TUHH
Hamburg University of Technology

The flipside of the coin - an investor-level investigation of blockchain-based fundraising (Initial Coin Offerings)

Vom Promotionsausschuss der
Technischen Universität Hamburg

zur Erlangung des akademischen Grades

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

genehmigte Dissertation

von

Carolin Petra Brückmann

aus

Kassel

2021

Gutachter

Prof. Dr. Christoph Ihl

Prof. Dr. Jörn Block

Prof. Dr. Christian Lüthje

Tag der mündlichen Prüfung: 12. Oktober 2021

https://orcid.org/0000-0003-2962-1284

Abstract

The flipside of the coin - an investor-level investigation of blockchain-based fundraising (Initial Coin Offerings)

by Carolin Petra Brückmann

Initial Coin Offerings (ICOs) are an early-venture financing innovation in which start-ups raise capital from investors on the blockchain. ICOs may democratize financing, thereby diversifying the start-up space, consequently benefitting society as a whole. Whether ICOs are able to add value sustainably depends on the crowd decision-making effectiveness. This thesis investigates whether investor crowds channel capital to prospering ventures. I approach this novel topic broadly and gradually zoom-in across *three* research studies, based on *26* million blockchain transactions by *1,628* ICOs between 2015-18. The first study explores ICO crowd composition. The findings point to pronounced crowd heterogeneity. While novice investors dominate the crowd, campaign success depends on a small number of high-contribution investors. The second study analyzes crowd dynamics and shows that investors revert to imitating others in the face of uncertainty, whom they assume to possess superior information. The third study analyzes whether crowds successfully *scout* prospering ventures from the pool of capital seeking ventures. The findings suggest that the crowd does not channel capital to higher performing ventures. Yet, we find weak evidence for superior decision-making by experienced crowds. Taken together, the findings point to two sides of the same coin. ICOs have not yet been adding value to society. On the flipside, our findings show that decision-making improves with investor experience. Thus, a maturing market aided by effective regulation may enable ICOs to create a sustainable positive impact in the long-run. This thesis contributes to theory by highlighting the importance of disentangling the crowd in ICOs. We also extend the understanding of the circumstances under which herding may be(come) rational and contribute an ICO perspective on crowd decision-making quality.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

In 2018, when I started working on this thesis, the market for Initial Coin Offerings (ICOs) was evolving rapidly. The ICO market value reached an astonishing *17* billion US Dollar in 2018, with an average ICO campaign raising *12* million US Dollar (Fromberger and Haffke, 2020). Our analysis shows that investors, on average, earned almost *60%* return on their investment in the short-run. If these statistics do not sound impressive enough on their own, placing them in context may highlight their peculiarity. An ICO raised *440x* more than an average crowdfunding campaign (Fundera.com, 2020), while the underlying fundraising concept and ventures are expected to be comparable. Moreover, the ICO market value tripled within a single year from 2017 to 2018 (Fromberger and Haffke, 2020).
I wanted to understand what distinguishes ICOs; consequently making them so attractive to investors. Moreover, when I started my research, market developments revealed the seemingly shiny coin's flipside. While the average ICO gained *60%* in value, the majority (median ICO) lost *90%* of its value on crypto-exchanges. Scandals of fraudulent ventures attempting to take advantage of the investment *Wild West* - opaque, unregulated, and inherently risky - came to light.

Indeed, crowdfunding in general and ICOs in particular place investors in a novel environment, one historically open primarily to professional investors such as venture capitalists. Early-venture investments are inherently risky, with around *90%* of start-ups failing within *10* years and less than *50%* making it to their *5*th year (Prosser, 2016). Investors face the challenge of predicting which ventures will turn into successful businesses. Most ICOing ventures do not even have a test version of their product available and usually no track record that would allow investors to predict the likelihood of success based on traditional measures. Entrepreneurs have the incentive to window-dress and not reveal any negative information to investors; information asymmetry between ventures and investors is the consequence. The venture capital industry has formed to reduce information asymmetry by bundling investor capital and investing significant amounts of capital in performing due diligence on venture prospects (Isenberg, 2012).

In contrast, ICO crowds dominated by many low-contribution investors must rely on the (marketing) information provided by ventures. Their low individual investment amounts cannot justify significant due diligence expenses. Furthermore, arguably, most retail investors lack the capabilities to perform thorough investment analyses. Plausibly, ICOs are subject to even higher levels of uncertainty than the average venture. Their tokens' value critically depends on the value of blockchain technology, which in itself is volatile and highly uncertain, further compounding the challenge to investor crowds. This thesis has been devoted to understanding how (retail) investors form investment decisions in this novel environment.

1.2 Research question and background

A growing body of research has started to explore Initial Coin Offerings, usually taking either a regulatory or venture lens. The *regulatory lens* explores how regulators can balance nurturing innovation and protecting investors. Early research investigated how voluntary disclosure affects ICO fundraising success (Parrino and Greenslade, 2014; Giudici and Adhami, 2019; Bourveau et al., 2018; Leone et al., 2007; Blaseg, 2018, e.g.,). Their findings show that ICOs benefit from voluntary disclosure by raising more capital and reducing ICO underpricing, detailed in the following section. Moreover, clear governance structures serve as a credible quality were found to signal and increase the probability of fundraising success. However, findings also indicate that many ICOs do not enclose sufficient information to form proper investment decisions, e.g., risk factors or use of funds.

Another stream of research discusses appropriate regulatory frameworks (Kher et al., 2020). Dell'Erba (2018) contrasts *do-no-harm* and *full-reinforcement* regulatory approaches and their consequences. In another example, Sherman (2018) advocates a behavioral economics regulatory framework, building on human decision-making imperfections due to bounded rationality to protect uninformed investors. Other researchers seek to create transparency of the underlying phenomenon and present a taxonomy of regulatory challenges that have to be tackled by regulators (e.g., Zetzsche et al., 2017; Mendelson, 2019; Wöckner et al., 2018).

A large stream of research examines how the regulatory environment affects the developments of the ICO market. Zhang et al. (2020) shows that regulatory uncertainty may be more harmful than imperfect regulation by lowering investor demand for ICOs. Other research shows how ICOing ventures converge towards geographical areas with clearly defined regulation (e.g., Malta) or favorable ICO jurisdictions (often tax havens). Indeed, such migration patterns point to a major challenge in regulating ICOs: as a global phenomenon, they are not bound by national borders. Thus, their regulation must be a truly global effort (e.g., Boreiko and Risteski, 2020).

Another strand of research adopts the *venture lens* on ICOs, focusing on how ventures can successfully raise capital on the blockchain (e.g., Fisch, 2019; Amsden and Schweizer, 2018; Momtaz, 2018b; Mollick, 2013; Cerchiello, 2018; Rhue, 2018). One larger stream of research investigates how ventures can credibly send quality signals to investors to reduce information asymmetry. Findings suggest that investors evaluate business and campaign-related quality signals, e.g., a higher share of retained equity and team background (Amsden and Schweizer, 2018). Investors positively interpret signals that increase transparency, such as publicly accessible code on *GitHub*[1].

A second research stream explores ICO underpricing (Felix, 2018; Kher et al., 2020; Bourveau et al., 2018; Momtaz, 2018a; Benedetti and Kostovetsky, 2018). The larger the underpricing, the more money that was *left on the table* by ventures during the campaign. Research indicates large underpricing in ICOs with a median underpricing of around *24%* (Lee et al., 2019). While the reasons for the large underpricing are not fully understood yet, several hypotheses exist. For example, underpricing may be intentional, serving as a risk discount for high information asymmetry. Alternatively, the large magnitude could be unintended and explained by demand uncertainty by ventures (Lee et al., 2019).

The third stream of research compares ICOs to alternative sources of financing, aiming to investigate what ventures can gain and what they risk by pursuing an ICO (e.g., Liu and Wang, 2018; Catalini and Gans, 2018; Hui et al., 2014; Anson, 2018; Momtaz, 2018b; Benedetti and Kostovetsky, 2018; Chod and Lyandres, 2018).

In our understanding, a third lens has been underrepresented in ICO research thus far, namely the *investor lens*. Research has started to explore investor decision-making during ICO campaigns (Amsden and Schweizer, 2018; Cerchiello, 2018; Adhami et al., 2018; Benedetti and Kostovetsky, 2018; Cordova et al., 2015; Felix, 2018; Momtaz, 2018b; Fisch, 2019). Their findings show that investors base

[1]GitHub is an open-source development platform that programmers use to showcase their work and co-develop along with other programmers (Github, 2020).

decisions on campaign attributes. However, they show that investors also scan the environment for cues, often social, to aid their decision-making. Momtaz (2018b) for instance, finds that investors place a higher decision-weight on alleged experts' ratings than on venture attributes. Similarly, Benedetti and Kostovetsky (2018) find that venture *Twitter* activity predicts fundraising success and Lee et al. (2019) shows that investors mimic other investors' behavior, a phenomenon called *herding*. Therefore, previous findings point to the importance of understanding investors individually as well as the dynamics among them. Research on *two* related areas, crowdfunding and stock markets, further supports the understanding that ICO investors' behavior could be of integral importance.

Crowdfunding research acknowledges the importance of disentangling investors by showing that investment crowds exhibit pronounced heterogeneity and that subgroups vary greatly in their decision-making. For example, Wallmeroth (2019) studies investor behavior in German crowdfunding campaigns. He finds evidence for variations in investment motives and strategies of investors. Abrams (2017) shows that early and late investors in campaigns form their investment decisions based on substantially different decision cues. Wick and Ihl (2019) build on their finding and show that herding behavior differs between investor groups.

Stock markets and the theory of speculative bubbles demonstrate how dynamics amongst investors can have disastrous effects on individual investors and the market as a whole. Binglin et al. (2017) show how historical stock market bubbles were systematically characterized by a high share and constant inflow of inexperienced investors into the market. As such, their findings highlight the importance of understanding crowd composition. Greenwood and Nagel (2009) show that inexperienced investors are more susceptible to trend-chasing behavior than experienced investors. Griffin et al. (2011) extend their finding by showing how professional investors sold their assets in a coordinated sales effort to uninformed investors at the peak of the dot-com bubble, effectively exploiting them.

I have devoted this dissertation to understanding ICOs from the investors' perspective, aiming to close a gap in the current academic understanding of ICOs. My guiding research question is as follows:

Who invests in ICOs and how do market participants form decisions in this novel environment, characterized by buzz and high uncertainty?

I aim to answer this research question based on *three* sub-questions analyzed across *three* research studies. The first study aims to profile ICO investors and understand both the individual investor and crowd composition. We perform cluster analysis to investigate the degree of crowd heterogeneity and identify behavioral investor clusters to develop a foundation of ICO investor research.

The second research study asks how ICO investors form investment decisions. We investigate to what extent herding is present in the ICO market and whether market participants herd irrationally or strategically reinterpret cues in the light of new information. The third research piece tackles the effectiveness of crowd-based decision-making in ICOs. We ask whether the *wisdom of the crowd* channels funding to the most promising start-ups?

