWL badge might be a rewarding strategy. This is probably because the PWL badge is considered a reliable quality signal for projects that only few entrepreneurs are able to obtain. The results suggest that investing time in effort in obtaining some type of validation from third parties might help entrepreneurs to improve the performance of their crowdfunding projects. However, this thesis also finds that the effect of third-party endorsements is considerably more moderate than that of herding behaviour. Herding behaviour is found to have the strongest impact on the count of backers compared to all inspected variables. The results of this thesis suggest that entrepreneurs can considerably improve the performance of their crowdfunding projects if they manage to attract at least 107 backers (used threshold in this thesis for herding behaviour).

Another implication of this finding is that it is considerably more difficult for entrepreneurs to attract early investors in the beginning of the crowdfunding campaign than additional investors after the campaign has reached a certain level of funding. Therefore, besides the presented findings on effective quality signals, this thesis also provides a useful link to other potential influencing factors that have been found by previous research to have a positive impact on early backers (cp. Table 1, p. 55). For example, different studies show that constructing an emotional and engaging campaign video can have a positive effect on early backers and, thereby, increase crowdfunding success (Davis et al. 2017; Frydrych et al. 2014). Also, Krishnan et al. (2015) find that crowdfunding projects that choose a fixed funding goal ("all-or-nothing" model) are more likely to succeed than projects that do not choose this approach ("keep-what-you-get" model). For more information see also Appendix B.

However, even if all recommendations presented in this thesis are implemented, this does not seem to be a guarantee that a crowdfunding project will eventually succeed. At the current state of research, no approach has been found that could predict with a high degree of certainty whether a crowdfunding project will ultimately achieve its goal or not. There are still too many variables that influence the outcome that have not yet been sufficiently researched (such as the quality of the campaign video, platform sorting algorithms, the perception of the entrepreneur, promotional activities outside the crowdfunding platform).

5.6 Limitations and Future Research

The presented findings in this thesis are subject to certain limitations. One limitation is that the thesis focuses exclusively on data extracted from one crowdfunding platform (Kickstarter). This focus might introduce a potential bias because Kickstarter's community, tools and processes might be unique to this specific platform and, therefore, impede the generalisability of the findings.

However, Kickstarter demonstrates the largest, most popular and most international reward-based crowdfunding platform to date (Statista, 2018) and has therefore been used by many prior studies (cp. Table 2, p. 65). The platform often serves as a role model for many other crowdfunding platforms, which is expressed in the fact that many competitors seem to follow Kickstarter's successful practices when designing and adapting their marketplaces. New features, functions and tools introduced by Kickstarter are often quickly copied by competitors. Therefore, the structure and design of other crowdfunding platforms greatly resembles the structure and design of

Kickstarter's platform.[76] Nevertheless, future research could consider examining Home Bias in other relevant reward-based crowdfunding platforms.[77]

Another concern is that the presented findings are likely to apply for reward-based crowdfunding, however, should only cautiously be transferred to other crowdfunding models (e.g., donation-, interest- and equity-based models). Depending on the model, differences might exist in the backers' motivation to participate (e.g., charity vs. profit), the legal restrictions (especially in equity-based crowdfunding) and the amount of capital invested (e.g., threshold for minimum investment). The crowdfunding market is diverse and continuously evolving. Over time, several specialized crowdfunding platforms have emerged.[78] Therefore, it is possible that other crowdfunding models follow different, undiscovered dynamics. Decisive differences within the various crowdfunding models in terms of participation, motivation, targeted reward or risk tolerance have already been addressed by some previous studies (cp. K. Kim & Hann, 2013; Gerber & Hui, 2013). Future research could consider comparing the effect of Home Bias across different crowdfunding models.

An additional potential limitation is that the findings of this thesis might be biased because the collected data only includes transactions that actually occurred and are described on Kickstarter but does not consider all possible country-country constellations that could have occurred. In other words, it is not possible to retrace how many backers, and from which countries, saw a specific campaign and made the conscious decision not to support it because the entrepreneur is based in a distant country. Therefore, the lack of information as well as the insufficient available computational capacity do not allow to consider these "unobserved" constellations.

[76] Examples: www.indiegogo.com, www.seedrs.com, www.startnext.com. All platforms present projects, number of investors and rewards in a highly similar manner.

[77] Another major reward-based crowdfunding platform is www.indiegogo.com.

[78] One example is Kiva.org, which is a charitable organization that demonstrates a hybrid version of donation- and interest-based crowdfunding. Another example is Syndicateroom.com, which is an equity-based crowdfunding model that includes so-called "Lead Investors" to reduce the information asymmetry problem.

Future research could try to address this limitation by, for example, using more powerful computers to consider all possible country combinations that could have occurred or by designing special experiments where the project exposure to backers and their conscious "negative" decision to back a specific project can be recorded and examined.

The collected data does not provide information on the amount of individual backers' monetary contributions because crowdfunding platforms typically do not publish this information on their websites. In other words, the collected data shows the aggregate count of backers from a specific country but does not consider whether their average monetary contributions were different from those of other countries. Future research could try to obtain this information directly from crowdfunding platform operators (e.g., by conducting joined research) and investigate whether geographically distant backers contribute to the same amount as geographically close backers. This approach would allow to examine if Home Bias manifests itself in a different form in crowdfunding. For example, it would be possible that although increasing distance does not have a relevant effect on the count of backers, it does negatively impact the individual monetary contributions of backers. Future research could examine this theoretically possible scenario and provide new insights on Home Bias in crowdfunding.

This thesis is the first to address potential problems of Big Data models in crowdfunding research. Future studies could extend this discussion, question prior findings, and develop new strategies to ensure that the obtained results are not only statistically significant but also practically relevant. Future research could also take an entirely different methodological approach and, for example, examine the effect of Home Bias in small qualitative studies, which are almost non-existent in the current crowdfunding literature (cp. Table 2, 65). Qualitative studies could provide a different insight into the individual cognitive processes and motivations associated with active crowdfunding backing decisions. However, they might be harder to execute due to the difficult access to participants.

Despite of the stated limitations, the findings of this thesis are highly relevant for both practice and theory. They provide important guidelines for entrepreneurs on how to design their campaigns more effectively and expand the theoretical discussion on the potential of reward-based crowdfunding to alleviate the relevance of Home Bias in business financing. This thesis offers a valuable insight into the current patterns of the crowdfunding industry.

5.7 Challenges in Crowdfunding Research

This research confronted numerous challenges which are briefly described in this section to aid future researchers in conducting crowdfunding research.

The first major challenge was the access to data. Although many crowdfunding projects are publicly available on the internet, only few provide geographical information on backers. Extracting this information from platforms is highly effort- and time consuming. Therefore, to capture sufficient data for this thesis, it was necessary to develop a special web-crawler. A web-crawler is a computer program that can access the internet via the "Hypertext Transfer Protocol" (HTTP) and copy specific data of interest from the website into a local database or spreadsheet. Web-crawling is a common approach in research on crowdfunding because it facilitates the collection of comparably large data samples with high levels of data validity (Frydrych et al., 2014; K. Kim & Hann, 2013; Mollick, 2014). However, developing a web-crawler requires programming knowledge that had to be acquired specifically for this thesis (which sometimes delayed the overall progress of the research). Future researchers should also expect that platform operators are difficult to reach and generally reserved to share any internal data. This experience was also confirmed by other scientists who have dealt with crowdfunding research (see Marom et al. 2014; Kim and Hann 2013). Therefore, future research might be limited to projects with data that can be collected without the help of platform operators via manual work or web-crawling.

This thesis is based on 211,695 crowdfunding projects and demonstrates, thereby, the largest crowdfunding analysis to date. However, the size of the sample introduced several challenges. For example, identifying data errors and conducting calculations was tedious, time consuming and required high computational power. Model estimations in particular required much computing time and therefore delayed the progress of this thesis at times. In many cases the used computer would not manage to converge the model, especially when more complex estimations were tried (e.g., hurdle models). This thesis uses R and Python programming languages interchangeably to conduct the estimations. The problem was that R, as it loads all the data into computer memory, was sometimes not able to conduct certain estimations due to limited memory capacity. For example, confidence interval calculations would return “N/A” values.

Python scripts provided a better solution in this case. The size of the data also makes it difficult to use traditional tests for model development (e.g., Hausmann Test). Tests that are based on p-values tend to strive towards zero and can lead to erroneous assumptions. The problems with large data samples and potential solutions have been addressed several times throughout this thesis. It is important to highlight that this was an unforeseen difficulty when this project started.

5.8 Final Remarks

The completion of this thesis has been a long, demanding, but also highly educational journey. It started with the intention to investigate whether advancements in the internet technology offer start-ups new opportunities to access financial resources to fund creative ideas and new business ventures. This is relevant because insufficient funding is considered one of the major growth obstacles for start-ups worldwide. In particular, the primary aim of this thesis was to examine the existence of the well-described Home Bias problem in international reward-based crowdfunding.

As the research developed, some unexpected challenges emerged, extending the focus of the thesis to addressing additional problems, such as working with large data samples (“Big Data”) that has been largely ignored in crowdfunding research. In particular, this thesis identifies the importance of going beyond the traditional analysis of statistical significance to assessing the practical relevance of findings.

Despite challenges, this thesis offers new and interesting insights into the crowdfunding industry. Overall, it provides evidence that geographical distance plays a rather subordinate role in international reward-based crowdfunding. Moreover, the thesis provides new insights into the influence of additional variables that have potential effect on the count of backers namely GDP per capita, project category, third-party endorsements, herding behaviour and Covid-19 pandemic.

Because the model results presented in this thesis might appear not entirely unambiguous to readers (due to the described statistical artefacts of large samples), this thesis devoted much time and effort in showing that, depending on the used research methods, different researchers can come to different conclusions when examining the same data. This echoes the fundamental discussion on the ontological and epistemological position of researchers as described in section 3.2. Furthermore, it reflects the current discussion in the literature on when scientific results truly contribute to the advancement of knowledge.

This thesis uses a pragmatic approach to explain the different methods to evaluate the usefulness of findings. By analysing the statistical significance, practical significance, and practical relevance of the results obtained through the negative binomial regression model, this thesis develops a comprehensible argumentation structure to support its final conclusions.

This thesis shows that measures of statistical significance can be affected by sample size and, thereby, do not always provide a reliable measure for the usefulness of results. Similarly, relative effect size measures (such as IRRs) can sometimes obscure the real magnitude of an effect, conveying that a discovered influence is more important than it is (cp. the examples from the pharmaceutical literature in section 3.6.3, p. 132). This problem applies in particular to count models that have a high number of small values in their distribution (i.e., a 50% reduction might lead to a factual difference of 1 backer).

Because of the described reasons, this thesis relies on absolute effect size measures (factual count of backers estimated via MEM) and the established approach introduced by Cohen (Cohen’s d) to evaluate the relevance of relationships, since these approaches have been found to be the most robust in the evaluation of practical relevance, given the conditions of this thesis. Moreover, it must be admitted that a certain degree of the researcher's subjective judgement affected the evaluation of practical relevance of the presented findings. This, however, is encouraged by different authors that are aware of the p-value problem (cp. Kirk 1996, Mohajeri et al. 2020) and is one of the reasons why this thesis went beyond the traditional NHST to evaluate the usefulness of its results.

Using the described methods of assessment, this thesis concludes that neither geographical distance, nor the GDP per capita, project category and Covid-19 pandemic have a practically relevant impact on the success of international crowdfunding campaigns. On the other hand, herding behaviour and third-party endorsements do seem to have a practically relevant impact on international reward-based crowdfunding projects because they act as reliable quality signals. However, more research is required with special focus on practical relevance to confirm (or disprove) the findings of this thesis.