1.3 Results and contributions

Foremost, we see a significant contribution of this dissertation in developing the first thorough empirical understanding of ICOs. Specifically, we contribute perspectives on the decision-making of ICO investors and the post-campaign success of ICO-backed ventures. While this may be interpreted as an empirical rather than a theoretical contribution, we hope that our work will serve as a groundwork for future research on ICOs and inspire fellow researchers and practitioners to fruitful discourse on the future of ICOs.

Our theoretical contributions are threefold. First, we contribute to the understanding of crowdfunding by highlighting the importance of disentangling the crowd. We show that high heterogeneity exists among ICO investors and demonstrate how decision-making differs between investors. We show that investigating decision formation for the crowd as a homogeneous group does not live up to the complexity of the phenomenon. While crowdfunding research has started to explore the heterogeneity of investors (e.g., Abrams, 2017; Wick and Ihl, 2019; Wallmeroth, 2019; Moritz and Block, 2016; Lin et al., 2014; Hervé, 2016; Kim and Viswanathan, 2018; Hervé, 2016; Li et al., 2016), we are the first to investigate heterogeneity in the context of ICOs, as called upon by numerous researchers (e.g., Fisch, 2019).

Second, we provide evidence for the existence of herding in ICOs. Our findings support initial evidence on crypto-exchange trading (Kallinterakis and Wang, 2019; Gurdgiev and O'Loughlin, 2020) and ICO campaigns (e.g., Lee et al., 2019; Boreiko and Risteski, 2020). Beyond that, we extend the understanding of how environmental factors affect herding. We show that irrational herding is reinforced by the ability to observe other investors and being fueled by social media sentiment. We also show how investor experience moderates the vulnerability of investors to socially induced herding behavior. This finding contributes to the herding theory by testing it in novel contexts (ICOs) and extending the understanding of moderating effects (experiential learning).

Last, we contribute an ICO perspective on crowd decision-making effectiveness. We show that ICO crowds generally do not channel capital to successful start-ups, thus pointing to the value of *scouting* by VCs, extending the work of Carnahan et al. (2010), Baum and Silverman (2004) and, Megginson and Weiss (1991). Beyond our theoretical contributions, we hope to invoke thought processes among regulators as well. The dominance and herding tendency of uninformed investors in the ICO market indicate the need for a clear and effective regulatory framework to protect investors' wealth.

1.4 Structure

The remainder of the thesis is structured around our *three* research papers. Before deep-diving into them, we briefly elaborate on the ICO phenomenon and the data used throughout this work in order to prevent repetitive elaborations and prepare the reader appropriately.

As ICO investors' topic is unexplored to date, we initiate our research with broadly and gradually adopt a more focused approach. First, we explore the composition of ICO crowds in *Chapter 4*: *Putting a face to the crowd - an investor-based exploration of Initial Coin Offerings*. We perform cluster analysis to explore crowd heterogeneity and characterize ICO investors. Subsequently, we show how investor groups differ in their decision-making by analyzing how their decision to invest is influenced by market sentiment. Second, we investigate crowd dynamics in *Chapter 5*: *Sheep or Shepherd? Rational herding in ICOs and the role of experience*. We investigate herding in ICOs and show how herding is moderated by investor experience. Third, we investigate crowd decision effectiveness in *Chapter 6*: *Wisdom of the crowds in ICOs? The decision-quality of ICO crowds*. We investigate ICO post-campaign success and investigate whether and which subgroups of the crowd channel capital to prospering ventures.

We place our findings into a broader context of theory and discuss their theoretical and practical implications in each research piece. Of course, we also elaborate on the limitations of our research and point to future research directions. I conclude this dissertation by synthesizing all *three* studies' findings, placing them into the context of recent developments in the ICO market, and pointing to future research directions.

Chapter 2

Context

2.1 Blockchain technology and the rise of ICOs

The decoupling of blockchain technology from its original peer-to-peer payment use case, *Bitcoin*, has enabled the innovation of Initial Coin Offerings (ICOs). The collaborative development of *Bitcoin* and blockchain (as its enabling technology) are based on a whitepaper published in 2008 by an anonymous syndicate who called themselves *Satoshi Nakamoto*. They proposed a decentralized peer-to-peer network approach to verify electronic payment transactions and consequently remove the need for trust-building financial intermediaries such as banks. Therefore, blockchain technology is designed to replace trust with cryptographic proof, verified by a network of independent nodes. It has been argued that the verification algorithm is resistant to manipulation as long as less than *50%* of the computational power is controlled by a single or collusive entity (Nakamoto, 2008). Since then, it has been shown that it is possible to manipulate the network with as little as *25%* of the computing power (Griffin and Shams, 2018). Moreover, the cryptographic proof algorithm represents a major challenge in developing a scalable blockchain solution. Specifically, the original cryptographic approach's computational requirements, *proof-of-work*, are significant, causing low transaction speed and extremely high energy consumption (Lee, 2016). As a result, the economic value of blockchain technology has yet to be determined. Its value will ultimately depend on the value of its use cases, one of which is Initial Coin Offerings (hereafter ICOs).

In ICOs, early-ventures employ blockchain technology to raise capital from many individual investors directly, without the need for intermediaries. Hence, ICOs democratize early-venture financing by redistributing decision power from a few (financial intermediaries such as venture capitalists or banks) to many (potentially everyone). ICOs may be regarded as a sub-form of conventional crowdfunding, a term used to describe the process of ventures raising capital from retail investors on crowdfunding platforms, which has gained in popularity since around 2009. I will further elaborate on crowdfunding in comparison to ICOs later in this chapter.

The first ICO was conducted by *MasterCoin*, now *Omni*, by *J.R. Willet* in 2013 (Merre, 2020). During a *bitcoinforum*, *Willet* presented the idea of *MasterCoin* as a project aiming to extend blockchain functionality and asking for the *Bitcoin*-donations to realize his idea. Donations would be rewarded with units of a new cryptocurrency, which contributors could use to pay for services on the *MasterCoin* platform, once functional (Merre, 2020). *MasterCoin* raised *600,000* US Dollar and *Willet* had conducted the first ICO. During the conference, he also shared his thoughts on the potential of blockchain technology as a means to raise capital for projects. Over the course of 2013 and 2014, *four* projects raised capital on the blockchain: *NextCoin*, *CounterParty*, *MaidSafeCoin* and *Swarm* (Merre, 2020).

The ICO-boom initiated with the record ICO of *Ethereum* in 2014. *Vitalik Buterin* raised close to *14* million US Dollar to build a new blockchain infrastructure, which would allow users to create customized add-on protocols to *Ethereum*, so-called *smart contracts*. *Smart contracts* are run on conventional blockchain infrastructure and are automatically executed on the pre-programmed terms and conditions (Merre, 2020). *Ethereum* quickly became the medium of choice for many ICOs, as *smart contract* made it surprisingly simple to custom-design ICO contracts. With its record ICO, *Ethereum* also contributed to raising the visibility of ICOs to the general public. *Ethereum* investors doubled their investment within *24* hours and earned *4,520x* their investment within *three* years, which arguable attracted many novice participants to the market.

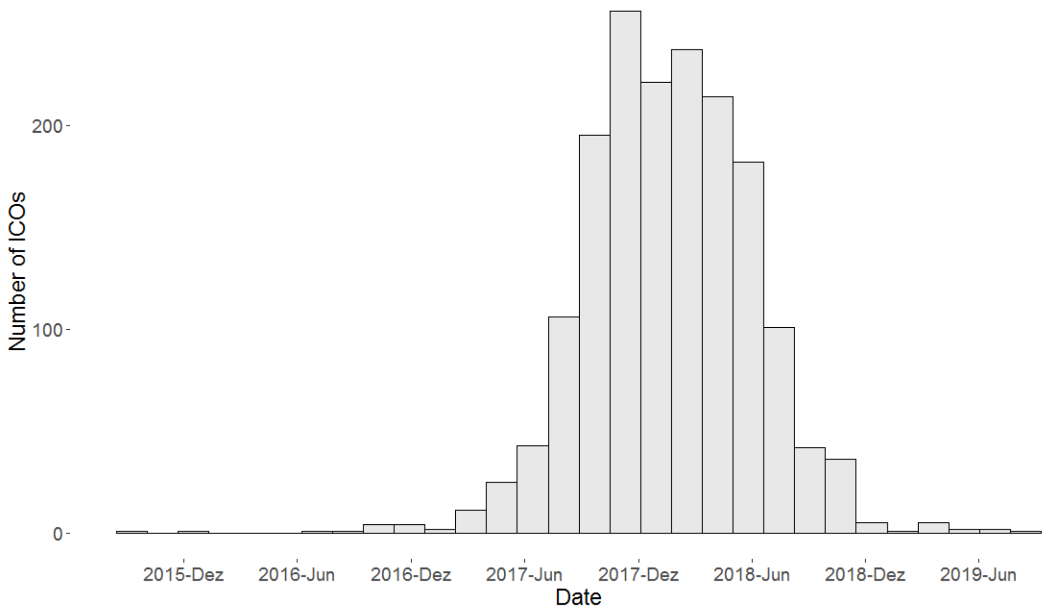
ICO market activity gradually increased over the next years, peaking in 2018. *Figure 2.1* shows the distribution of ICOs in our sample and demonstrates the sharp rise in market activity in 2017. An analysis by Fromberger and Haffke (2020) reports a total of *3,000* ICOs in 2017-18. According to their research, more than *14* billion US Dollar were raised by ICOs in 2018 (Fromberger and Haffke, 2020) and the average campaign raised *13* million US Dollar, approximately *440x* more than a conventional crowdfunding campaign in 2020 (Fundera.com, 2020). Mega ICOs during this period raised much more than the *Ethereum*. The largest ICO in history was conducted by *EOS*, a development platform for decentralized Apps (so-called *dApps*), which raised more than *4* billion US Dollar (Penke, 2018).

The ICO-boom combined with the nature of the market, i.e., risky ventures and the absence of regulation, also attracted scammers. Indeed, our data show that the median ICO has lost *90%* of value on crypto-exchanges. The rising investor and regulatory scrutiny led to a sharp decline in market activity at the end of 2018, as observable in *figure 2.1*. Today, few ventures attempt to raise capital on the blockchain (Suberg, 2018).

2.2 The ICO-process

The process at the core of an ICO is relatively simple: a venture attempting to raise capital writes a *smart contract* in a selected blockchain environment. A *smart contract* specifies the terms and conditions for a new crypto token to be created. For example, the contract specifies the token design, e.g., how many tokens are issued

FIGURE 2.1: Distribution of ICO campaigns in the data sample



Note: Above figure shows the monthly distribution of ICO-campaigns in our data sample over the course of our observation period between September 2015 and year-end 2018. Campaigns are counted in the month of listing of the pre-ICO, if conducted, or main ICO-campaign otherwise.

and at what price. The contract is smart as the protocol automatically executes it upon triggering events. For example, all tokens are distributed to investors if, and only if a minimum amount is raised during fundraising. Otherwise, funds are returned to investors. As such, *smart contracts* enable anonymous transactions between unrelated parties without requiring intermediaries (e.g., banks) to provide legitimization. The development of programmable open-source blockchains, most notably *Ethereum*, has made it surprisingly simple to program *smart contracts*. *Ethereum* even offers a standard token contract *ERC-20*, a *smart contract* template (Rhue, 2018).