Overall, this thesis addresses multiple research gaps in the field of international crowdfunding and makes valuable contributions to academic and practice-oriented knowledge. The author hopes that the insights gained in this thesis will be useful for future entrepreneurs and researchers.

6 References

Ackerberg, D. A. (2001). Empirically Distinguishing Informative and Prestige Effects of Advertising. *The RAND Journal of Economics*, *32*(2), 316. https://doi.org/10.2307/2696412

Agrawal, A., Catalini, C., & Goldfarb, A. (2010). Entrepreneurial Finance and the Flat-World Hypothesis: Evidence from Crowd-Funding Entrepreneurs in the Arts.

Agrawal, A., Catalini, C., & Goldfarb, A. (2011). *The Geography of Crowdfunding*. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w16820

Agrawal, A., Catalini, C., & Goldfarb, A. (2013). *Some Simple Economics of Crowdfunding*. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w19133

Agrawal, A., Catalini, C., & Goldfarb, A. (2015). Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions. *Journal of Economics & Management Strategy*, *24*(2), 253–274. https://doi.org/10.1111/jems.12093

Agresti, A. (2002). *Categorical data analysis* (2. ed.). *Wiley series in probability and statistics*. Hoboken, NJ: Wiley-Interscience. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=85496 https://doi.org/10.1002/0471249688

Agresti, A. (2007). *An introduction to categorical data analysis* (2. ed.). *Wiley series in probability and statistics*. Hoboken, NJ: Wiley-Interscience. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10278250 https://doi.org/10.1002/0470114754

Ahearne, A. G., Griever, W. L., & Warnock, F. E. (2004). Information costs and home bias: an analysis of US holdings of foreign equities. *Journal of International Economics*, *62*(2), 313–336. https://doi.org/10.1016/S0022-1996(03)00015-1

Ahlers, G. K. C., Cumming, D., Günther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship, Theory and Practice : ET & P*, *39*(4), 955–980.

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, *84*(3), 488. https://doi.org/10.2307/1879431

Alibhai, S., Bell, S., & Conner, G. (2017). What's Happening in the Missing Middle? : Lessons from Financing SMEs. Retrieved from https://openknowledge.worldbank.org/handle/10986/26324 License: CC BY 3.0 IGO.

Allison, P. (2009). *Fixed Effects Regression Models*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc. https://doi.org/10.4135/9781412993869

Alois, J. (2021, April 3). Germany: Investment Crowdfunding Declines to €400 Million in 2020, €100 Million Reported in Q1 2021. *Crowded Media Group*. Retrieved from https://www.crowdfundinsider.com/2021/04/173706-germany-investment-crowdfunding-declines-to-e400-million-in-2020-e100-million-reported-in-q1-2021/

Alvesson, M. (1995). *The Meaning and Meaninglessness of Postmodernism: Some Ironic Remarks*: Organization Studies, 16(6): 1047-75.

Amel, D. F., & Brevoort, K. P. (2005). THE PERCEIVED SIZE OF SMALL BUSINESS BANKING MARKETS. *Journal of Competition Law & Economics*, *1*(4), 771–784. https://doi.org/10.1093/joclec/nhi021

Anderson, J. E., & van Wincoop, E. (2004). Trade Costs. *Journal of Economic Literature*, *42*(3), 691–751. https://doi.org/10.1257/0022051042177649

Arnaboldi, V., Passarella, A., Conti, M., & Dunbar, R. I. M. [Robin Ian MacDonald] (2015). *Online social networks: Human cognitive constraints in Facebook and Twitter personal graphs*. *Computer science reviews and trends*. Amsterdam, Netherlands: Elsevier. Retrieved from http://proquest.tech.safaribooksonline.de/9780128030424

Ayyagari, M., Juarros, P., Martinez Peria, M. S., & Singh, S. (2016). *Access to Finance and Job Growth: Firm-Level Evidence across Developing Countries*: The World Bank. https://doi.org/10.1596/1813-9450-7604

Babbie, E. R. (2013). *The practice of social research* (13. ed., internat. ed.). Independence, KY: Wadsworth Cengage Learning.

Bade, M., & Walther, M. (2021). Local preferences and the allocation of attention in equity-based crowdfunding. *Review of Managerial Science.* Advance online publication. https://doi.org/10.1007/s11846-020-00429-6

Baltzer, M., Stolper, O. A., & Walter, A. (2012). Is Local Bias a Cross-Border Phenomenon? Evidence from Individual Investors' International Asset Allocation. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2200821

Barasinska, N., & Schäfer, D. (2014). Is Crowdfunding Different?: Evidence on the Relation between Gender and Funding Success from a German Peer-to-Peer Lending Platform. *German Economic Review*, n/a-n/a. https://doi.org/10.1111/geer.12052

Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning. *Organizational Behavior and Human Decision Processes*, *42*(1), 88–110. https://doi.org/10.1016/0749-5978(88)90021-0

Bates, T. [T.], & Bradford, W. [W.] (2007). Traits and Performance of the Minority Venture-Capital Industry. *The ANNALS of the American Academy of Political and Social Science*, *613*(1), 95–107. https://doi.org/10.1177/0002716207303580

Bates, T. [Timothy], & Bradford, W. D. (2008). Venture-Capital Investment in Minority Business. *Journal of Money, Credit and Banking*, *40*(2-3), 489–504. https://doi.org/10.1111/j.1538-4616.2008.00123.x

Beaujean, A. A., & Grant, M. B. (2016). *Tutorial on Using Regression Models with Count Outcomes using R.* https://doi.org/10.7275/pj8c-h254

Beck, T., & Demirguc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & Finance*, *30*(11), 2931–2943. https://doi.org/10.1016/j.jbankfin.2006.05.009

Belleflamme, P., Lambert, T., & Schwienbacher, A. (2013). Individual crowdfunding practices. *Venture Capital*, *15*(4), 313–333. https://doi.org/10.1080/13691066.2013.785151

Benjamin, D. J., Berger, J., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., . . . Johnson, V. (2017). *Redefine statistical significance.* https://doi.org/10.31234/osf.io/mky9j

Berger, P. L., & Luckmann, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge* (Anchor books ed). *Anchor books: A589*. Garden City, N.Y.: Doubleday.

Bergman, B. (2020, June 4). 'Keep the American Dream Alive': Equity Crowdfunding Is Surging From an Appeal to Patriotism and Altruism. *Dot.LA*. Retrieved from https://dot.la/equity-crowdfunding-2646151982.html

Bernanke, B. S., & Rogoff, K. (2001). *The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?* Cambridge, Mass: MIT Press. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=62625

Bernstein, S., Korteweg, A., & Laws, K. (2014). *Attracting early stage investors: Evidence from a randomized field experiment*. *Working paper series / Rock Center for Corporate Governance: Vol. 185*. Stanford, Calif.: Univ. Rock Center for Corporate Governance. Retrieved from http://dx.doi.org/10.2139/ssrn.2432044 https://doi.org/10.2139/ssrn.2432044

Biletzki, Anat, & Matar (2019). Ludwig Wittgenstein. Retrieved from https://plato.stanford.edu/entries/wittgenstein/#Lan

Bird, A. (2013). "Thomas Kuhn",. Retrieved from URL = <http://plato.stanford.edu/archives/fall2013/entries/thomas-kuhn/>

Blanchflower, D., Levine, P., & Zimmerman, D. (1998). *Discrimination in the Small Business Credit Market*. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w6840

Blum, B. S., & Goldfarb, A. (2006). Does the internet defy the law of gravity? *Journal of International Economics*, *70*(2), 384–405. https://doi.org/10.1016/j.jinteco.2005.10.002

Boeuf, B., Darveau, J., & Legoux, R. (2014). Financing Creativity: Crowdfunding as a New Approach for Theatre Projects. *International Journal of Arts Management*. (Vol.16(3)), pp.33-48,79.

Bojanova, I. (2014). The Digital Revolution: What's on the Horizon? *IT Professional*, *16*(1), 8–12. https://doi.org/10.1109/MITP.2014.11

Borgatti, S. P., Brass, D. J., & Halgin, D. S. (2014). Social Network Research: Confusions, Criticisms, and Controversies. In D. J. Brass, G. Labianca, A. Mehra, D. S. Halgin, & S. P. Borgatti (Eds.), *Research in the Sociology of Organizations. Contemporary Perspectives on Organizational Social Networks* (Vol. 40, pp. 1–29). Emerald Group Publishing Limited. https://doi.org/10.1108/S0733-558X(2014)0000040001

Box, G. E. P. (1953). Non-Normality and Tests on Variances. *Biometrika*, *40*(3/4), 318. https://doi.org/10.2307/2333350

Breaugh, J. (2003). Effect Size Estimation: Factors to Consider and Mistakes to Avoid. *Journal of Management*, *29*(1), 79–97. https://doi.org/10.1016/S0149-2063(02)00221-0

Breznitz, S. M., & Noonan, D. S. (2020). Crowdfunding in a not-so-flat world. *Journal of Economic Geography*, *20*(4), 1069–1092. https://doi.org/10.1093/jeg/lbaa008

British Psychological Society (2017). Ethics Guidelines for Internet-mediated Research. Retrieved from https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-mediated%20Research%20(2017).pdf

Brooks, C. (2008). *Introductory Econometrics for Finance* (2nd ed.). Leiden: Cambridge University Press.

Bruhn, M., & Love, I. (2009). *The Economic Impact Of Banking The Unbanked: Evidence From Mexico*: The World Bank. https://doi.org/10.1596/1813-9450-4981

Brunnermeier, M. K. (2009). Deciphering the Liquidity and Credit Crunch 2007–2008. *Journal of Economic Perspectives*, *23*(1), 77–100. https://doi.org/10.1257/jep.23.1.77

Brush, G., Greene, P., Balachandra, L., & Davis, A. (2014). Diana Report Women Entrepreneurs 2014: Bridging the Gender Gap in Venture Capital. Retrieved from http://www.babson.edu/Academics/centers/blank-center/global-research/diana/Documents/diana-project-executive-summary-2014.pdf

Brynjolfsson, E., Hu, Y., & Rahman, M. S. (2009). Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition. *Management Science*, *55*(11), 1755–1765. https://doi.org/10.1287/mnsc.1090.1062

Burr, V. (2003). *Social constructionism* (2nd ed). London, New York: Routledge.

Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*. London: Heinemann.

Burtch, G., Ghose, A., & Wattal, S. (2013). An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets. *Information Systems Research*, *24*(3), 499–519. https://doi.org/10.1287/isre.1120.0468

Burtch, G., Ghose, A., & Wattal, S. (2014). Cultural differences and geography as determinants of online prosocial lending. *Management Information Systems : Mis Quarterly*, *38*(3), 773–794.

Cairncross, F. (1998). *The death of distance: How the communications revolution will change our lives* (2. impr). London: Orion Business Books.

Cameron, A. C., & Trivedi, P. K. [Pravin K.] (2010). *Microeconometrics using Stata* (Rev. ed.). *A Stata Press publication*. College Station, Tex.: Stata Press.

Cameron, A. C., & Trivedi, P. K. [P. K.] (2013). *Regression analysis of count data* (Second edition). *Econometric Society monographs: Vol. 53*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139013567

Campbell, D. T., & Stanley, J. C. (1967). *Experimental and quasi-experimental designs for research* (2. print; Reprinted from "Handbook of research on teaching"). Boston: Houghton Mifflin Comp.