An ICO token is usually offered to the general public in a fundraising campaign, which is investable via company websites and often advertised on several ICO platforms. Ventures publish marketing material to promote their campaign. While no legally binding enclosure requirements exist, companies will often publish a so-called *whitepaper*, summarizing business plan and campaign design. The

whitepaper is commonly supplemented with videos and extensive social media communication. Ventures interact with potential investors via social media platforms (e.g., *Telegram*, *bitcointalk*).
ICO tokens are often listed on crypto-exchanges (e.g., *Binance*), allowing investors to trade tokens, making ICOs a liquid investment. Tokens can be purchased against major cryptocurrencies such as *Ether* or *Bitcoin*, sometimes also against fiat currency (legal tender issued by governments).
Marketing is commonly the main expense of an ICO. During the ICO-boom, ventures competed against many other ICOs. Gaining visibility and signaling quality became increasingly important. In response, many ventures hired ICO servicing companies to assist with optimal campaign design, produced high-quality marketing content, and hired investor relation teams to keep investors engaged.
The token in ICOs, similar to different forms of crowdfunding, may serve several purposes. Although issuers may freely assign rights to a token, most tokens belong in one of *three* categories. *Utility tokens* provide access rights to a product, service, or platform. *Cryptocurrencies* represent a storage of value used as a means of payment, independent from the underlying company. *Security tokens* offer some form of performance-based reward, comparable to dividends (Amsden and Schweizer, 2018). Most tokens, $\sim$*90%* in our sample, are *utility tokens*. The reason is simple: by offering a *utility token*, a venture operates outside the existing regulatory framework, as their tokens are not regarded as an investment. Rather, they pre-sell their product. This places ventures in a comfortable position, as Isenberg (2012) argues: ventures do not have to give up equity, yet still receive *risk-free* funding, as investors have no legal claim if the venture fails. If an ICO issued a *security token*, it would be required to register with the *Security and Exchange Commission (SEC)*, which increases both cost and time to market. Thus, as long as investors are willing to invest without acquiring equity, ICOing ventures seem to have little incentive to offer equity.

Regulating ICOs is difficult for *three* reasons. First, as just described, a token technically is not an investment and is not covered by any existing regulatory framework. Second, the regulation of ICOs requires global collaboration. Third, governments struggle between protecting investors and at the same nurturing innovation. Jurisdictions around the world have adopted varying approaches to regulating ICOs. Some countries, such as China, have prohibited (unregistered) ICOs altogether. Other countries, such as Malta, have proactively developed clear regulatory frameworks, reducing uncertainty and risk for ventures and investors alike (Wöckner et al., 2018).

2.3 Alternative early-venture financing methods

Before the crowdfunding era, entrepreneurs seeking capital to start-up their ventures had *three* choices: seek financing from family & friends, obtain a bank loan or seek capital from professional investors, primarily angel investors or venture capitalists (VCs). Funds from family & friends are often insufficient to start a business, and banks are resistant to lend money in the absence of collateral. As a result, venture capital has become a major source of start-up financing. According to the *OECD*, venture capital investments in 2019 reached above *60* billion US Dollar globally (OECE, 2020).

In contrast to venture capital, crowdfunding democratizes start-up financing, thereby potentially enabling a more diverse set of start-ups to raise capital. Research indicates that VC decision-making is biased, favoring a certain set of start-ups over others. Research has shown, for example, that VCs tend to over-rely on team performance (Schefczyk and Gerpott, 2001; Baum and Silverman, 2004), are subject to natural behavioral biases such homophily (Cumming and Dai, 2010; Franke et al., 2006; Malmström et al., 2020) and suffer from over-confidence (Zacharakis and Shepherd, 2001).

Crowdfunding may also help ventures gain publicity and test product demand. Indeed, crowdfunding could be regarded as customer voting on business ideas. Ventures may tap into the *wisdom of the crowd* to co-develop products. Customers, who are now also investors, naturally have a high(er) incentive to devote time to the start-up, as they personally benefit from it.

Crowdfunding is generally divided into equity-, reward-based- and donation-based crowdfunding. ICOs may be regarded as a subform of crowdfunding. While both share many characteristics, they differ in *four* important ways. First, fundraising in ICOs is by definition accompanied by the issuance of tokens. The token-contract is programmed as a self-executing contract, eliminating the dependence on intermediaries. In contrast, crowdfunding still requires the interference of crowdfunding platforms (e.g., *Kickstarter*), which serve *two* functions: they provide a technical platform to raise and redistribute funds, and they pre-screen campaigns based on platform-specific quality criteria. Consequently, they partially substitute the role to reduce information asymmetry from VCs. ICOs make fundraising platforms redundant. Interestingly, however, alternative intermediaries quickly emerged, effectively taking the place of providing quality assurance, i.e., rating platforms (Lee et al., 2019; Rhue, 2018).

Second, while start-up investments were historically illiquid, realizable only at exit often through an IPO, ICO tokens can be traded on crypto-exchanges, allowing investors to exit at any time. While this increases liquidity and reduces risk, it also makes ICO tokens target for short-term speculation. Investors' constant scrutiny may also reduce entrepreneurial willingness to take risks and instead maximize the valuation on crypto-exchanges. Such a myopic behavior is well-documented for large corporations trading on stock markets (Zhang and Gimeno, 2016; Bolton, 2006). Arguably, start-ups require a longer-termed perspective from investors, leaving room for experimentation and development.

Third, crowdfunding is regulated by the *JOBS Act*[1] enacted by the Obama Administration in 2005. The JOBS Act exempts crowdfunding ventures from registering with the *SEC* if they fulfill certain requirements (SEC, 2015). ICOs, in contrast, are currently the *Wild West* of financing. This may, in turn, create high regulatory uncertainty for ventures and investors alike.

Last, the ICO hype led to a rapid rise of ICO activity, inflating the market to *14* billion US Dollar in 2018 (Fromberger and Haffke, 2020). In comparison, the crowdfunding market reached *11* billion US Dollar in 2018 (pnswire, 2019). As such, the opportunity to raise capital is higher for ventures pursuing an ICO, which may, in turn, attract systematically different ventures and investors to the market.

[1] Jumpstart Our Business Startups.

Chapter 3

Data

3.1 Transaction dataset

The analyses presented throughout this work are based on a dataset of *37* million *Ethereum* blockchain transactions, which we have extracted from the blockchain explorer Ethplorer (*ethplorer.io*). While they are employed differently in each *chapter*, they represent the baseline for all analyses. Accordingly, the dataset construction and cleaning process deserves some upfront elaboration, which are summarized in *figure 3.1*. We will elaborate on each processing step in the following sections.

3.1.1 Identifying blockchain-linkable ICO campaigns

The initial step of our data preparation is to identify ICOs conducted during the period of our analysis. No comprehensive database exists, we thus consolidate ICOs listed on the *nine* most comprehensive ICO platforms: *ICOdata*, *ICORating*, *ICObench*, *ICOdrops*, *Coingecko*, *Cryptoslate*, *Tokendata*, *ICOwatchlist* and *Coincheckup*. We identify *3,209* ICOs between September 2015 and 2018 and subsequently match them with their blockchain accounts on *Ethplorer*. We are able to link *1,705* of *3,209*

FIGURE 3.1: Construction of blockchain transaction dataset

Steps not relevant for all Chapters (steps 5 and 6)

	①	②	③	④	⑤	⑥
Processing step	**Identified ICO-campaigns**	**Linkable to blockchain accounts**	**Issuer ID identifiable**	**Investment amount available**	**Meet cluster criteria *(Chapter 4 only)***	**Accurate transaction timestamps *(Chapter 5 & 6)***
Number of:						
Campaigns	3,209	1,705	1,705	1,628	1,628	1,274
Investors (mn)	na	nm	9.4	4.9	1.4	3.6
Transactions (mn)	na	37	37	26	19	17
Description	All ICOs listed on: ICO data, ICORating, ICObench, ICOdrops, Coingecko, Cryotoslate, Tokendata, ICOwatchlist, Coincheckup	ICOs linkable to their Ethereum blockchain accounts via ethplorer.io	Identify venture blockchain IDs in order to classify investments into primary (ICO) and secondary (crypto-exchange) phase	Impute transactions with missing transaction amounts in accordance with imputation strategy	Remove investors with > 3,500 investments (exchanges) and single investments (treated as pre-defined cluster)	Remove ICOs whose blockchain transactions are recorded with time delay

Note: Above figure visualizes the construction of our core dataset of Ethereum-based blockchain transactions the the cleaning steps performed as a basis for all research pieces presented in this work. nm means not meaningful, mn refers to millions.

Ethereum-based ICOs with their blockchain-accounts. The *50%* ratio seems rather low. However, matching is only possible based on venture names, which are often not unique, and different forms of writing may exist. Matching is performed based on *Levenshtein-distance*[1] between campaign name and blockchain contract name, aided by manual correction for distances below *90%*.

We extract all transactions for the remaining *1,705* ICOs recorded until year-end 2018, resulting into *37* million transactions. For each transaction, we obtain a timestamp, information on the sending and recipient ID, the transfer amount, and a transaction *hash* (a unique identifier for each blockchain transaction).

[1]Matching is performed using *Fuzzymatch*, a string-matching algorithm in the *Python* package *Fuzzywuzzy*. It matches strings based on their *Levenshtein-distance*, which is the number of edits required to match two strings. The larger the number of required edits, the lower the likelihood that these strings describe the same ICO (Cohen, 2020; Yujian and Bo, 2007).