Campennì, M., & Cecconi, F. (2019). Modeling the Dynamics of Reward-Based Crowdfunding Systems: An Agent-Based Model of Kickstarter. In F. Cecconi & M. Campennì (Eds.), *Computational Social Sciences. INFORMATION AND COMMUNICATION TECHNOLOGIES (ICT) AND ECONOMIC MODELING* (Vol. 33, pp. 91–116). [S.l.]: SPRINGER NATURE. https://doi.org/10.1007/978-3-030-22605-3_6

Carpenter, R. E., & Petersen, B. C. (2002). Is the Growth of Small Firms Constrained by Internal Finance? *Review of Economics and Statistics*, *84*(2), 298–309. https://doi.org/10.1162/003465302317411541

Carver, R. (1978). The Case Against Statistical Significance Testing. *Harvard Educational Review*, *48*(3), 378–399. https://doi.org/10.17763/haer.48.3.t490261645281841

Caseau, C., & Grolleau, G. (2020). Impact Investing: Killing Two Birds with One Stone? *Financial Analysts Journal*, *76*(4), 40–52. https://doi.org/10.1080/0015198X.2020.1779561

Cavalluzzo, K. S., Cavalluzzo, L. C., & Wolken, J. D. (1999). Competition, Small Business Financing, and Discrimination: Evidence from a New Survey. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.171733

Chaney, T. (2018). The Gravity Equation in International Trade: An Explanation. *Journal of Political Economy*, *126*(1), 150–177. https://doi.org/10.1086/694292

Chen, X.-P., Yao, X., & Kotha, S. (2009). Entrepreneur Passion And Preparedness In Business Plan Presentations: A Persuasion Analysis Of Venture Capitalists' Funding Decisions. *Academy of Management Journal*, *52*(1), 199–214. https://doi.org/10.5465/AMJ.2009.36462018

Chia, R. (1995). *"From modern to postmodern organizational analysis": Organization Studies, 16 (4)*, pp. 579–604.

Choi, J., & Bell, D. R. (2011). Preference Minorities and the Internet. *Journal of Marketing Research*, *48*(4), 670–682. https://doi.org/10.1509/jmkr.48.4.670

Clauss, T., Breitenecker, R. J., Kraus, S., Brem, A., & Richter, C. (2018). Directing the wisdom of the crowd: the importance of social interaction among founders and the crowd during crowdfunding campaigns. *Economics of Innovation and New Technology*, *27*(8), 709–729. https://doi.org/10.1080/10438599.2018.1396660

Cochran, W. G. (1947). Some Consequences When the Assumptions for the Analysis of Variance are not Satisfied. *Biometrics*, *3*(1), 22. https://doi.org/10.2307/3001535

Cohen, J. (1995). The earth is round (p <.05): Rejoinder. *American Psychologist*, *50*(12), 1103. https://doi.org/10.1037/0003-066X.50.12.1103

Coleman, J. S. (2000). *Foundations of social theory* (3. print). Cambridge, Mass.: Belknap Press of Harvard Univ. Press.

Coleman, S., & Robb, A. [Alicia] (2009). A comparison of new firm financing by gender: Evidence from the Kauffman Firm Survey data. *Small Business Economics*, *33*(4), 397–411. https://doi.org/10.1007/s11187-009-9205-7

Colombo, M. G., Franzoni, C., & Rossi-Lamastra, C. (2015). Internal Social Capital and the Attraction of Early Contributions in Crowdfunding. *Entrepreneurship Theory and Practice*, *39*(1), 75–100. https://doi.org/10.1111/etap.12118

Cooper, H. M. (1981). On the significance of effects and the effects of significance. *Journal of Personality and Social Psychology*, *41*(5), 1013–1018. https://doi.org/10.1037/0022-3514.41.5.1013

Cooper, I., & Kaplanis, E. (1994). Home Bias in Equity Portfolios, Inflation Hedging, and International Capital Market Equilibrium. *Review of Financial Studies*, *7*(1), 45–60. https://doi.org/10.1093/rfs/7.1.45

Corbetta, P. (2003). *Social research: Theory, methods and techniques*. London, Thousand Oaks, Calif.: SAGE Publications.

Coval, J. D., & Moskowitz, T. J. (1999). Home Bias at Home: Local Equity Preference in Domestic Portfolios. *The Journal of Finance*, *54*(6), 2045–2073. https://doi.org/10.1111/0022-1082.00181

Coxe, S. (2018). Effect size measures for nonlinear count regression models. Retrieved from https://stefanycoxe.weebly.com/uploads/2/4/3/7/24372766/coxe_asa_csp_020618.pdf

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to poisson regression and its alternatives. *Journal of Personality Assessment*, *91*(2), 121–136. https://doi.org/10.1080/00223890802634175

Cumming, D., & Dai, N. (2010). Local bias in venture capital investments. *Journal of Empirical Finance*, *17*(3), 362–380. https://doi.org/10.1016/j.jempfin.2009.11.001

Cumming, D., Meoli, M., & Vismara, S. (2019). Does equity crowdfunding democratize entrepreneurial finance? *Small Business Economics.* Advance online publication. https://doi.org/10.1007/s11187-019-00188-z

Currie, D. (2005). *Developing and applying study skills: Writing assignments, dissertations and management reports*. London: Chartered Institute of Personnel and Development.

Datar, R., & Garg, H. (2019). *Hands-On Exploratory Data Analysis with R* (1st edition): Packt Publishing.

Date, S. (2019). Negative Binomial Regression: A Step by Step Guide. Retrieved from https://towardsdatascience.com/negative-binomial-regression-f99031bb25b4

Davis, B. C., Hmieleski, K. M., Webb, J. W., & Coombs, J. E. (2017). Funders' positive affective reactions to entrepreneurs' crowdfunding pitches: The influence of perceived product creativity and entrepreneurial passion. *Journal of Business Venturing*, *32*(1), 90–106. https://doi.org/10.1016/j.jbusvent.2016.10.006

De', R., Pandey, N., & Pal, A. (2020). Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. *International Journal of Information Management*, *55*, 102171. https://doi.org/10.1016/j.ijinfomgt.2020.102171

Dekel, O., Amon, S., & Filatov, G. (2016). Venture Community: Democratisation of Entrepreneurship in Developing Economies: Published on Conference: ISMD international Society of Markets and Development, Lima, Peru.

Della Porta, D., & Keating, M. (2008). *Approaches and methodologies in the social sciences: A pluralist perspective*. Cambridge, N.Y.: Cambridge University Press.

Demidenko, E. (2016). The p-Value You Can't Buy. *The American Statistician*, *70*(1), 33–38. https://doi.org/10.1080/00031305.2015.1069760

Dinh, H. T., Mavridis, D. A., & Nguyen, H. B. (2010). *The binding constraint on firms' growth in developing countries*: The World Bank. https://doi.org/10.1596/1813-9450-5485

Dunbar, R. I. M. [R. I. M.] (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*(4), 681–694. https://doi.org/10.1017/S0140525X00032325

Dunbar, R.I.M. [R.I.M.] (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*(6), 469–493. https://doi.org/10.1016/0047-2484(92)90081-J

Dziuda, W., & Mondria, J. (2012). Asymmetric Information, Portfolio Managers, and Home Bias. *Review of Financial Studies*, *25*(7), 2109–2154. https://doi.org/10.1093/rfs/hhs063

Elston, D. M. (2018). Sample size. *Journal of the American Academy of Dermatology*, *79*(4), 635. https://doi.org/10.1016/j.jaad.2017.11.006

European Comission (2020). An SME Strategy for a sustainable and digital Europe. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0103

Ferrando, A., Pal, R., & Durante, E. (2019). Financing and obstacles for high growth enterprises: The European case.

Fisher, R. A. [R. A.] (1925). Theory of Statistical Estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, *22*(5), 700–725. https://doi.org/10.1017/S0305004100009580

Fisher, R. A. [Ronald A.] (1973). *Statistical methods and scientific inference* (3. ed., rev. and enl). New York, N.Y.: Hafner Press.

Florida, R. L., & King, K. M. (2016). *Rise of the global startup city: The geography of venture capital investment in cities and metros across the globe*. Toronto, ON: Martin Prosperity Institute Rotman School of Management University of Toronto. Retrieved from http://martinprosperity.org/content/rise-of-the-global-startup-city/

Forman, C., Ghose, A., & Goldfarb, A. (2009). Competition Between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where You Live. *Management Science*, *55*(1), 47–57. https://doi.org/10.1287/mnsc.1080.0932

French, K., & Poterba, J. (1991). *Investor Diversification and International Equity Markets*. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w3609

Frey, B. B. (2018). Pragmatic Paradigm. In B. B. Frey (Ed.), *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation* (1st ed.). Thousand Oaks: SAGE Publications. https://doi.org/10.4135/9781506326139.n534

Fried, V. H., & Hisrich, R. D. (1994). Toward a Model of Venture Capital Investment Decision Making. *Financial Management*, *23*(3), 28. https://doi.org/10.2307/3665619

Frydrych, D., Bock, A. J., Kinder, T., & Koeck, B. (2014). Exploring entrepreneurial legitimacy in reward-based crowdfunding. *Venture Capital*, *16*(3), 247–269. https://doi.org/10.1080/13691066.2014.916512

Gallemore, C., Nielsen, K. R., & Jespersen, K. (2019). The uneven geography of crowdfunding success: Spatial capital on Indiegogo. *Environment and Planning a: Economy and Space*, *51*(6), 1389–1406. https://doi.org/10.1177/0308518X19843925

Gelman, A., & Stern, H. (2006). The Difference Between “Significant” and “Not Significant” is not Itself Statistically Significant. *The American Statistician*, *60*(4), 328–331. https://doi.org/10.1198/000313006X152649

Gerber, E. M., & Hui, J. (2013). Crowdfunding. *ACM Transactions on Computer-Human Interaction*, *20*(6), 1–32. https://doi.org/10.1145/2530540

Gergen, K. J. (1992). *"Organiztion theory in the postmodern era": In M. Reed and M. Hughes, ed., Rethinking Organizations. London: Sage Publications*, p. 207.

Giannetti, M., & Laeven, L. (2012). The flight home effect: Evidence from the syndicated loan market during financial crises. *Journal of Financial Economics*, *104*(1), 23–43. https://doi.org/10.1016/j.jfineco.2011.12.006

GIIN (2019). What You Need to Know about Impact Investing: Global Impact Investing Network.

Glassman, D. A., & Riddick, L. A. (2001). What causes home asset bias and how should it be measured? *Journal of Empirical Finance*, *8*(1), 35–54. https://doi.org/10.1016/S0927-5398(00)00026-8

Gompers, P. A., & Lerner, J. [Joshua] (2004). *The venture capital cycle* (2. ed.). Cambridge, Mass.: MIT Press.

Google Webmaster Blog (2019, July 15). Formalizing the Robots Exclusion Protocol Specification. Retrieved from https://webmasters.googleblog.com/2019/07/rep-id.html

Gorman, M., & Sahlman, W. A. (1989). What do venture capitalists do? *Journal of Business Venturing*, *4*(4), 231–248. https://doi.org/10.1016/0883-9026(89)90014-1

Gorton, G. (2010). *Slapped by the Invisible Hand: The Panic of 2007*: Oxford University Press, USA.

Govindaraju, V., Raghavan, V., & Rao, C. R. (2015). *Big Data Analytics*. *Handbook of Statistics: v.33*. Burlington: Elsevier Science. Retrieved from http://gbv.eblib.com/patron/FullRecord.aspx?p=2122443

Graham, J. R., Harvey, C. R., & Huang, H. (2009). Investor Competence, Trading Frequency, and Home Bias. *Management Science*, *55*(7), 1094–1106. https://doi.org/10.1287/mnsc.1090.1009

Grasso, M., Klicperová-Baker, M., Koos, S., Kosyakova, Y., Petrillo, A., & Vlase, I. (2021). The impact of the coronavirus crisis on European societies. What have we learnt and where do we go from here? – Introduction to the COVID volume. *European Societies*, *23*(sup1), S2-S32. https://doi.org/10.1080/14616696.2020.1869283

Grauer, R. R., & Hakansson, N. H. (1987). Gains from international diversification: 1968 - 85 returns on portfolios of stocks and bonds. *The Journal of Finance : The Journal of the American Finance Association*, *45*.