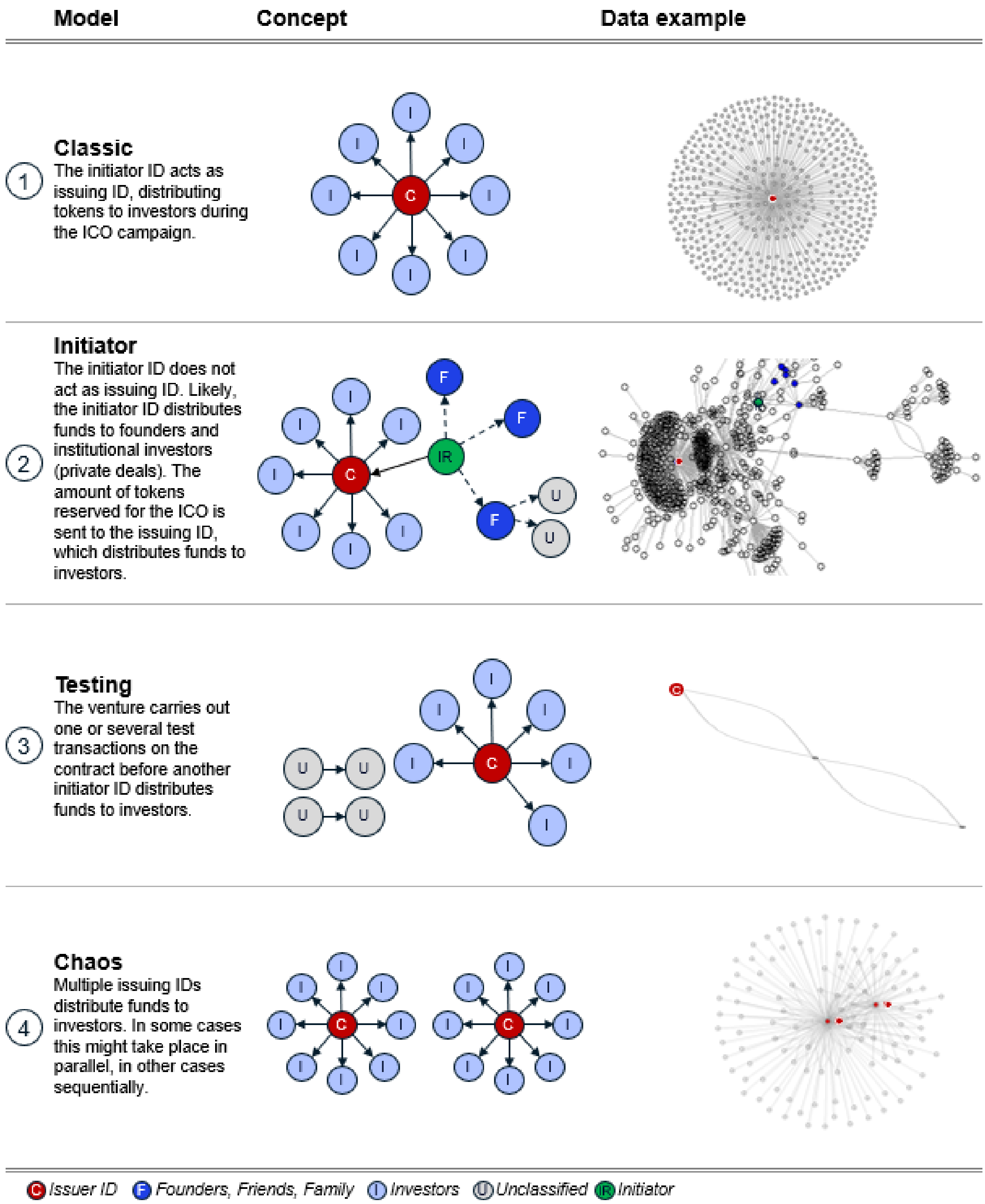
3.1.2 Identifying issuer IDs

The identification of issuer IDs is a critical prerequisite for further analysis. As described, our dataset contains information on the sender and recipient of each transaction. This information does not enable us to differentiate transactions initiating from the venture versus investors. Even though each contract has a creator and a contract ID, it might be either none or both of these IDs that act as issuer ID, namely the ID distributing tokens to investors. Distinguishing between issuer and investor is of critical importance for *two* reasons. First, treating issuers (hence ventures) as investors will introduce significant bias into our investor-based analyses. Second, the distinction allows for the differentiation of primary (fundraising) and secondary investments (purchases and sales on crypto-exchanges), critical for all *three* research studies. One might think that the differentiation between issuers and investors is simple by looking at the transaction volume or date. Surprisingly, this is often not the case.

Figure 3.2 visualizes *four* issuing patterns commonly observed in our data and shows a real-life example for each pattern. In *model 1* a unique issuing ID initiates trading on the ICO contract. Identifying the issuing ID is straightforward. The initiator ID, the ID performing the first transaction, is the issuing ID.

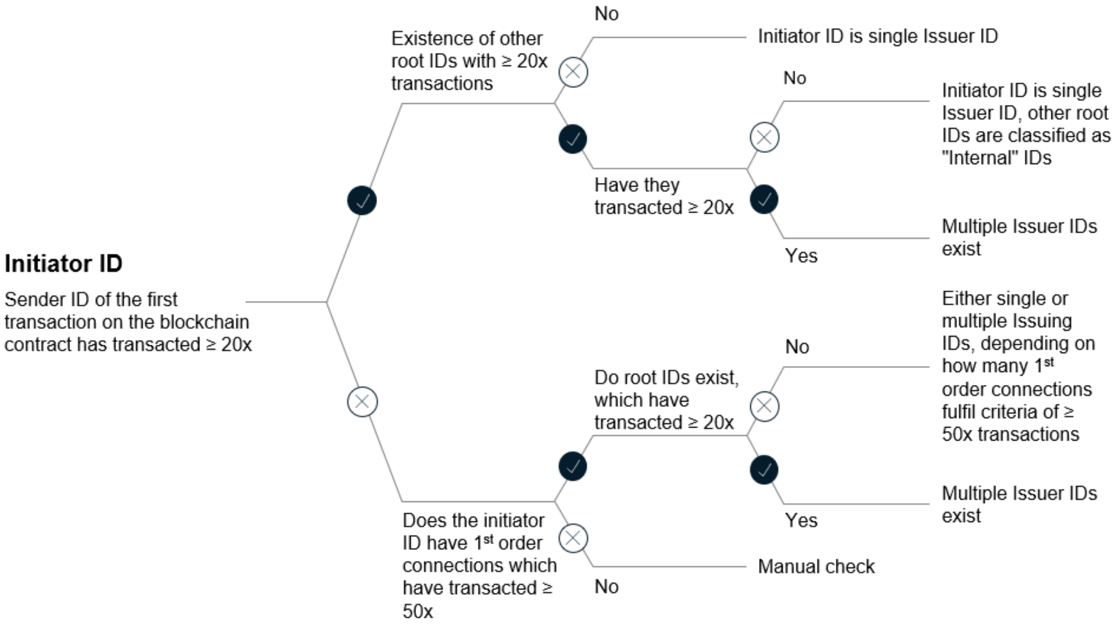
In *model 2* one ID distributes initial funds to several IDs, one of them being the issuing ID. The automated identification of the issuing ID becomes trickier compared to *model 1*. To identify the issuing ID, one must consider the initiator ID and its first-order connections (all IDs it directly transacted with). One or several, but not all, of these first-order connections, act as issuing IDs, identifiable only by the number of transactions performed. To add more difficulty, *model 3* has multiple root IDs (IDs with no predecessor), only one of them acting as issuing ID. Consequently, we must look at all root IDs and identify the issuer ID by the largest number of transactions. Last, under *model 4*, *two* independent issuer streams may exist. Therefore, our process must be flexible to allow for multiple issuing IDs.

FIGURE 3.2: Conceptualization of ICO issuing patterns



Note: Above figure shows the four archetypal ICO issuing patterns observed in our sample of 37 million blockchain transactions by 1,705 ICOs between September 2015 and year-end 2018. The data examples are based on Carboneum, Cartaxi, Caviar and Cartel coin ICO respectively.

FIGURE 3.3: Approach to identifying issuer IDs in ICOs



We have conceptualized the issuer identification process as shown in *figure 3.3*. The cut-off levels define whether the number of transactions performed by an ID is sufficiently large to assume it is an issuing ID. They have been set to at least *20* transactions for all root IDs and at least *50* transactions for all first-order connection of the root IDs. The rationale is the following: root IDs are always related to the venture. As such, they are by default classified as internal IDs and not considered an investor. If root IDs have transacted a few times, it could be for either of *three* reasons. First, transactions are test transactions, and no account should require more than *20* test transactions. Second, the root ID transfers to founders, family, friends & institutional investors, and the actual issuing ID. In this case, it is assumed that no more than *20* of these individuals exist. Third, the ICO has failed, and few investors purchased tokens. In this case, it is likely that with an average of *25,000* investors per ICO, *20* investors still showed interest in the campaign.

The cut-off level for first-order IDs is set to at least *50*, as in the case the ICO performs a *model 1*, *3* or *4* ICO. It is likely that first-order connections of the root ID, which are indeed investors, sell their tokens in more than *20* transactions. This step is only performed given that no root IDs have more than *20* transactions. This approach balances the Type I and II error of classifying investors as issuers versus correctly identifying an issuing ID.
We identify issuing IDs of *1,358* ICOs using this procedure. The remaining *347* ICOs are manually analyzed using a combination of network visualization and pattern identification of relative transactions between root IDs and their first- and second-order connections. Based on this classification, all IDs not classified as issuing IDs are regarded as investors. Unique wallet blockchain-IDs allow us to group investment and sales transactions per investor. For a given campaign, all investors who received their first investment in a given ICO from an issuer ID are classified as primary (fundraising) investors. All other investors are classified as secondary investors, as they bought shares from other investors on crypto-exchanges. We identify *9.4* million investors in *1,705* ICOs and *37* million transactions.

3.1.3 Imputation methodology

Our dataset had missing transaction values for *360* thousand transactions and transaction values smaller than *five* US Dollar for *6.9* million transactions. Although minimum investment amounts required by ICOs vary substantially, most ICOs set a minimum investment of at least *five* US Dollar. We assume that transaction values smaller than *five* US Dollar indicate data inaccuracies rather than reflecting the actual investment amount. The US Dollar transaction value is, unlike other variables, not directly observable in the blockchain. Rather, tokens are purchased in exchange for cryptocurrencies. Our dataset contains *Ethereum* transactions only, therefore, the final transaction is always made against *Ether* (the *Ethereum* currency). The value of a transfer expressed in US Dollar is a function of the historical *Ether* exchange rate and the token price, both of which vary during the fundraising phase. Neither *Ether* inflow nor token price is directly observable on *Ethplorer*. We thus rely on the US Dollar value reported by *Ethplorer*. We impute values for transactions with missing or inaccurate values as follows. We distinguish between investors with available values for some investments and those with missing values on all transactions. For the former group, we impute transaction value as the mean investment sum of their remaining investments. We decide to exclude the latter group's transactions, as too little data is available to estimate their investment values in this highly heterogeneous environment accurately. The reduced dataset consists of *4.9* million investors in *1,628* ICOs who conducted *26* million transactions.

3.1.4 Clustering subset of data

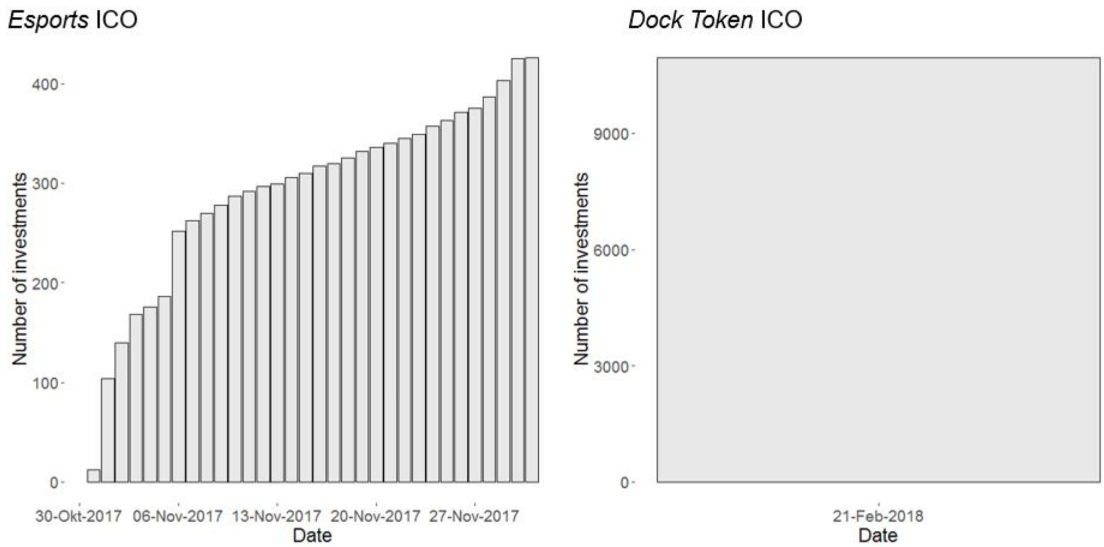
The cluster analysis presented in *chapter 4* presents a special case, as they are based on a subset of the full dataset and deserve some brief elaboration, further detailed in *chapter 4*. We exclude all transactions from one-off investors and investors with more than *3,500* investments across all ICOs in our sample in preparation for the cluster analysis. A single investment contains insufficient information on the investor to assign them to a cluster. IDs with more than *3,500* transactions are assumed to be exchanges executing transactions for a large number of investors. We will further elaborate on the underlying reasons in *chapter 4*. Our final dataset for clustering investors consists of *1.4* million investors in *1,628* ICOs and *19* million transactions.

3.1.5 Accurate recording of transaction time

While all transactions are recorded on the blockchain, our analysis shows that their timestamp may be inaccurate in some cases. *Figure 3.4* contrasts the issuing pattern for *two* ICOs, *Esports* and *Dock Token*, and shows how timestamp accuracy can vary between campaigns. *Esports* fundraising process is in line with a pattern commonly observed in crowdfunding (e.g. Hornuf and Schwienbacher, 2018) and consequently in line with our expectations. Investments are spread across the entire fundraising period, with steeper slopes (i.e., more investments) at the beginning and end of fundraising. For *Dock Token* however, our data lead us to believe that contributions were made on a single day. Our records also show that all transactions were executed within seconds of each other. We do some research and find that the ICO was announced to last *14* days. We suspect that orders of *Dock Token* were consolidated off the record, and the blockchain contract was used to distribute tokens to investors after the campaign ended, all at once.

While the order of investments is preserved for scenarios such as the one of *Dock Token*, timestamps of individual investments are inaccurate, lagging by a few days. The accurate measurement of investment timing is, however, crucial for the analyses presented in this dissertation. For this reason, we decide to exclude all ICOs which collected funds on a single day and scan the issuing pattern of the remaining ICOs for abnormalities, as shown in *figure 3.4*. While this may lead us to wrongly exclude some ICOs with accurate timestamps, the bias introduced by including inaccurate timestamps has more severe consequences for our analyses. The final dataset employed for the analyses presented in *chapter 5* and *chapter 6* contains *17* million transactions by *3.6* million investors across *1,274* ICOs, we excluded *354* ICOs in this cleaning step. We decide not to exclude these campaigns for our cluster analysis (*chapter 4*). Investors are, among others, clustered based on investment timing, makings it is accurate measurement crucial. However, we employ investment timing to distinguish primary versus secondary phase investments and investors' relative investment positions. Both measures are not affected by delayed transaction visibility. Consequently, the analyses presented in *chapter 4* continue to be based on the set of *4.9* million investors in *1,628* ICOs who conducted *26* million transactions.

FIGURE 3.4: Timestamp inaccuracy at the example of Esports and Dock Token



Note: Above figure visualizes the fundraising process of two representative ICOs, as observed in our blockchain transaction dataset. While Esports transactions are spread over a period of one month, Dock Token raised 100% of its capital in a single day.

3.2 Additional datasets

We enrich our blockchain transaction dataset with *four* additional datasets, used to varying degrees across our analyses, as summarized in *figure 3.5*. We will briefly elaborate on each dataset in this section. We detail discussions for the *ICO attributes*, and *social media sentiment* dataset in this section, as both are employed in at least *two* chapters, hence to circumvent repetitive discussion in each chapter. The remaining datasets are employed in a single chapter. They will be briefly introduced here and detailed in the respective chapters.

3.2.1 ICO attributes

We consolidate ICO attributes from *nine* ICO platforms: *ICOdata*, *ICORating*, *ICObench*, *ICOdrops*, *Coingecko*, *Cryptoslate*, *Tokendata*, *ICOwatchlist* and *Coincheckup*, prioritized in this order if a variable is available on more than one source.

Coincheckup was employed for a specific set of venture attributes and does not overlap with the other sources.

ICO platforms have no formal role within the campaign but are used as the first point of information by many investors. In their basic form, they summarize campaign design attributes (e.g., ICO-start date) as well as venture attributes (e.g., industry) and provide links to further sources of information (e.g., firm website or social media channels). Platforms may extend the basic content with (expert) ratings or discussion boards. We collect information on *41* attributes and match them with our blockchain transaction dataset using *Levenshtein-distance*, as detailed in *section 3.1.1*. A full list of all *41* attributes is shown in the *Appendix* on *page 163*. They can be broadly classified into campaign (*18*) and venture attributes (*23*).

Unfortunately, the dataset has missing values for a large number of ICOs for most variables. Furthermore, some of the variables measure similar characteristics, which may introduce multicollinearity issues to our analyses. We select *seven* attributes to be included in the analyses of *chapter 5* and *chapter 6*. These *three* campaign and *four* venture attributes are selected in accordance with *three* criteria: distinct explanatory power, meaningfulness and data availability.

We select campaign attributes in accordance with prior ICO research, e.g., Fisch and Momtaz (2020), and select *Pre ICO*, *Token price* and *Accept fiat*. The use of venture attributes in ICO research is novel, to our knowledge. Thus, we base our selection on venture capital research. Baum and Silverman (2004) argue that venture capitalists analyze *three* types of capital possessed by a venture when forming an investment decision: human, intellectual, and alliance capital. We measure human capital by *CEO experience* and intellectual capital by *GitHub* popularity of the project (*GitHub stars*). We attempted to measure alliance capital by including *Advisor large corp*, i.e., whether one or more of the ICO advisors have worked for a large corporation before. Unfortunately, the data quality is low as it is, based on a low number of overall observations.

FIGURE 3.5: Overview of datasets employed within this work

	Description	Source	Size	Paper 1	Paper 2	Paper 3
Blockchain transactions	Individual investor transactions as recorded on the blockchain, each attributed to a specific ICO campaign. Transactions by the same investor are consolidated by a unique investor ID across ICOs.	ethplorer.io	26 mn transactions across 1,628 campaigns by 4.9 mn investors	✓	✓	✓
ICO attributes	Set of 41 variables describing ICO campaign design, ICO marketing activities and venture attributes at the time of fundraising. 7 variables selected for our analyses: Pre ICO, Token price, Accept fiat, CEO experience, GitHub stars, Blockchain industry and Project status.	ICO data, ICORating, ICObench, ICOdrops, Coingecko, Cryotoslate, Tokendata, ICOwatchlist, Coincheckup	1,628 campaigns (3,209 initially identified, of which 1,628 can be linked with transaction records), missing data points along individual variables	(✓)	✓	✓
Social media sentiment	Weekly investor sentiment for the 100 largest cryptocurrencies by market capitalization, as measured by positive and negative referencing in user-generated social media sentiment.	MarketPsych TRMI	183 weeks between 2015-18 to match the observation period of our blockchain transaction data	✓	✓	✗
ICO post-campaign success				✗	✗	✓
Survival	Measures whether a venture is still active at our selected cut-off date (*12-08-2020*) .	twitter.com	198 campaigns			
BHAR	Buy-and-hold abnormal return earned by an investor who purchased the token upon listing and held it for 6 months, adjusted by the weighted market return.	coinmarketcap.com	149 campaigns			
Crypto market data	Weekly crypto market environment, based on price and trading volume of Ether, second largest crypto-currency by market	finance.yahoo.com	183 weeks between 2015-18 to match the observation period of our blockchain	✓	✗	✗

Note: Above figure provides an overview of all datasets employed within this work and points to the analyses they are included in. The use of campaign attributes varies across chapters. Within chapter 4 they are not directly included, but employed to calculate investor portfolio diversification.

Therefore, we decide to combine the approximation of human and alliance capital in the variable *CEO experience*. A more experienced CEO is expected to possess both a larger number of network ties and proven leadership capability (based on track record, i.e., survivorship).

We also include *Blockchain industry*, specifying whether a ventures core business is related to blockchain technology. This measures technology focus and may be used as a proxy for risk. Last, we include *Project status*, defined as *four* levels: *Just an idea*, *Minimum viable product (MVP)* available, *Beta version* available and *Working product*.

3.2.2 Sentiment dataset

We measure social media cryptocurrency sentiment on the basis of a dataset kindly provided by *Thomson Reuters' Refinitiv* in the form of their *MarketPsych* cryptocurrency index (*TRMI*). The *TRMI* measures general market sentiment as well as a large number of sentiment sub-dimensions for the largest *100* cryptocurrencies by market capitalization, as obtained from *coinmarketcap.com* and adjusted periodically. We employ this dataset for the purpose of our analyses in *chapter 4* and *chapter 5*. While *chapter 5* only employs general sentiment, *chapter 4* also includes *four* sub-dimensions: *Uncertainty*, *Emotion*, *Innovation* and *Anonymity*. These will be detailed in *chapter 4*.

We choose to focus on social media sentiment rather than traditional media sentiment (also available in the *TRMI* dataset). Prior ICO research, e.g., by Benedetti and Kostovetsky (2018) and Cerchiello (2018), shows that social media activity affects the fundraising success of ICOs. This supports anecdotal evidence that investors employ user-generated content in their decision-making.

TRMI social media sentiment is calculated based on a *LexisNexis* text corpus generated from social media content. The text corpus is analyzed for positive and negative connotations, depending on the specific dimension to be measured, to calculate (net) sentiment. Some indices are measured on a scale [0,1]. These indices count references to a certain sub-dimension (e.g., uncertainty). Other indices (e.g., overall sentiment) are measured on a scale *[-1,1]*. These indices are two-sided, meaning negative references are subtracted from positive references.

Based on the above-described sentiment dataset, we calculate a value-weighted index. The sentiment-weighting for each of the *100* coins is determined by their daily market capitalization, as obtained from *coinmarketcap.com*. For the purpose of *chapter 4* we convert daily measurements into week average, while daily estimates are used in *chapter 5*.

3.2.3 Other datasets

Venture success (ICO post-campaign success), is approximated for the analyses presented in *chapter 6* by *two* variables: *Survival* and *BHAR*. *Survival* is measured based on *Twitter* activity. A venture is said to have survived the observation period if they have tweeted a least once within *six* months of the cut-off date (*25-08-2020*). *BHAR* is the buy-and-hold abnormal return earned by an investor who purchased a token upon listing and held it for *six* months, adjusted for the market return during the same holding period.

We also include the general crypto market environment (*Crypto-market data*) by measuring *Ether*'s price level and volatility. *Ether* is the second-largest cryptocurrency by market capitalization and means of conducting transactions on *Ethereum* blockchain, which is the basis for our blockchain transaction dataset. This dataset is employed in the analyses presented in *chapter 4*. *Venture success* and *Crypto-market data* are employed in a single research chapter and will be discussed in more detail in the appropriate sections.