Greene, P. G., Brush, C. G. [C. G.], Balachandra, L., & Davis, A. E. (2014). Diana Report. Women Entrepreneurs 2014: Bridging the Gender Gap in Venture Capital.

Greene, P. G. [Patricia G.], Brush, C. G. [Candida G.], Hart, M. M., & Saparito, P. (2001). Patterns of venture capital funding: Is gender a factor? *Venture Capital*, *3*(1), 63–83. https://doi.org/10.1080/13691060118175

Greene, W. (2008). Functional forms for the negative binomial model for count data. *Economics Letters*, *99*(3), 585–590. https://doi.org/10.1016/j.econlet.2007.10.015

Guenther, C., Johan, S., & Schweizer, D. (2018). Is the crowd sensitive to distance?—how investment decisions differ by investor type. *Small Business Economics*, *50*(2), 289–305. https://doi.org/10.1007/s11187-016-9834-6

Guiso, L., Sapienza, P., & Zingales, L. (2002). *Does Local Financial Development Matter?* Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w8923

Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural Biases in Economic Exchange? *Quarterly Journal of Economics*, *124*(3), 1095–1131. https://doi.org/10.1162/qjec.2009.124.3.1095

Guo, L., Guo, D., Wang, W., Wang, H., & Wu, Y. (2018). Distance Diffusion of Home Bias for Crowdfunding Campaigns between Categories: Insights from Data Analytics. *Sustainability*, *10*(4), 1251. https://doi.org/10.3390/su10041251

Harrison, R. T., & Mason, C. M. (2007). Does Gender Matter?: Women Business Angels and the Supply of Entrepreneurial Finance. *Entrepreneurship Theory and Practice*, *31*(3), 445–472. https://doi.org/10.1111/j.1540-6520.2007.00182.x

Harvey, L. A. (2014). Statistical power calculations reflect our love affair with P-values and hypothesis testing: Time for a fundamental change. *Spinal Cord*, *52*(1), 2. https://doi.org/10.1038/sc.2013.117

Hatcher, L. (2013). *Advanced statistics in research: Reading, understanding, and writing up data analysis results*. Saginaw, MI: ShadowFinch Media LLC.

Hecht, J. (2020, August 31). Is Crowdfunding A Good Option For Businesses During The Pandemic? *Forbes*. Retrieved from https://www.forbes.com/sites/jaredhecht/2020/08/31/is-crowdfunding-a-good-option-for-businesses-during-the-pandemic/?sh=615e341765f0

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating Trade Flows: Trading Partners and Trading Volumes *Quarterly Journal of Economics*, *123*(2), 441–487. https://doi.org/10.1162/qjec.2008.123.2.441

Hertwig, R. (2012). Psychology. Tapping into the wisdom of the crowd--with confidence. *Science (New York, N.Y.)*, *336*(6079), 303–304. https://doi.org/10.1126/science.1221403

Higgins, J. P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*: Wiley. https://doi.org/10.1002/9781119536604

Hilbe, J. M. (2011). *Negative Binomial Regression*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511973420

Hill, C. R., & Thompson, B. (2005). Computing and Interpreting Effect Sizes. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research. Higher Education: Handbook of Theory and Research* (Vol. 19, pp. 175–196). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2456-8_5

Hoffmann, A. O.I., Post, T., & Pennings, J. M.E. (2013). Individual investor perceptions and behavior during the financial crisis. *Journal of Banking & Finance*, *37*(1), 60–74. https://doi.org/10.1016/j.jbankfin.2012.08.007

Holmquist, C., & Carter, S. (2009). The Diana Project: Pioneering women studying pioneering women. *Small Business Economics*, *32*(2), 121–128. https://doi.org/10.1007/s11187-008-9151-9

Hornuf, L., & Schmitt, M. (2016). Does a Local Bias Exist in Equity Crowdfunding? The Impact of Investor Types and Portal Design. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2801170

Hornuf, L., & Schwienbacher, A. (2018). Market mechanisms and funding dynamics in equity crowdfunding. *Journal of Corporate Finance*, *50*, 556–574. https://doi.org/10.1016/j.jcorpfin.2017.08.009

Hortaçsu, A., Martínez-Jerez, F. A., & Douglas, J. (2009). The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre. *American Economic Journal: Microeconomics*, *1*(1), 53–74. https://doi.org/10.1257/mic.1.1.53

Hryniewicz, O. (2018). Statistical properties of the fuzzy p-value. *International Journal of Approximate Reasoning*, *93*, 544–560. https://doi.org/10.1016/j.ijar.2017.12.003

Hsu, D. H. (2004). What Do Entrepreneurs Pay for Venture Capital Affiliation? *The Journal of Finance*, *59*(4), 1805–1844. https://doi.org/10.1111/j.1540-6261.2004.00680.x

Huang, X. J., & Hoeber, O. (2010). *Ieee/wic/acm International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010: Joint conference] ; Aug. 31, 2010 - Sept. 3, 2010, Toronto, Ontario, Canada ; proceedings ; [including workshop papers*. Piscataway, NJ: IEEE. Retrieved from https://www.computer.org/csdl/proceedings/wi-iat/2010/4191/03/index.html

Huberman, G. (2001). Familiarity Breeds Investment. *Review of Financial Studies*, *14*(3), 659–680. https://doi.org/10.1093/rfs/14.3.659

Hui, J., Gerber, E. M., & Greenberg, M. (2012). Easy Money? The Demands of Crowdfunding Work.

Hwang, V., Desai, S., & Baird, R. (2019). Access to Capital for Entrepreneurs: Removing Barriers.

International Finance Corporation (2017). MSME Finance Gap: Assessment of the shortfalls and opportunities in financing micro, small and medium enterprises in emerging markets. Retrieved from https://openknowledge.worldbank.org/handle/10986/28881

International Labour Organization (2020). ILO Monitor: Covid-19 and the world of work. Second edition: Updated estimates and analysis. Retrieved from https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/documents/briefingnote/wcms_740877.pdf

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*, *108*(3), 577–598. https://doi.org/10.2307/2118401

Jeffers, J. N. R. (1994). The Importance of Exploratory Data Analysis before the Use of Sophisticated Procedures. *Biometrics*, *50*(3), 881. https://doi.org/10.2307/2532803

Johnson, P., & Duberley, J. (2000). *Understanding management research: An introduction to epistemology*. London, Thousand Oaks: SAGE Publications.

Kahneman, D. (Ed.) (2008). *Judgment under uncertainty: Heuristics and biases* (24. printing). Cambridge: Cambridge Univ. Press.

Kahneman, D. (2013). *Thinking, fast and slow*. London: Penguin.

Kampenes, V. B., Dybå, T., Hannay, J. E., & Sjøberg, D. I.K. (2007). A systematic review of effect size in software engineering experiments. *Information and Software Technology*, *49*(11-12), 1073–1086. https://doi.org/10.1016/j.infsof.2007.02.015

Kang, L., Jiang, Q., & Tan, C.-H. (2017). Remarkable advocates: An investigation of geographic distance and social capital for crowdfunding. *Information & Management*, *54*(3), 336–348. https://doi.org/10.1016/j.im.2016.09.001

Kaplan, R. M., Chambers, D. A., & Glasgow, R. E. (2014). Big data and large sample size: A cautionary note on the potential for bias. *Clinical and Translational Science*, *7*(4), 342–346. https://doi.org/10.1111/cts.12178

Karlsson, A., & Nordén, L. (2007). Home sweet home: Home bias and international diversification among individual investors. *Journal of Banking & Finance*, *31*(2), 317–333. https://doi.org/10.1016/j.jbankfin.2006.04.005

Karney, C. F. F. (2013). Algorithms for geodesics. *Journal of Geodesy*, *87*(1), 43–55. https://doi.org/10.1007/s00190-012-0578-z

Katzner, D. W. (2017). *Models, Mathematics, and Methodology in Economic Explanation*: Cambridge University Press. https://doi.org/10.1017/9781108291422

Kenney, M. (1999). Technology, entrepreneurship and path dependence: Industrial clustering in Silicon Valley and Route 128. *Industrial and Corporate Change*, *8*(1), 67–103. https://doi.org/10.1093/icc/8.1.67

Keren, G., & Lewis, C. (Eds.) (1993). *A handbook for data analysis in the behavioral sciences: Statistical issues*. Hillsdale, N.J: L. Erlbaum Associates. Retrieved from http://www.loc.gov/catdir/enhancements/fy0745/92049442-d.html

Kickstarter (2010). Trends in Pricing and Duration. Retrieved from https://www.kickstarter.com/blog/trends-in-pricing-and-duration?lang=de

Kickstarter (2016a). Campus — Kickstarter: Success rate depending on project's country of origin? Retrieved from https://www.kickstarter.com/campus/questions/success-rate-depending-on-projects-country-of-origin

Kickstarter (2016b). The First 100,000 Funded Kickstarter Projects in 100 Numbers. Retrieved from https://www.kickstarter.com/blog/the-first-100000-funded-kickstarter-projects-in-100-numbers?lang=de

Kickstarter (2018). Kickstarter Privacy Policy. Retrieved from https://www.kickstarter.com/privacy

Kickstarter (2019a). How to Get Featured on Kickstarter. Retrieved from https://www.kickstarter.com/blog/how-to-get-featured-on-kickstarter?lang=de

Kickstarter (2019b). Our Rules — Kickstarter. Retrieved from https://www.kickstarter.com/rules

Kickstarter (2019c). What are the basics? Retrieved from https://help.kickstarter.com/hc/en-us/articles/115005028514-What-are-the-basics-

Kickstarter (2020). Kickstarter-Statistiken — Kickstarter. Retrieved from https://www.kickstarter.com/help/stats?lang=de

Kickstarter (2021). How to Get Featured on Kickstarter — Kickstarter. Retrieved from https://www.kickstarter.com/blog/how-to-get-featured-on-kickstarter

Kickstarter Support (2020, November 30). How does Advanced Search work? Retrieved from https://help.kickstarter.com/hc/en-us/articles/115005066253-How-does-Advanced-Search-work-

Kim, H., & Kim, J. (2017). Geographic proximity between lender and borrower: how does it affect crowdfunding? *Review of Accounting and Finance*, *16*(4), 462–477. https://doi.org/10.1108/RAF-02-2016-0017

Kim, K., & Hann, I.-H. (2013). Does Crowdfunding Democratize Access to Capital?: A Geographical Analysis. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2334590

Kirk, R. E. (1996). Practical Significance: A Concept Whose Time Has Come. *Educational and Psychological Measurement*, *56*(5), 746–759. https://doi.org/10.1177/0013164496056005002

Kish, Z., & Fairbairn, M. (2018). Investing for profit, investing for impact: Moral performances in agricultural investment projects. *Environment and Planning a: Economy and Space*, *50*(3), 569–588. https://doi.org/10.1177/0308518X17738253

Klaes, M. (Ed.) (2017). *Identifying market and regulatory obstacles to cross-border development of crowdfunding in the EU: Final report*. Luxembourg: Publications Office of the European Union. https://doi.org/10.2874/65957

Klapper, L., Amit, R., Guillen, M. F., & Quesada, J. M. (2007). *Entrepreneurship And Firm Formation Across Countries*: The World Bank. https://doi.org/10.1596/1813-9450-4313

Kortum, S., & Lerner, J. [Josh] (2000). Assessing the Contribution of Venture Capital to Innovation. *The RAND Journal of Economics*, *31*(4), 674. https://doi.org/10.2307/2696354

Kraus, S., Richter, C., Brem, A., Cheng, C.-F., & Chang, M.-L. (2016). Strategies for reward-based crowdfunding campaigns. *Journal of Innovation & Knowledge*, *1*(1), 13–23. https://doi.org/10.1016/j.jik.2016.01.010

Krishnan, K., Lohmer, L., & Wang, P. (2015). What Do Crowds Really Want? Crowdfunding and Early Product Access. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2684223

Kühberger, A., Fritz, A., Lermer, E., & Scherndl, T. (2015). The significance fallacy in inferential statistics. *BMC Research Notes*, *8*, 84. https://doi.org/10.1186/s13104-015-1020-4

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed). Chicago, IL: University of Chicago Press.