Chapter 4

Putting a face to the crowd - an investor-based exploration of Initial Coin Offerings

4.1 Abstract

Crowdfunding significantly lowers barriers to entry early-stage investments. ICOs pair this new asset class with high uncertainty, buzz, and liquidity through crypto-exchanges. This novel environment offers a unique opportunity to study crowd heterogeneity, of which not much is known today. We analyze *26* million blockchain transactions by *4.9* million investors in *1,628* ICOs. Employing cluster analysis, we detect behavior-based investor groupings and analyze their decision-making processes. We find evidence for pronounced investor heterogeneity. The crowd is dominated by one-off investors, while capital contribution is concentrated with *3%* of investors contributing *75%* of capital. Three distinct clusters exist: *Crypto Enthusiasts*, scanning the market for opportunities and actively managing their portfolio; *Risk-avoidant Rationalists*, concentrating their investments in a few hyped campaigns and *Passive Investors*, seeking long-term value.

4.2 Introduction

4.2.1 Motivation and background

Technological advances have the potential to transform the way early ventures raise capital by enabling and accelerating innovations such as crowdfunding. Crowdfunding significantly lowers barriers to entering early venture financing by providing digital platforms easily accessible to anyone and lowering the minimum investment - making it affordable to the broad masses. Consequently, a new high-risk asset class (early-venture financing), which professional investors historically dominated, opens up to retail investors.

As the dot-com bubble has taught us, the influx of uninformed investors combined with high uncertainty may trigger a dangerous market dynamic. Greenwood and Nagel (2009) show that informed investors deliberately drove up prices to sell overpriced shares to uninformed investors. It follows that high investor heterogeneity can have far-reaching implications on overall market dynamics.

Initial Coin Offerings (ICOs), a recent crowdfunding innovation in which early ventures raise capital through the issuance of cryptocurrencies, creates a particularly challenging investment environment for *four* reasons. First, the buzz surrounding cryptocurrencies led to a surge in ICOs through 2017-18. This led to a large number of inexperienced investors entering the market. With many ICOing ventures also actively seeking investment from professional investors, ICO crowds are likely characterized by pronounced investor heterogeneity. Second, ICOing ventures tend to be early-stage, many without a prototype in place, amplifying investors' valuation challenge. The absence of regulation also enabled an unfiltered cohort of ventures to attempt fundraising at a low cost. This likely results in a considerable variation in the quality of investment opportunities.

Third, the uncertainty surrounding the potential of blockchain technology led to tremendous market volatility. Even if a venture itself is not operating in the

blockchain industry, the value of its tokens (shares) depends on it. This uncertainty increases the difficulty of evaluating venture prospects and may reinforce the attraction of a crowd with different motives, background, and investment strategy, as compared to conventional crowdfunding.
Fourth, in contrast to conventional crowdfunding, ICO tokens are tradable on crypto-exchanges. As a result, ICOs lower barriers to entry early-venture investing through increased liquidity. We expect investors, especially inexperienced ones, to (over)value liquidity, as they tend to undervalue cashflows further in the future in favor of near-term payoffs. This phenomenon has been extensively studied in stock markets, e.g., by Black and Fraser (2002). As noted by Mochkabadi and Volkmann (2018) as well as Turan (2015), secondary markets could also reduce information asymmetries by introducing a mechanism for cascading information among investors. Nevertheless, they may also spur speculation and attract sophisticated investors aiming at exploiting arbitrage opportunities. In that sense, the hunter (crowd) could become the hunted.

4.2.2 Research question and methods

We infer from the above discussion that the ICO crowd is expected to be highly heterogeneous, which may have severe implications on market dynamics. Thus far, little is known about the composition of ICO crowds. Much of the existing research on ICOs has treated investors as a homogeneous group. Several research studies employ signaling theory to analyze how ventures can credibly signal quality (Amsden and Schweizer, 2018; Cerchiello, 2018; Benedetti and Kostovetsky, 2018). Findings suggest that investors employ campaign attributes (e.g., retained equity) as credible signals. Another stream of research has started to explore environmental cues (Momtaz, 2018b; Lee et al., 2019). Their findings show that investors also incorporate environmental cues in the form of investment decisions by others into their decision-making. We argue that in order to comprehend

ICO market dynamics, we first need to understand crowd composition and the decision-making processes. This paper aims to answer the following research question:

What is the composition of investor crowds in ICOs and how do investor groups differ in their investment behavior?

We attempt to answer this research question with a two-stage analysis. First, we profile ICO investors based on their investment behavior and cluster investors with similar behavior. This analysis enables us to answer the first part of our research question. Second, we investigate how environmental cues are employed in the decision-making of investor clusters. This analysis enables us to develop a more nuanced understanding of investor behavior, thereby answer the second part of our research question. Our analysis is built on a novel dataset of *26* mn ICO investments by *4.9* million investors between 2015 and year-end 2018.

4.2.3 Results and contributions

Our findings reveal that the ICO market is highly concentrated, with *75%* of capital contributed by only *3%* of investors. While novice investors dominate the crowd, our analysis suggests that significant heterogeneity exists among investors. Employing cluster analysis, we distinguish *three* investor clusters. *Crypto Enthusiasts*, who scan the ICO market for opportunities and actively manage their portfolios by diversifying and trading on exchanges. *Risk-avoidant Rationalists*, who concentrate their investment in a few campaigns that have gained market traction to minimize risk and realize short-term profits. *Passive Investors* who follow market trends and tend to invest for the longer-term.

Our findings contribute to the understanding of investor heterogeneity and decision-making in a novel investment environment (ICOs). We believe that our findings have practical implications for regulatory authorities. Ultimately, they aim at striking a balance between investor protection and nurturing innovation. Therefore, it is crucial to understand market actors and how individual behavior reflects on overall market dynamics.

4.2.4 Structure

The remainder of this paper is structured as follows. The first section discusses the theoretical background of ICOs and crowdfunding. We then draw on stock market research to develop a nuanced understanding of investor dynamics. The third section introduces the data sample and presents initial descriptive insights. Our analysis section is divided into *two* complementary parts. First, we perform cluster analysis in *section 4.5* to investigate the composition of the ICO crowd. Second, we employ regression analysis in *section 4.6*, allowing us to refine our understanding of investor clusters' investment behavior. We end with a discussion, elaborating on future research directions and limitations of this study.

4.3 Theoretical background

4.3.1 Initial Coin Offerings

Surprisingly little is known about the ICO crowd. The majority of research in this field takes on either of *two* lenses, the regulatory or the venture lens. Regulatory related research explores how regulators can strike a balance between fostering innovation while at the same time adequately protecting investors (An et al., 2018; Dell'Erba, 2018; Parrino and Greenslade, 2014; Bourveau et al., 2018; Chiu, 2018; Cumming and Johan, 2013; Hornuf and Schwienbacher, 2017).
A larger stream of venture related research applies signaling theory to analyze how ventures can send credible quality signals to investors (Amsden and Schweizer, 2018; Cordova et al., 2015; Fisch, 2019; Mollick, 2013; Cerchiello, 2018). Findings suggest that investors evaluate business and campaign-related quality signals,

such as a higher share of retained equity and team background (Amsden and Schweizer, 2018). Furthermore, investors positively interpret signals that reduce uncertainty and increase transparency, such as publishing publicly accessible code on *GitHub*[1]. One surprising finding is that of Amsden and Schweizer (2018), who find evidence that investors favor ICOs that report plans to list on exchanges subsequently. Their finding suggests that the existence of secondary markets attracts investors to ICOs.

Another stream of venture research has started to explore individual investor decision-making during ICOs (Amsden and Schweizer, 2018; Cerchiello, 2018; Adhami et al., 2018; Benedetti and Kostovetsky, 2018; Cordova et al., 2015; Felix, 2018; Momtaz, 2018b; Fisch, 2019). Their findings support the theory that investors include venture and campaign attributes in their decision-making. Yet, they also show that investors scan the environment for cues to aid their decision-making. Momtaz (2018b) and Lee et al. (2019) for example, find that investors place a higher decision-weight on alleged experts' ratings than on ICO-related cues. Similarly, Benedetti and Kostovetsky (2018) find that venture *Twitter* activity predicts fundraising success. Lee et al. (2019) build on these findings by investigating herding in ICOs. Their findings suggest that the number of first-day subscriptions is a strong determinant of follow-up subscriptions and ICO fundraising success.

While the understanding of funding success factors is advancing, we argue that a step forward is needed. To comprehend ICOs, we need to further the understanding of investors. All research to date has looked at ICO investors as a homogeneous group. The ICO *crowd* is likely more multifaceted. To aid our understanding of investor heterogeneity, we look at the related field of crowdfunding, where research on investor behavior has already advanced.

[1] A well-known coding repository, which enables developers to share and co-develop their projects within a community of developers (Github, 2020).

4.3.2 Crowdfunding

While much of the early research in crowdfunding has treated investors as a homogeneous group, much like in ICO research, an emerging stream of research explores investor-based data to investigate crowd heterogeneity (Abrams, 2017; Wick and Ihl, 2019; Wallmeroth, 2019; Moritz and Block, 2016; Lin et al., 2014; Hervé, 2016; Kim and Viswanathan, 2018; Hervé, 2016; Li et al., 2016).

Several studies build on herding literature, studying how cue interpretation depends on investors' (perceived) expertise. Kim and Viswanathan (2018) analyze investor dynamics on a crowdfunding platform for mobile applications. They find that investors tend to identify and follow early investors with expertise in the start-up's relevant stage. Similarly, Li et al. (2016) research Chinese crowdfunding campaigns and show that prior investments by lead investors positively affect funding success[2].

Wick and Ihl (2019) explore the heterogeneity of decision-making amongst expert versus novice investors. They differentiate investors by investment amount, timing, and frequency. Their findings suggest that investors interpret the behavior of other investors differently, depending on their background. Expert investors tend to herd after other expert investors and actively avoid crowds dominated by novice investors.

Other studies differentiate investors through investment timing. For example, Abrams (2017) investigates US crowdfunding campaigns. His findings show how decision-making differs between early and late investors in a campaign. While investments after the first week are closely tied to business fundamentals, first-week investors' investments are not. He concludes that this behavioral difference is attributable to investor backgrounds. Unsophisticated investors invest early, while more sophisticated investors closely evaluate business fundamentals, leading them to invest later.

[2]Lead investors are classified by the crowdfunding platform and are defined as experienced investors possessing substantial area-specific expertise.

Investment timing is particularly interesting for ICOs, as they present investors with an alternative option. While investment in conventional crowdfunding is only possible during the fundraising, ICO investors may opt to wait and invest on crypto-exchanges. This option may be costly for *two* reasons. First, ventures often provide discounts to attract investors during the early fundraising phase and gain market momentum. Second, research has documented high underpricing in ICOs (on average, *24%*) (e.g., Momtaz, 2018b; Lee et al., 2019). Even if the option to wait is costly, it allows investors to assess the general market opinion on a given ICO as expressed by the market price, consequently resolving information asymmetries.
Another research stream takes an investment strategy perspective on crowdfunding investors. These studies profile investors based on behavioral characteristics. Lin et al. (2014) study *1,904* projects on the crowdfunding platform *Kickstarter*. They find evidence for substantial heterogeneity amongst investors and classify them into *four* archetypes: active backers, trend followers, the altruistic, and the crowd. Their classification is based on behavioral characteristics, such as the popularity of each project, the number of projects backed, investor activity intensity, and diversification. Wallmeroth (2019) studies investor behavior in German crowdfunding campaigns. He approximates investor motivation by investment amount and finds that investors with contributions more than *5,000* Euro make up *50%* of capital, yet only represent *3%* of investors.
While we expect ICO and crowdfunding investors to share important characteristics, several features differentiate both markets; leading us to believe that ICOs may attract systematically different crowds, as discussed in *section 4.2.1*. Therefore, we investigate the crowds' composition in ICO markets and qualitatively compare it to the crowd in conventional crowdfunding.