Kuppuswamy, V., & Bayus, B. L. (2013). Crowdfunding Creative Ideas: The Dynamics of Project Backers in Kickstarter. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2234765

Lai, S., & Teo, M. (2008). Home-Biased Analysts in Emerging Markets. *Journal of Financial and Quantitative Analysis*, *43*(03), 685. https://doi.org/10.1017/S0022109000004257

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lambert, T., & Schwienbacher, A. (2010). An Empirical Analysis of Crowdfunding.

Langley, P., & Leyshon, A. (2017). Capitalizing on the crowd: The monetary and financial ecologies of crowdfunding. *Environment and Planning a: Economy and Space*, *49*(5), 1019–1039. https://doi.org/10.1177/0308518X16687556

Lansiti, M., & Lakhani, K. (2017). THE TRUTH ABOUT BLOCKCHAIN. (95(1)), 119–127.

Last, J. M., & Porta, M. (2018). *A dictionary of public health* (2. edition). *Oxford Reference*. Oxford, New York: Oxford University Press. Retrieved from http://rzblx10.uni-regensburg.de/dbinfo/detail.php?bib_id=alle&colors=&ocolors=&lett=fs&tid=0&titel_id=7606 https://doi.org/10.1093/acref/9780191844386.001.0001

Leeper, T. J. (2017). Interpreting Regression Results using Average Marginal Effects with R ’ s margins. Retrieved from https://cran.r-project.org/web/packages/margins/vignettes/TechnicalDetails.pdf

Leland, J. (2020). Are People Still Supporting Projects? Our Data Says Yes. Retrieved from https://www.kickstarter.com/articles/covid-19-coronavirus-kickstarter-impact-pledges-projects-launch

Lewis, K. K. (1999). Trying to Explain Home Bias in Equities and Consumption. *Journal of Economic Literature*, *37*(2), 571–608. https://doi.org/10.1257/jel.37.2.571

Lin, M., Lucas, H. C., & Shmueli, G. (2013). Research Commentary —Too Big to Fail: Large Samples and the p -Value Problem. *Information Systems Research*, *24*(4), 906–917. https://doi.org/10.1287/isre.2013.0480

Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending. *Management Science*, *59*(1), 17–35. https://doi.org/10.1287/mnsc.1120.1560

Lin, M., & Viswanathan, S. (2016). Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market: An Empirical Study of an Online Crowd Funding Market. *Management Science*, *62*(5), 1393–1414. https://doi.org/10.1287/mnsc.2015.2206

Lin, Y., Boh, W. F., & Goh, K. H. (2014). How Different are Crowdfunders?: Examining Archetypes of Crowdfunders and Their Choice of Projects. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2397571

Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of Assumption Violations Revisited: A Quantitative Review of Alternatives to the One-Way Analysis of Variance "F" Test. *Review of Educational Research*, *66*(4), 579. https://doi.org/10.2307/1170654

Long, J. S. (1997). *Regression models for categorical and limited dependent variables* ([Nachdr.]). *Advanced quantitative techniques in the social sciences: Vol. 7*. Thousand Oaks, Calif.: Sage Publ.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, *70*(3), 151–159. https://doi.org/10.1037/h0026141

MacInnes, J. (2017). *An Introduction to secondary data analysis with IBM SPSS statistics* (First published 2017). Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: Sage.

Manyika, J., & Roxburgh, C. (2011). The great transformer: The impact of the internet on economic growth and prosperity. Retrieved from https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-great-transformer

Marom, D., Robb, A. [Alicia], & Sade, O. (2014). Gender Dynamics in Crowdfunding (Kickstarter): Evidence on Entrepreneurs, Investors, Deals and Taste Based Discrimination. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2442954

Marom, D., & Sade, O. (2013). Are the Life and Death of a Young Start-Up Indeed in the Power of the Tongue?: Lessons from Online Crowdfunding Pitches. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2255707

Mason, C. (2007). Venture Capital: A Geographical Perspective. In H. Landström (Ed.), *Handbook of Research on Venture Capital*. Edward Elgar Publishing. https://doi.org/10.4337/9781847208781.00010

Massolution (2015). *2015CF: The Crowdfunding Industry Report*. Los Angeles, CA.

Mather, K. (1951). R. A. Fisher's Statistical Methods for Research Workers : An Appreciation. *Journal of the American Statistical Association*, *46*(253), 51–54. https://doi.org/10.1080/01621459.1951.10500767

McAuley, J., Duberley, J., & Johnson, P. (2014). *Organization theory: Challenges and perspectives* (Second edition). Harlow: Pearson education limited.

McCullagh, P., & Nelder, J. A. (2018). *Generalized Linear Models* (2nd ed.). *Chapman and Hall/CRC Monographs on Statistics and Applied Probability: v.37*. Boca Raton: Routledge. Retrieved from https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5631551

McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. *ProQuest Ebook Central*. Boca Raton, London, New York: CRC Press. Retrieved from http://gbv.eblib.com/patron/FullRecord.aspx?p=4648054

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Meehl, P. E. (1990). Why Summaries of Research on Psychological Theories are Often Uninterpretable. *Psychological Reports*, *66*(1), 195–244. https://doi.org/10.2466/pr0.1990.66.1.195

Mendes-Da-Silva, W., Rossoni, L., Conte, B. S., Gattaz, C. C., & Francisco, E. R. (2016). The impacts of fundraising periods and geographic distance on financing music production via crowdfunding in Brazil. *Journal of Cultural Economics*, *40*(1), 75–99. https://doi.org/10.1007/s10824-015-9248-3

Miller, C. (2010). Why So Few Women in Silicon Valley? Retrieved from http://www.nytimes.com/2010/04/18/technology/18women.html

Mohajeri, K., Mesgari, M., & Lee, A. S. (2020). When Statistical Significance Is Not Enough: Investigating Relevance, Practical Significance, and Statistical Significance. *MIS Quarterly*, *44*(2), 525–559. https://doi.org/10.25300/MISQ/2020/13932

Mollick, E. (2013). Swept Away by the Crowd?: Crowdfunding, Venture Capital, and the Selection of Entrepreneurs. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2239204

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing*, *29*(1), 1–16. https://doi.org/10.1016/j.jbusvent.2013.06.005

Mollick, E., & Kuppuswamy, V. (2014). After the Campaign: Outcomes of Crowdfunding. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2376997

Mollick, E., & Nanda, R. (2016). Wisdom or Madness? Comparing Crowds with Expert Evaluation in Funding the Arts. *Management Science*, *62*(6), 1533–1553. https://doi.org/10.1287/mnsc.2015.2207

Mollick, E., & Robb, A. [A.] (2016). Democratizing Innovation and Capital Access: The Role of Crowdfunding. *California Management Review*, *58*(2), 72–87. https://doi.org/10.1525/cmr.2016.58.2.72

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2015). *Introduction to Linear Regression Analysis*: Wiley. Retrieved from https://books.google.de/books?id=27kOCgAAQBAJ

Morgan, D. L. (2007). Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research*, *1*(1), 48–76. https://doi.org/10.1177/2345678906292462

Moritz, A., Block, J. H., & Lutz, E. (2014). Investor Communication in Crowdfunding: A Qualitative-Empirical Study. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2462282

Myers, D. G. (2009). *Social psychology* (10. ed.). New York: McGraw-Hill.

Nebehay, S. (2013, February 28). Higher cancer risk after Fukushima nuclear disaster: WHO. *Reuters Media*. Retrieved from https://www.reuters.com/article/us-japan-nuclear-cancer-idUSBRE91R0D420130228

Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7. ed., internat. ed.). Boston, Mass.: Pearson.

Newman, R., Chang, V., Walters, R. J., & Wills, G. B. (2016). Web 2.0—The past and the future. *International Journal of Information Management*, *36*(4), 591–598. https://doi.org/10.1016/j.ijinfomgt.2016.03.010

Neyman, J., & Pearson, E. S. (1928). On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference: Part I. *Biometrika*, *20A*(1/2), 175. https://doi.org/10.2307/2331945

Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Niemand, T., Angerer, M., Thies, F., Kraus, S., & Hebenstreit, R. (2018). Equity crowdfunding across borders: a conjoint experiment. *International Journal of*

Entrepreneurial Behavior & Research, *24*(4), 911–932. https://doi.org/10.1108/IJEBR-07-2017-0256

Noordzij, M., van Diepen, M., Caskey, F. C., & Jager, K. J. (2017). Relative risk versus absolute risk: One cannot be interpreted without the other. *Nephrology, Dialysis, Transplantation : Official Publication of the European Dialysis and Transplant Association - European Renal Association*, *32*(suppl_2), ii13-ii18. https://doi.org/10.1093/ndt/gfw465

Noton, C. (2015). On the size of home bias. *Applied Economics*, *47*(2), 123–128. https://doi.org/10.1080/00036846.2014.964832

Nuzzo, R. (2014). Scientific method: Statistical errors. *Nature*, *506*(7487), 150–152. https://doi.org/10.1038/506150a

Oberheim, Eric, Hoyningen-Huene, & Paul (2015). The Incommensurability of Scientific Theories. Retrieved from http://plato.stanford.edu/entries/incommensurability/

OECD (2009). *National Accounts of OECD Countries 2009, Volume I, Main Aggregates*: OECD. Retrieved from http://www.oecd.org/berlin/44681640.pdf https://doi.org/10.1787/na_vol_1-2009-en-fr

Parker, M. (1992). Post-Modern Organizations or Postmodern Organization Theory? *Organization Studies*, *13*(1), 1–17. https://doi.org/10.1177/017084069201300103

Patzer, G. L. (1995). *Using secondary data in marketing research: United States and worldwide*. Westport, Conn: Quorum Books.

Payne, E. H., Gebregziabher, M., Hardin, J. W., Ramakrishnan, V., & Egede, L. E. (2018). An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data. *Communications in Statistics: Simulation and Computation*, *47*(6), 1722–1738. https://doi.org/10.1080/03610918.2017.1323223

Peikoff, L. (1971). *Nazism and Subjectivism*. in The Objectivist.

Pernet, C. (2015). Null hypothesis significance testing: A short tutorial. *F1000Research*, *4*, 621. https://doi.org/10.12688/f1000research.6963.3

Petersen, M. A., & Rajan, R. G. (2002). Does Distance Still Matter? The Information Revolution in Small Business Lending. *The Journal of Finance*, *57*(6), 2533–2570. https://doi.org/10.1111/1540-6261.00505

Pharoah, P. (2007). How not to interpret a P value? *Journal of the National Cancer Institute*, *99*(4), 332–333. https://doi.org/10.1093/jnci/djk061

Poldrack, R. A. (2018). Statistical Thinking for the 21st Century.