4.3.3 Stocks markets and bubbles

To extend our understanding of dynamics among heterogeneous investors in markets characterized by high uncertainty and buzz, we turn to a well-studied phenomenon, stock market bubbles. Long before the development of crowdfunding, stock markets started the democratization of financing by allowing private individuals to provide financing to firms in return for ownership share. On stock markets, a large number of heterogeneous investors interact with each other to determine fair market prices.

Over the past decades, we have seen multiple episodes of bubbles, meaning sharply rising prices associated with subsequent price crashes, such as the housing bubble in 2006 and the dot-com bubble in 2000. Investor behavior during stock market bubbles has been extensively studied, amongst others by Morris (1996); Scheinkman and Xiong (2003); Chiang et al. (2011); Greenwood and Nagel (2009); Brennan (2004); Griffin et al. (2011). *Two* common and well-documented characteristics for recent bubbles are buzz and the consequent inflow of novice investors (Greenwood and Nagel, 2009; Binglin et al., 2017; Griffin et al., 2011). Research findings suggest that the composition of market participants changes in favor of inexperienced investors during bubbles (Griffin et al., 2011). This is interesting, as anecdotal evidence leads us to believe that both conditions hold for ICOs as well.

One bubble is of particular interest is the dot-com bubble. Between 1995 and 2000, the US equity market tripled in value, driven by internet companies (Clarida et al., 2013) whose technology and market value was uncertain. During the dot-com bubble, the stock market was characterized by similar circumstances as the ICO market today. The value of ICO tokens is also dependent on an uncertain technology - blockchain technology. The uncertainty surrounding ICOs is arguably even higher than that of internet companies in the 90s. Companies pursuing an IPO[3] have to obtain regulatory approval and usually have a track record. In contrast,

[3]Initial Public offering refers to the fundraising process on stock markets.

ICOs are mostly unregulated and usually at a very early stage.

Several studies have investigated investor heterogeneity and investor groups' interplay during stock market bubbles (Brennan, 2004; Greenwood and Nagel, 2009; Griffin et al., 2011). Greenwood and Nagel (2009) study how experience influences the investment behavior of fund managers. They find that inexperienced fund managers were more likely to exhibit trend-chasing behavior during the dot-com bubble. They conclude that experience, especially experience of prior similar circumstances (bubbles), makes investors less prone to chase trends.

Griffin et al. (2011) analyze the dynamics between institutional and retail investors during the dot-com bubble. They show that trend-chasing behavior by institutional investors may be attributable to rational speculation. Specifically, institutional investors purchase shares early, actively increasing prices. At the peak of the bubble, they observe a coordinated selling effort by institutional investors. Their findings suggest that experienced investors, opposing the efficient market hypothesis, do not trade against irrational market developments. Instead, they actively manipulate the market to exploit uninformed investors.

Binglin et al. (2017) study the phenomenon of persistent heterogeneity in markets. At the example of the Baosteel Call Warrant bubble, they show that a consistent inflow of novice investors causes persistent heterogeneity and bubble development. While individual investors learn, market-wide learning is prevented by the constant influx of new investors.

We expect ICO crowd composition and dynamics to resemble stock market bubbles due to the buzz and high uncertainty. Nevertheless, our current understanding of the actors involved is insufficient to draw conclusions on crowd dynamics. This study aims to develop a fundamental understanding of the heterogeneity of ICO investors. We aim to lay the groundwork to comprehend ICO crowd dynamics, direct future research, and aid decision makers by developing fact-based insights into this highly dynamic market environment.

4.4 Data and descriptive results

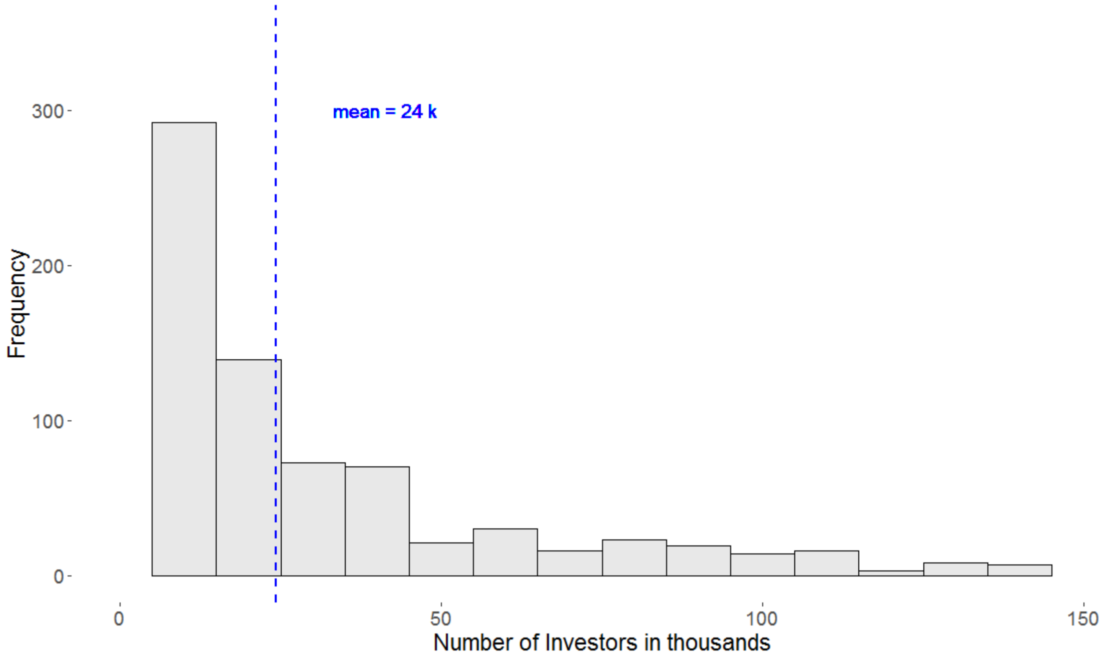
4.4.1 Data sample

All analyses presented in this work are based on *26* million blockchain transactions by *4.9* mn investors in *1,628* ICOs extracted from the blockchain explorer (*ethplorer.io*) for the period between September *2015* until year-end 2018.
We have imputed transactions with missing investment amounts whenever we had information on the average contribution of a given investor (based on other transactions by the same investor) and exclude transactions otherwise. This led to the exclusion of *11* mn transactions by *4.5* mn investors and *77* ICOs. Each transaction contains information on the *timestamp*, *blockchain ID* of sender and recipient, *transaction amount* and a unique transaction ID (called *Hash*). Please see *section 3.1* for a detailed discussion of dataset construction and cleaning steps.

4.4.2 Cluster descriptive results

We run some basic analyses to better understand the data before diving into our main analyses; some insights stand out. On the campaign level, the majority of ICOs in our data sample took place in 2018 (*60%*), *40%* in 2017 and only *12* in 2015-16. This is in line with our expectations. 2016 ICOs are likely underrepresented in our dataset. Data availability on ICO platforms is low for campaigns before the ICO-boom in 2017. Our sample is based on platform data and replicates this bias.
Figure 4.1 shows the distribution of investors per ICO. An average campaign attracted *25,000* investors. This result is driven by a small number of large ICOs, given that the median ICO only attracted *3,500* investors. While the latter appears low compared to *25,000* investors, a comparison to conventional crowdfunding highlights the impressive magnitude. According to Wallmeroth (2019), the average German crowdfunding campaign had *727* investors. The median ICO attracted *5x* more investors than the mean crowdfunding campaign.

FIGURE 4.1: Investors per ICO campaign



Note: The figure above shows the distribution of investor participation per ICO campaign, including investors in the primary stage (during fundraising) and those who purchased in the secondary market.

On the investor level, we find that the majority of investors have invested a few times during the observation period. Indeed, *70%* have invested once, *93%* have invested no more than *five* times. Building on this finding, as shown in *figure 4.2*, the ICO market, although large in size, is driven by the investments of a few hundred thousand investors only. *75%* of ICO funding is driven by *100,000* investors, equivalent to *3%* of investors[4]. This finding is in line with Wallmeroth (2019)'s finding for the German crowdfunding market. His data show that *50%* of investments are driven by *5%* of investors. Wick and Ihl (2019) also show that campaign success is driven by a few large investors.

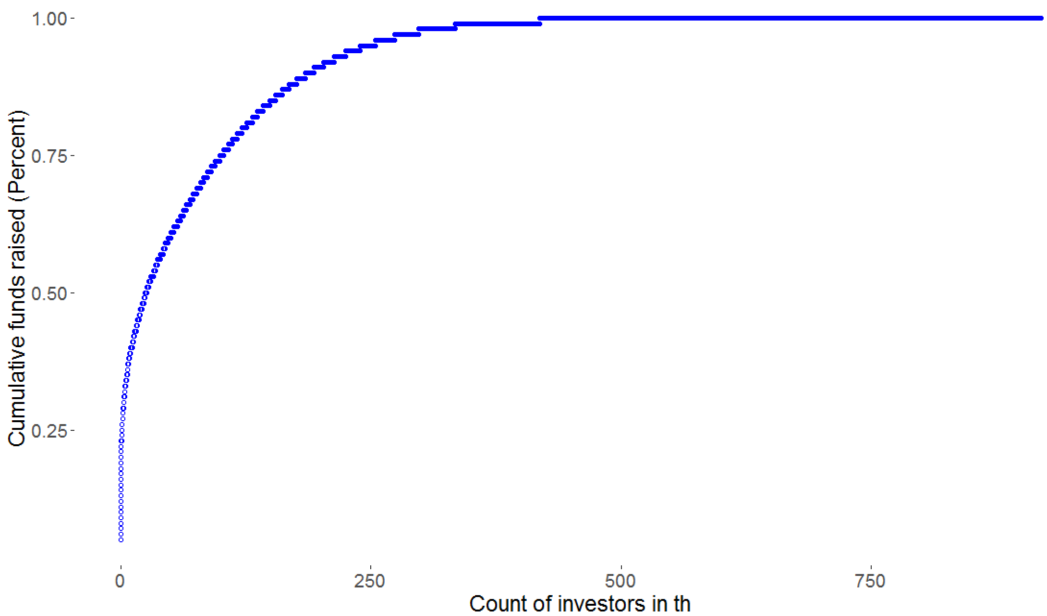
Our finding suggests that the ICO market might be even more concentrated than the crowdfunding market. One possible explanation is that only a rather knowledgeable and privileged subgroup of investors gain access to the often highly

[4]Considering primary investments (during the fundraising phase) only. Thus, we ignore the existence of secondary markets (crypto-exchanges) for the purpose of this analysis.

discounted tokens during the fundraising phase. The vast majority of investors subsequently buy shares on crypto-exchanges.