Porter, M. E. (2000). Location, Competition, and Economic Development: Local Clusters in a Global Economy. *Economic Development Quarterly*, *14*(1), 15–34. https://doi.org/10.1177/089124240001400105

Powell, W. W., Koput, K. W., Bowie, J. I., & Smith-Doerr, L. (2002). The Spatial Clustering of Science and Capital: Accounting for Biotech Firm-Venture Capital Relationships. *Regional Studies*, *36*(3), 291–305. https://doi.org/10.1080/00343400220122089

Pratt, D. D. (1998). *Five perspectives on teaching in adult and higher education* (Original ed). Malabar, Fla.: Krieger Pub. Co.

Privacy Policy — Kickstarter (2021, July 9). Retrieved from https://www.kickstarter.com/privacy/jun2016

Punch, K. F. (2014). *Introduction to social research: Quantitative and qualitative approaches* (Third edition). Los Angeles, London, Washington DC: Sage.

PwC, & CB Insights (2019). MoneyTree Report: Q4 2018.

Qiu, C. (2013). Issues in Crowdfunding: Theoretical and Empirical Investigation on Kickstarter. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2345872

Rand, A., & Binswanger, H. (1988). *The Ayn Rand lexicon: Objectivism from A to Z*. New York, N.Y.: New American Library.

Rao, D. (2013). Why 99.95% Of Entrepreneurs Should Stop Wasting Time Seeking Venture Capital. Retrieved from https://www.forbes.com/sites/dileeprao/2013/07/22/why-99-95-of-entrepreneurs-should-stop-wasting-time-seeking-venture-capital/#c29ba9f46eb5

Robertson, E. N., & Wooster, R. B. (2015). Crowdfunding as a Social Movement: The Determinants of Success in Kickstarter Campaigns. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2631320

Rorty, R. (1982). *Consequences of Pragmatism*: Minneapolis: University of Minnesota Press.

Rose, N. (2019). How the World Regulates Equity Crowdfunding | The Regulatory Review. Retrieved from https://www.theregreview.org/2019/06/26/rose-how-world-regulates-equity-crowdfunding/

Rosenau, P. V. (1992). *Post-modernism and the social sciences: Insights, inroads, and intrusions*. *Princeton paperbacks*. Princeton, N.J.: Princeton University Press.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2011). *Contrasts and Effect Sizes in Behavioral Research*: Cambridge University Press. https://doi.org/10.1017/CBO9780511804403

Roth, R. (1998). Did Class Matter in American Politics?: The Importance of Exploratory Data Analysis (EDA). *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *31*(1), 5–25. https://doi.org/10.1080/01615449809600090

Rowland, P. F. (1999). Transaction costs and international portfolio diversification. *Journal of International Economics*, *49*(1), 145–170. https://doi.org/10.1016/S0022-1996(98)00059-2

Ruhnau, J. (2019). What happened to Crowdfunding. Retrieved from https://www.n-tv.de/ratgeber/Was-aus-Crowdfunding-geworden-ist-article21365313.html

Russell A. Poldrack (2018). Statistical Thinking for the 21st Century. Retrieved from http://web.stanford.edu/group/poldracklab/statsthinking21/

Sakia, R. M. (1992). The Box-Cox Transformation Technique: A Review. *The Statistician*, *41*(2), 169. https://doi.org/10.2307/2348250

Salkind, N. (Ed.) (2010). *Encyclopedia of Research Design*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. https://doi.org/10.4135/9781412961288

Salkind, N. J. (2010). Exploratory Data Analysis. In N. Salkind (Ed.), *Encyclopedia of Research Design*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n143

Saunders, M., Lewis, P., & Thornhill, A. (2012). *Research methods for business students* (6. ed.). *Always learning*. Harlow: Pearson. Retrieved from http://lib.myilibrary.com/detail.asp?id=385301

Sawilowsky, S. S. (2009). New Effect Size Rules of Thumb. *Journal of Modern Applied Statistical Methods*, *8*(2), 597–599. https://doi.org/10.22237/jmasm/1257035100

Sawyer, A. G., & Peter, J. P. (1983). The Significance of Statistical Significance Tests in Marketing Research. *Journal of Marketing Research*, *20*(2), 122–133. https://doi.org/10.1177/002224378302000203

Schneble, C. O., Elger, B. S., & Shaw, D. (2018). The Cambridge Analytica affair and Internet-mediated research. *EMBO Reports*, *19*(8). https://doi.org/10.15252/embr.201846579

Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. *Nature*, *515*(7525), 9. https://doi.org/10.1038/515009a

Shane, S., & Stuart, T. (2002). Organizational Endowments and the Performance of University Start-ups. *Management Science*, *48*(1), 154–170. https://doi.org/10.1287/mnsc.48.1.154.14280

Shapiro, C., & Varian, H. R. (2008). *Information rules: A strategic guide to the network economy* ([Nachdr.]). Boston, Mass.: Harvard Business School Press.

University Research Ethics Committee of Sheffiled Hallam University. (2017).

Sibanda, K., Hove-Sibanda, P., & Shava, H. (2018). The impact of SME access to finance and performance on exporting behaviour at firm level: A case of furniture manufacturing SMEs in Zimbabwe. *Acta Commercii*, *18*(1), 215. https://doi.org/10.4102/ac.v18i1.554

Sidhu, I., & Doyle, T. C. (2016). *The digital revolution: How connected digital innovations are transforming your industry, company, and career*. Old Tappan, New Jersey: Pearson Education Inc.

Silva, J. M. C. S., & Tenreyro, S. (2006). The Log of Gravity. *Review of Economics and Statistics*, *88*(4), 641–658. https://doi.org/10.1162/rest.88.4.641

Singh Chawla, D. (2017). Big names in statistics want to shake up much-maligned P value. *Nature*, *548*(7665), 16–17. https://doi.org/10.1038/nature.2017.22375

Smart, J. C. (Ed.) (2005). *Higher Education: Handbook of Theory and Research*. *Higher Education: Handbook of Theory and Research*. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2456-8

Sondhi, N. (2011). Jonathan Wilson, Essentials of Business Research—A Guide to Doing Your Research Project. New Delhi: SAGE Publications India Pvt Ltd, 2010, 316 pp. Rs 495 (ISBN: 978-81-321-0567-1[Pb]). *Global Business Review*, *12*(2), 343–344. https://doi.org/10.1177/097215091101200211

Sorenson, O. (2018). Social networks and the geography of entrepreneurship. *Small Business Economics*, *51*(3), 527–537. https://doi.org/10.1007/s11187-018-0076-7

Sorenson, O., Assenova, V., Li, G.-C., Boada, J., & Fleming, L. (2016). Expand innovation finance via crowdfunding. *Science (New York, N.Y.)*, *354*(6319), 1526–1528. https://doi.org/10.1126/science.aaf6989

Sorenson, O., & Stuart, T. E. (2001). Syndication Networks and the Spatial Distribution of Venture Capital Investments. *American Journal of Sociology*, *106*(6), 1546–1588. https://doi.org/10.1086/321301

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355. https://doi.org/10.2307/1882010

Stadler, F., Baier, K. E., Barone, F., Hempel, C. G., Körner, S., Mulder, H., . . . Stöltzner, M. (2003). *The Vienna Circle and Logical Empiricism* (Vol. 10). Dordrecht: Springer Netherlands. https://doi.org/10.1007/0-306-48214-2

Staněk, R. (2017, March). Home bias in sport betting: Evidence from Czech betting market: Judgment and Decision Making. (Vol.12(2)), pp.168-172.

Statista (2018). Leading crowdfunding platforms by funds raised 2016 | Statista. Retrieved from https://www.statista.com/statistics/757442/leading-crowdfunding-platforms-by-value-of-funds-raised/

Steigenberger, N. (2017). Why supporters contribute to reward-based crowdfunding. *International Journal of Entrepreneurial Behavior & Research*, *23*(2), 336–353. https://doi.org/10.1108/IJEBR-04-2016-0117

Stevenson, R. M., Kuratko, D. F., & Eutsler, J. (2019). Unleashing main street entrepreneurship: Crowdfunding, venture capital, and the democratization of new venture investments. *Small Business Economics*, *52*(2), 375–393. https://doi.org/10.1007/s11187-018-0097-2

Stiglitz, J. E. (1974). *The theory of "screening," education, and the distribution of income.*

Strong, N., & Xu, X. (2003). Understanding the Equity Home Bias: Evidence from Survey Data. *Review of Economics and Statistics*, *85*(2), 307–312. https://doi.org/10.1162/003465303765299837

Stuart, A. (1976). *Basic ideas of scientific sampling* (2. ed.). *Griffin's statistical monographs and courses: Vol. 4*. London: Griffin.

Stuart, T., & Sorenson, O. (2003). The geography of opportunity: Spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy*, *32*(2), 229–253. https://doi.org/10.1016/S0048-7333(02)00098-7

Stuart, T., & Sorenson, O. (2005). The evolution of venture capital investment networks. Federal Reserve Bank of Atlanta.

Stuart, T. E., & Sorenson, O. (2005). Social Networks and Entrepreneurship. In S. A. Alvarez, R. Agarwal, & O. Sorenson (Eds.), *International Handbook Series on Entrepreneurship. Handbook of Entrepreneurship Research* (Vol. 2, pp. 233–252). New York: Springer-Verlag. https://doi.org/10.1007/0-387-23622-8_11

Stuenkel, O. (2013). South Africas BRICS membership: A win-win situation? *African Journal of Political Science and International Relations*, *7*(7), 310–319. https://doi.org/10.5897/AJPSIR2013.0625

Teng, S.-Y., Yeh, M.-Y., & Chuang, K.-T. (2015). Toward Understanding the Mobile Social Properties. In J. Pei, F. Silvestri, & J. Tang (Eds.), *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and*

Mining 2015 - ASONAM '15 (pp. 266–269). New York, New York, USA: ACM Press. https://doi.org/10.1145/2808797.2809416

Tesar, L. L., & Werner, I. M. (1995). Home bias and high turnover. *Journal of International Money and Finance*, *14*(4), 467–492. https://doi.org/10.1016/0261-5606(95)00023-8

Thierer, A. D., Koopman, C., Hobson, A., & Kuiper, C. (2015). How the Internet, the Sharing Economy, and Reputational Feedback Mechanisms Solve the 'Lemons Problem'. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2610255

Thompson, B. (2002). “Statistical,” “Practical,” and “Clinical”: How Many Kinds of Significance Do Counselors Need to Consider? *Journal of Counseling & Development*, *80*(1), 64–71. https://doi.org/10.1002/j.1556-6678.2002.tb00167.x

Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools*, *44*(5), 423–432. https://doi.org/10.1002/pits.20234

Thompson, N., Wang, X., & Baskerville, R. (2021). Improving IS Practical Significance through Effect Size Measures. *Journal of Computer Information Systems*, 1–8. https://doi.org/10.1080/08874417.2020.1837036

Thorbecke, W. (2020). The Impact of the COVID-19 Pandemic on the U.S. Economy: Evidence from the Stock Market. *Journal of Risk and Financial Management*, *13*(10), 233. https://doi.org/10.3390/jrfm13100233

Tinbergen, J. (1962). *An Analysis of World Trade Flows*: Shaping the World Economy, New York, NY: Twentieth Century Fund.

Trusty, J., Thompson, B., & Petrocelli, J. V. (2004). Practical Guide for Reporting Effect Size in Quantitative Research in the Journal of Counseling & Development. *Journal of Counseling & Development*, *82*(1), 107–110. https://doi.org/10.1002/j.1556-6678.2004.tb00291.x

Tukey, J. W. (1977). *Exploratory data analysis*. *Addison-Wesley series in behavioral science*. Reading, Mass.: Addison-Wesley.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/BF00122574

UCLA (2019a). Poisson Regression | Stata Annotated Output. Retrieved from https://stats.idre.ucla.edu/stata/output/poisson-regression/

UCLA (2019b). Regression Models with Count Data. Retrieved from https://stats.idre.ucla.edu/stata/seminars/regression-models-with-count-data/

UNICEF (2019). UNICEF - Definitions. Retrieved from https://www.unicef.org/infobycountry/stats_popup7.html

Vacha-Haase, T., & Thompson, B. (2004). How to Estimate and Interpret Various Effect Sizes. *Journal of Counseling Psychology*, *51*(4), 473–481. https://doi.org/10.1037/0022-0167.51.4.473

Vasconcelos, G. (2013). The problems with UK equity crowdfunding and seed investment. Retrieved from https://www.syndicateroom.com/blog/what-is-the-problem-with-equity-crowdfunding-and-seed-investment

Vinturella, J., & Erickson, S. (2013). *Raising Entrepreneurial Capital, 2nd Edition* (2nd edition): Elsevier.

Vulkan, N., Åstebro, T., & Sierra, M. F. (2016). Equity crowdfunding: A new phenomena. *Journal of Business Venturing Insights*, *5*, 37–49. https://doi.org/10.1016/j.jbvi.2016.02.001

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA Statement on p -Values: Context, Process, and Purpose. *The American Statistician*, *70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Wellman, B., Salaff, J., Dimitrova, D., Garton, L., Gulia, M., & Haythornthwaite, C. (1996). Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community. *Annual Review of Sociology*, *22*(1), 213–238. https://doi.org/10.1146/annurev.soc.22.1.213

Wheat, R. E., Wang, Y., Byrnes, J. E., & Ranganathan, J. (2013). Raising money for scientific research through crowdfunding. *Trends in Ecology & Evolution*, *28*(2), 71–72. https://doi.org/10.1016/j.tree.2012.11.001

WHO (2020). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. Retrieved from https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Wiklund, J., & Shepherd, D. (2005). Entrepreneurial orientation and small business performance: a configurational approach. *Journal of Business Venturing*, *20*(1), 71–91. https://doi.org/10.1016/j.jbusvent.2004.01.001

Woodfield, K. (Ed.) (2018). *Advances in research ethics and integrity: volume 2*. *Ethics of online research*. Bingley: Emerald Publishing. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1577428 https://doi.org/10.1108/S2398-6018201802

World Bank (2012). World Development Report 2013.

World Bank (2013). IFC Jobs Study - Asessing Private Sector Contributions to Job Creation and Poverty Reduction. Retrieved from http://www.ifc.org/wps/wcm/connect/0fe6e2804e2c0a8f8d3bad7a9dd66321/IFC_FULL+JOB+STUDY+REPORT_JAN2013_FINAL.pdf?MOD=AJPERES

World Bank (2016a). *What's Holding Back the Private Sector in MENA?*: World Bank. https://doi.org/10.1596/24738

World Bank (Ed.) (2016b). *World Development Report 2016: Digital Dividends*: The World Bank. https://doi.org/10.1596/978-1-4648-0671-1

Yi, S. K. M., Steyvers, M., Lee, M. D., & Dry, M. J. (2012). The wisdom of the crowd in combinatorial problems. *Cognitive Science*, *36*(3), 452–470. https://doi.org/10.1111/j.1551-6709.2011.01223.x

Zane, L. J., & DeCarolis, D. M. (2016). Social networks and the acquisition of resources by technology-based new ventures. *Journal of Small Business & Entrepreneurship*, *28*(3), 203–221. https://doi.org/10.1080/08276331.2016.1162048

Zar, J. H. (1999). *Biostatistical analysis* (4th ed.). Upper Saddle River, N.J: Prentice Hall.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression Models for Count Data in R. *Journal of Statistical Software*, *27*(8). https://doi.org/10.18637/jss.v027.i08

Zhang, J., & Liu, P. (2012). Rational Herding in Microloan Markets. *Management Science*, *58*(5), 892–912. https://doi.org/10.1287/mnsc.1110.1459

Zhao, J., Wu, J., Liu, G., Tao, D., Xu, K., & Liu, C. (2014). Being rational or aggressive? A revisit to Dunbar's number in online social networks. *Neurocomputing*, *142*, 343–353. https://doi.org/10.1016/j.neucom.2014.04.025

Zook, M. A. (2002). Grounded capital: venture financing and the geography of the Internet industry, 1994-2000. *Journal of Economic Geography*, *2*(2), 151–177. https://doi.org/10.1093/jeg/2.2.151

Zucker, L., Darby, M., & Brewer, M. (1994). *Intellectual Capital and the Birth of U.S. Biotechnology Enterprises*. Cambridge, MA. https://doi.org/10.3386/w4653

Zvilichovsky, D., Inbar, Y., & Barzilay, O. (2013). Playing Both Sides of the Market: Success and Reciprocity on Crowdfunding Platforms. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.2304101

Appendix

Appendix A. Country Capital Coordinates used for Distance Estimation

Country	Capital Coordinates
Afghanistan	[34.5260131, 69.1776476],
Aland Islands	[60.0900, 19.94000],
Albania	[41.3279457, 19.8185323],
Algeria	[+36.7755, +3.0597],
American Samoa	[-14.2754786, -170.7048298],
Andorra	[42.5069391, 1.5212467],
Angola	[-8.8271656, 13.2436665],
Anguilla	[18.2145861, -63.0517759],
Antarctica	[-82.86275189, 135],
Antigua and Barbuda	[17.1184569, -61.8448509],
Argentina	[-34.6075616, -58.437076],
Armenia	[40.1776121, 44.5125849],
Aruba	[12.5268736, -70.0356845],
Australia	[-35.2975906, 149.1012676],
Austria	[48.2083537, 16.3725042],
Azerbaijan	[40.3759646, 49.8325808],
Bahrain	[26.2235041, 50.5822436],
Bangladesh	[23.7593572, 90.3788136],
Barbados	[13.0977832, -59.6184184],
Belarus	[53.902334, 27.5618791],
Belgium	[50.8465573, 4.351697],
Belize	[17.250199, -88.770018],
Benin	[6.4990718, 2.6253361],
Bermuda	[+32.2930, -64.7820],
Bhutan	[27.5583211, 89.5524789],

Bolivia	[-16.4956371, -68.1336346],
Bosnia and Herzegovina	[43.8519774, 18.3866868],
Botswana	[-24.655319, 25.908728],
Brazil	[-15.7801, -47.9292],
British Virgin Islands	[18.4257128, -64.6232227],
Brunei	[4.8895453, 114.9417574],
Bulgaria	[42.6978634, 23.3221789],
Burkina Faso	[12.368148, -1.527085],
Burundi	[-3.3638125, 29.3675028],
Cambodia	[11.568271, 104.9224426],
Cameroon	[3.8689867, 11.5213344],
Canada	[45.421106, -75.690308],
Cape Verde	[14.9160169, -23.5096132],
Caribbean Netherlands	[12.183333, -68.25],
Cayman Islands	[19.2953549, -81.3807776],
Central African Republic	[4.3907153, 18.5509126],
Chad	[12.1191543, 15.0502758],
Chile	[-33.4377968, -70.6504451],
China	[39.9059631, 116.391248],
Colombia	[4.5980772, -74.0761028],
Comoros	[-11.6931255, 43.2543044],
Congo	[-4.2694407, 15.2712256],
Cook Islands	[-21.2074736, -159.7708145],
Costa Rica	[+9.9402, -84.1002],
Croatia	[45.813177, 15.977048],
Cuba	[23.138028, -82.3658671],
Curaçao	[12.1091242, -68.9316546],
Cyprus	[35.1739302, 33.364726],
Czech Republic	[50.0874654, 14.4212535],
Democratic Republic of Congo	[-4.3217055, 15.3125974],

Denmark	[55.6867243, 12.5700724],
Djibouti	[11.85677545, 42.7577845199437],
Dominica	[15.2991923, -61.3872868],
Dominican Republic	[18.4801972, -69.942111],
East Timor	[-8.5536809, 125.5784093],
Ecuador	[-0.2201641, -78.5123274],
Egypt	[30.048819, 31.243666],
El Salvador	[+13.7034, -89.2073],
Equatorial Guinea	[3.752828, 8.780061],
Estonia	[59.4372155, 24.7453688],
Ethiopia	[9.0, 38.75],
Falkland Islands	[-51.6950575, -57.8491693],
Faroe Islands	[62.012, -6.768],
Fiji	[-18.1415884, 178.4421662],
Finland	[60.1674086, 24.9425683],
France	[48.8566101, 2.3514992],
French Guiana	[4.9371143, -52.3258307],
French Polynesia	[-17.5384014, -149.5661868],
Gabon	[0.390002, 9.454001],
Gambia	[13.45535, -16.575646],
Georgia	[41.6935247, 44.8015019],
Germany	[52.5170365, 13.3888599],
Ghana	[5.5600141, -0.2057437],
Gibraltar	[36.1451465, -5.3484848],
Greece	[37.9841493, 23.7279843],
Greenland	[64.175029, -51.7355386],
Grenada	[12.0535331, -61.751805],
Guadeloupe	[16.0000778, -61.7333373],
Guam	[13.4727446, 144.752018225443],
Guatemala	[14.6417889, -90.5132239],

Guinea	[9.5170602, -13.6998434],
Guyana	[+6.8046, -58.1548],
Haiti	[18.547327, -72.3395928],
Honduras	[14.0931919, -87.2012631],
Hong Kong	[22.2793278, 114.1628131],
Hungary	[47.4983815, 19.0404707],
Iceland	[64.145981, -21.9422367],
India	[28.6141793, 77.2022662],
Indonesia	[-6.1753942, 106.827183],
Iran, Islamic Republic of	[35.7006177, 51.401375],
Iraq	[33.3024309, 44.3787992],
Ireland	[53.3497645, -6.2602732],
Israel	[31.78911765, 35.2229730135452],
Italy	[41.8933203, 12.4829321],
Ivory Coast	[+6.8067, -5.2728],
Jamaica	[17.9712148, -76.7928128],
Japan	[+35.6785, +139.6823],
Jordan	[31.9515694, 35.9239625],
Kazakhstan	[51.15092055, 71.4388595154859],
Kenya	[-1.2832533, 36.8172449],
Kiribati	[1.4845542, 172.968964820009],
Korea, Democratic People's Republic of	[39.019444, 125.738056],
Kosovo	[42.6638771, 21.1640849],
Kuwait	[29.3797091, 47.9735629],
Kyrgyzstan	[42.8767446, 74.6069949],
Laos	[17.9640988, 102.6133707],
Latvia	[56.9716562, 24.1665986020787],
Lebanon	[33.8959203, 35.47843],
Lesotho	[-29.310054, 27.478222],

Liberia	[+6.3106, -10.8047],
Libya	[32.896672, 13.1777923],
Liechtenstein	[47.1392862, 9.5227962],
Lithuania	[54.6870458, 25.2829111],
Luxembourg	[+49.6100, +6.1296],
Macao	[22.166667, 113.55],
Macedonia	[41.9960924, 21.4316495],
Madagascar	[-18.9100122, 47.5255809],
Malawi	[-13.973456, 33.7878122],
Malaysia	[3.1546872, 101.7136362],
Maldives	[4.1779879, 73.5107387],
Mali	[12.649319, -8.000337],
Malta	[35.8989818, 14.5136759],
Marshall Islands	[7.0909924, 171.3816354],
Martinique	[14.6027962, -61.0676724],
Mauritania	[18.0792379, -15.9780071],
Mauritius	[-20.1637281, 57.5045331],
Mayotte	[-12.7805856, 45.2279908],
Mexico	[19.4326009, -99.1333416],
Micronesia, Federated States of	[6.920744, 158.1627143],
Moldova	[47.0122737, 28.8605936],
Monaco	[43.7311424, 7.4197576],
Mongolia	[47.949809, 106.966724193881],
Montenegro	[42.4415238, 19.2621081],
Montserrat	[16.7053047, -62.2130447],
Morocco	[34.022405, -6.834543],
Mozambique	[-25.9686, +32.5804],
Myanmar	[19.7540045, 96.1344976],
Namibia	[-22.5530386, 17.0543787928491],
Nepal	[27.708796, 85.320244],

Netherlands	[52.3745403, 4.89797550561798],
New Caledonia	[-22.2745264, 166.442419],
New Zealand	[-41.2887639, 174.7772239],
Nicaragua	[12.1461244, -86.273717],
Niger	[13.4998139, 2.10948597212168],
Nigeria	[9.0643305, 7.4892974],
Northern Mariana Islands	[15.1909828, 145.746808767694],
Norway	[59.9132694, 10.7391112],
Oman	[23.5125498, 58.55137343736],
Pakistan	[33.666667, 73.166667],
Palau	[7.5006193, 134.6243012],
Palestine	[31.9030821, 35.1951741],
Panama	[8.9714493, -79.5341802],
Papua New Guinea	[-9.4743301, 147.1599504],
Paraguay	[-25.2959916, -57.6311196],
Peru	[-12.0621065, -77.0365256],
Philippines	[14.5906216, 120.9799696],
Pitcairn	[-25.066667, -130.1002054],
Poland	[52.2319237, 21.0067265],
Portugal	[38.7077926, -9.1365061],
Puerto Rico	[18.4500, -66.0667],
Qatar	[25.3014957, 51.4996673988223],
Réunion	[-20.8799889, 55.448137],
Romania	[44.4361414, 26.1027202],
Russia	[55.7507178, 37.6176606],
Rwanda	[-1.9441, +30.0619],
Saint Kitts and Nevis	[17.2960919, -62.722301],
Saint Lucia	[14.0095966, -60.9902359],
Saint Martin (French part)	[18.0668544, -63.0848869],
Saint Pierre and Miquelon	[46.7774872, -56.1767706],

Saint Vincent and the Grenadines	[13.1561864, -61.2279621],
Samoa	[-13.8343691, -171.7692793],
San Marino	[43.9458623, 12.458306],
Sao Tome and Principe	[0.3389242, 6.7313031],
Saudi Arabia	[24.6319692, 46.7150648],
Senegal	[14.693425, -17.447938],
Serbia	[44.8178131, 20.4568974],
Seychelles	[-4.6232085, 55.452359],
Sierra Leone	[8.479004, -13.26795],
Singapore	[1.2904753, 103.8520359],
Sint Maarten (Dutch part)	[18.0250713, -63.0483073],
Slovakia	[48.1359085, 17.1597440625],
Slovenia	[46.049865, 14.5068921],
Solomon Islands	[-9.4312971, 159.9552773],
Somalia	[2.042778, 45.338564],
South Africa	[-25.7459374, 28.1879444],
South Korea	[37.5666791, 126.9782914],
South Sudan	[4.8472017, 31.5951655],
Spain	[40.4167047, -3.7035825],
Sri Lanka	[6.9023184, 79.9240749],
Sudan	[15.593325, 32.53565],
Suriname	[5.8216198, -55.1771974],
Svalbard and Jan Mayen	[78.6352, 21.9939],
Swaziland	[-26.325745, 31.144663],
Sweden	[59.3251172, 18.0710935],
Switzerland	[46.9482713, 7.4514512],
Syrian Arab Republic	[33.5130695, 36.3095814],
Taiwan	[25.0375167, 121.5637],
Tajikistan	[38.5425835, 68.8152131865304],
Tanzania, United Republic of	[-6.1791181, 35.7468174],

Thailand	[13.7542529, 100.493087],
The Bahamas	[25.0783456, -77.3383331],
Togo	[6.130419, 1.215829],
Tonga	[-21.1343401, -175.2018085],
Trinidad and Tobago	[10.6572678, -61.5180173],
Tunisia	[+36.8117, +10.1761],
Turkey	[39.9215219, 32.8537929],
Turkmenistan	[37.9396678, 58.3874263],
Turks and Caicos Islands	[21.4607723, -71.1399956],
Uganda	[0.3177137, 32.5813539],
Ukraine	[50.4501071, 30.5240501],
United Arab Emirates	[+24.4764, +54.3705],
United Kingdom	[51.5073219, -0.1276474],
United States	[38.8949549, -77.0366456],
Uruguay	[-34.9059039, -56.1913569],
Uzbekistan	[41.3123363, 69.2787079],
Vanuatu	[-17.7414972, 168.3150163],
Venezuela	[10.506098, -66.9146017],
Vietnam	[21.0292095, 105.85247],
Virgin Islands, U.S.	[18.341137, -64.932789],
Yemen	[15.342101, 44.2005197],
Zambia	[-15.416697, 28.281381],
Zimbabwe	[-17.831773, 31.045686],

Appendix B. Factors that Influence the Success of Crowdfunding Campaigns

Funding Target	Mollick (2014) finds that also the choice of the target amount can influence the success probability. Backers evaluate whether a funding goal is realistic. Project goals that are perceived too high or too low are more likely to fail. Accordingly, Krishnan et al. (2015) show in their research that crowdfunding campaigns that have a fixed funding goal, the “all or nothing” approach, tend to be more successful than projects with variable funding targets. It appears that backers have considerably more confidence in the “all or nothing” approach as the money will only be forwarded to the creators if the pre-defined target amount is reached. Kickstarter only allows the “all or nothing” approach, while creators on the Indiegogo crowdfunding platform can chose between the “all or nothing” and the “flexible” funding target.
Funding Duration	Frydrych et al. (2014), for example, find in their research on Kickstarter that the choice of the duration of the crowdfunding campaign can fluence its performance. Surprisingly, their findings suggest that entrepreneurs should implement shorter funding periods, ranging from 20-30 days. The scientists provide different explanations for this recommendation: First, shorter funding periods communicate a tone of confidence and urgency for backers to engage. In contrast, long durations may appear less urgent and encourage procrastination. Second, the authors argue that the attention towards a specific project seems to diminish with time. Entrepreneurs find it often difficult to keep the momentum of their campaigns alive. The longer the fundraising period, the more difficult it is to motivate backers to participate (Mollick, 2014). In accordance with the findings from the literature, platform operators equally recommend a rather short funding cycle. In general, the recommendation is to create a project that does not last longer than 30 days (Robertson & Wooster, 2015). This duration has become a standard recommendation for crowdfunding projects hosted on Kickstarter and has been adopted by other crowdfunding platforms.

Update Frequency	Xu et al. (2014) find in their research that frequent status updates are crucial for the performance of crowdfunding campaigns. Project owners that frequently update the community on their progress are able to significantly increase the probability of successful fundraising (32.6% vs. 58.7%). Moreover, the researchers show that intensive communication between the project creators and the community is more predictive of success than the project description (Xu et al., 2014). This finding is in accordance with Kuppuswamy and Bayus (2014), who discover that recent updates, especially in the final stage of the crowdfunding project, have a positive influence on the achievement of the funding target. The authors suggest that entrepreneurs should try to awaken emotions and excitement, especially in the final stage of the fundraising to increase the overall funding.

Appendix C. Model Performance Comparison for Inclusion of LP-Variable

Table 1: Results of Negative Binomial Regression Including LP-variable

*This table shows the results of the Negative Binomial Regression which describes the effect of different variables (Column 1) on the count of backers (dependent variable) that supported a specific crowdfunding project from a specific country. Columns 2-5 show the coefficients, IRRs and 95% confidence intervals for the different variables. The IRRs are obtained from the coefficients through exponentiation (cp. Eq. 10). They can be interpreted as the multiplicative effect for a one-unit increase of the respective variable. In this context, all IRRs < 1 indicate a negative effect relationship between dependent and independent variables. Accordingly, IRRs > 1 reveal a positive relationship. The symbols *, **, and *** denote statistical significance of the respective variables at a p-value of 0.01, 0.001, and <0.0001, respectively (R software default settings). The fourth and fifth column present the lower (2.5%) and upper limit (97.5%) of the 95% confidence interval for the IRRs.*

Dependent Variable = Count of Backers				
Independent Variable	**Coef.**	**IRR**	**Conf. Limit 2.5%**	**Conf. Limit 97.5%**
Distance	-0.04	0.96***	0.96	0.96
GDP-B	0.03	1.03***	1.03	1.03
GDP-E	-0.01	0.99***	0.99	0.99
PWL	0.64	1.89***	1.88	1.91
Large Project	2.43	11.41***	11.32	11.47
Covid-19 Pandemic	0.21	1.24***	1.21	1.27
Category = Art	-0.87	0.42***	0.41	0.42
Category = Comics	-0.94	0.39***	0.38	0.39
Category = Crafts	-0.97	0.38***	0.37	0.39
Category = Dance	-1.24	0.29***	0.28	0.30
Category = Design	-0.01	0.99	0.98	1.00
Category = Fashion	-0.65	0.52***	0.51	0.53
Category = Film & Video	-0.83	0.44***	0.43	0.44
Category = Food	-1.11	0.33***	0.32	0.34
Category = Journalism	-0.89	0.41***	0.39	0.42
Category = Music	-1.17	0.31***	0.31	0.32
Category = Photography	-0.95	0.38***	0.38	0.39
Category = Publishing	-1.00	0.37***	0.36	0.37
Category = Technology	-0.14	0.87***	0.86	0.88
Category = Theatre	-1.22	0.29***	0.29	0.30

Table 2: Results of Negative Binomial Regression excluding LP-variable

*This table shows the results of the Negative Binomial Regression which describes the effect of different variables (Column 1) on the count of backers (dependent variable) that supported a specific crowdfunding project from a specific country. Columns 2-5 show the coefficients, IRRs and 95% confidence intervals for the different variables. The IRRs are obtained from the coefficients through exponentiation (cp. Eq. 10). They can be interpreted as the multiplicative effect for a one-unit increase of the respective variable. In this context, all IRRs < 1 indicate a negative effect relationship between dependent and independent variables. Accordingly, IRRs > 1 reveal a positive relationship. The symbols *, **, and *** denote statistical significance of the respective variables at a p-value of 0.01, 0.001, and <0.0001, respectively (R software default settings). The fourth and fifth column present the lower (2.5%) and upper limit (97.5%) of the 95% confidence interval for the IRRs.*

Dependent Variable = Count of Backers				
Independent Variable	**Coef.**	**IRR**	**Conf. Limit 2.5%**	**Conf. Limit 97.5%**
Distance	-0.06	0.94***	0.94	0.94
GDP-B	0.04	1.04***	1.04	1.04
GDP-E	-0.01	0.99***	0.99	0.99
PWL	1.16	3.19***	3.17	3.21
Covid-19 Pandemic	0.26	1.30***	1.27	1.34
Category = Art	-1.80	0.17***	0.16	0.17
Category = Comics	-1.46	0.23***	0.23	0.24
Category = Crafts	-1.88	0.15***	0.15	0.16
Category = Dance	-2.60	0.07***	0.07	0.08
Category = Design	0.04	1.04***	1.03	1.05
Category = Fashion	-1.14	0.32***	0.31	0.32
Category = Film & Video	-1.34	0.26***	0.26	0.26
Category = Food	-1.53	0.22***	0.21	0.22
Category = Journalism	-1.39	0.25***	0.24	0.26
Category = Music	-1.95	0.14***	0.14	0.14
Category = Photography	-1.77	0.17***	0.17	0.17
Category = Publishing	-1.62	0.20***	0.19	0.20
Category = Technology	-0.26	0.77***	0.76	0.78
Category = Theatre	-2.24	0.11***	0.10	0.11