Indeed, *figure 4.3* shows the development of transactions over time and differentiates between primary and secondary market investments. Notably, across the entire period, secondary market transactions account for a share of about *80%* of all transactions. The decline in overall transactions starting in September 2018 as observable in *figure 4.3*, is influenced by, although likely not entirely attributable to, the fact that we only considered campaigns concluded in 2018 in our dataset. Therefore, we excluded campaigns starting in 2018 and running until 2019. Towards the end of 2018, the number of fundraising campaigns is declining, and crypto-exchange trading of concluded campaigns does not compensate for the declining primary activity (i.e., grey bars would remain at a stable level or decline at flatter slopes). We infer an interesting insight from this observation: exchange

FIGURE 4.2: Marginal contribution of ICO investors during fundraising (primary phase)



Note: This figure shows the marginal contribution per investor to overall capital raised during the fundraising phase (ICO). The calculation excludes any investments made in the secondary market (crypto-exchanges).

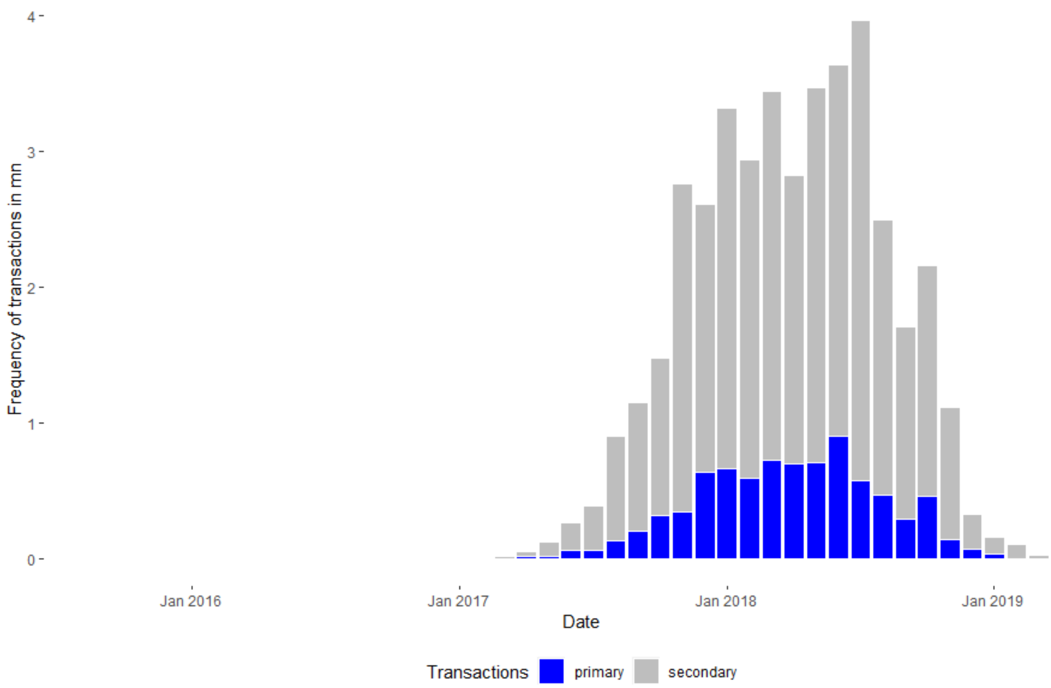
trading activity seems to be the highest right after listing on exchanges and declines afterward. In summary, our descriptive analysis suggests that pronounced heterogeneity exists among ICO investors. We set out to analyze this heterogeneity through cluster analysis.

4.5 Cluster analysis

4.5.1 Cluster data sample

The cluster analysis is based on a subset of the transaction dataset, as described in *section 4.4.1*, which excludes investors with a single investment during the observation period and investors with more than *3,500* investments across ICOs.

FIGURE 4.3: Distribution of transactions between 2015-18



Note: The figure above shows the transaction volume of 1,628 ICOs between 2015 and 2018. Grey bars represent overall transaction volume, including primary (fundraising) and secondary market (crypto-exchanges) transactions. The blue bars show primary transactions (fundraising) only.

A single investment contains insufficient information on the investor to assign them to a cluster. First-time investors belong to either of *three* groups: novice investors still exploring their investment strategy, sensation seekers attracted by the (social) media hype, or family & friends of the venture. This leads us to exclude *3.5* million one-time investors. As this group represents a large share of ICO investors, we do not disregard it but treat them as a pre-defined cluster of homogeneous investors.

We also exclude investors with more than *3,500* investments. Some IDs may represent exchanges, carrying out transactions on behalf of individual investors. A large number of smaller transactions should characterize such accounts. We assume that an individual investor would not invest in more than *10%* of ICOs in our dataset *(163)* and carry out more than *20* investments per ICO, on average. This means investors with more than *3,500* investments are excluded. *203* investors are excluded in this step. Our final clustering dataset consists of *1.4* million investors and *19* million transactions in *1,628* ICOs.

4.5.2 Cluster variables

We profile investors based on *four* behavioral dimensions: *Invested stake*, *Diversification*, *Timing* of investments and *Reselling activity*.

Invested stake measures how heavily an individual is invested in ICOs, *two* variables operationalize it: the number of ICOs an individual has invested in (*Invest campaigns*) and the total amount invested across ICOs, measured in US Dollar (*Invest amount*). Both, Wick and Ihl (2019), and Wallmeroth (2019) find evidence that investor behavior differs with the invested amount for crowdfunding campaigns. To our knowledge, this behavior has not been studied in ICOs.

Next, we measure the value-weighted industry diversification of each investors' ICO portfolio (*Diversification*). We extract industry tags from the ICO platform *icobench.com* and supplement tags for missing ICOs manually. An ICO can have several industry tags. For example, *Ripio* is a venture attempting to create a global credit network to connect lenders and borrowers around the world, based on *smart contracts* (icorating.com, 2017). According to *icobench.com*, *Ripio* operates in both the *platform* and *banking* industry. Simply counting the number of industry tags in an investor's ICO portfolio would not sufficiently approximate true investor diversification for *two* reasons. First, if an ICO has, let's say, *four* industry tags, then counting the tags would lead us to believe that the investor has achieved diversification with a single investment. Second, some industries are more similar than others. For example, we argue that an investor who has invested in one *banking* and one *payment* ICO is less diversified than an investor who has invested in one *art* and one *banking* ICO. To overcome this challenge, we follow the approach of Goldberg et al. (2015). They employ *modified Hausdorff distance* to measure variety, which measures the distance of objects (investments) based on assigned categories and their co-occurrence in the past in three steps.

First, we calculate pairwise distance of the *29 icobench.com* industry tags, which is based on co-occurence patterns of each pair within our data sample. We follow Goldberg et al. (2015)'s approach to measure pairwise distance by combining *Jaccard Similarity* (*J*) with *Shepard's Universal Law*. *Jaccard Similarity* is defined as the co-occurence of *two* objects divided by the number of individual occurrences yielding a range [0,1] where zero denotes perfect dissimilarity (no co-occurence):

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

Combining it with *Shepard's Universal Law* and solving for the distance between k and l we obtain:

$$d(k,l) = -\frac{ln(J(k,l))}{\gamma}$$

where the sizing parameter γ is, in accordance with prior literature (e.g., Goldberg et al., 2015), set to *0.5*. Second, we calculate the average pairwise distance between ICOs in our sample using *modified Hausdorff distance*, originally introduced by Burago et al. (2001), as extended by Dubuisson and Jain (1994) and also followed in a socio-cultural context by Goldberg et al. (2015). It calculates the average distance between all industry tags of *ICO A* and the closest industry tag of *ICO B* and vice versa. Letting C_A denote an industry tag set of *ICO A*, c_A be a specific industry tag and N_A the number of industry tags of set A, we get:

$$h'(C_A, C_B) = \frac{1}{|N_A|} \sum_{c_A \epsilon C_A} min(d(c_A, C_B))$$

for *set A* and vice versa for *set B*. From this we obtain the modified *Hausdorff distance* between the sets as:

$$HD'(C_A, C_B) = max(h'(C_A, C_B), h'(C_B, C_A))$$

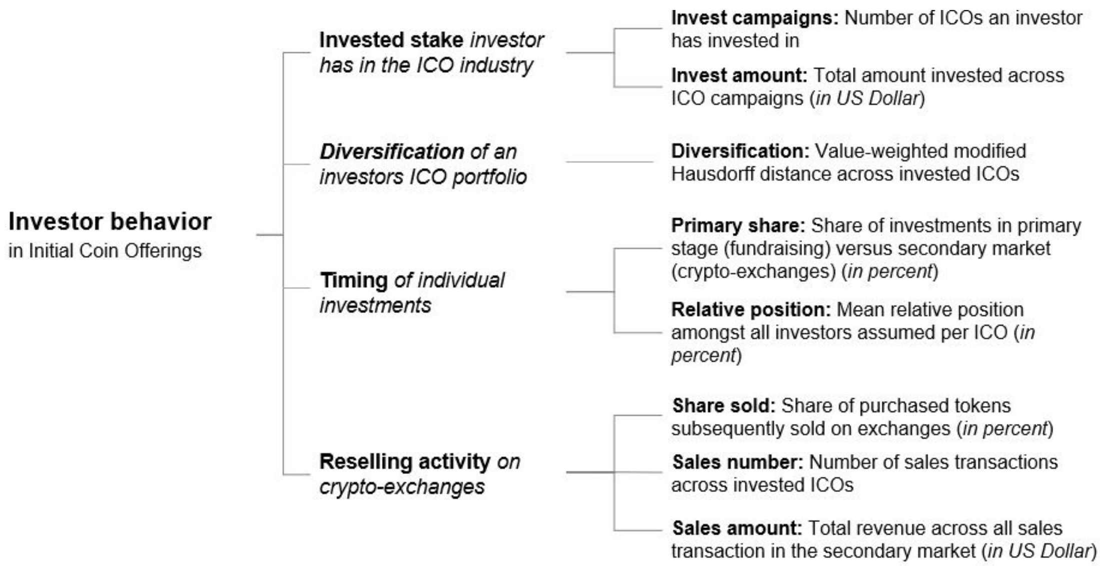
Third, we calculate investor diversification as the value-weighted pairwise distance between ICOs in an investors' portfolio.

We need to make an important remark. It has been shown in finance research that diversification of investments is beneficial to risk reduction of an investment portfolio (Markowitz, 1939). Diversification is most effective when applied to asset classes with low correlation. Even though individual ICOs may differ in their underlying industry, arguably all ventures attempting an ICO are dependent on the prospects of blockchain technology. Investors would therefore be better off not only to diversify within their ICO portfolio, but rather to also diversify it with other asset classes. This information is not observable in our data. While we believe that we achieve a good approximation of ICO portfolio diversification with our approach, the informative value of our *Diversification* variable is limited to the ICO market and thus a restricted measure of true investor diversification.

Our third behavioral dimension is investment *Timing*. It measures how early an investor tends to join ICOs. Abrams (2017) analyzes investment timing in crowdfunding campaigns and finds that early investors' decision-making cues differ from those who join later. We measure *Timing* by *two* variables, share of primary investments (*Primary share*) and average relative position (*Relative position*). *Primary share* measures the percentage of investments a given investor made during fundraising versus in the secondary market (crypto-exchanges). Mean *Relative position* describes the relative position an investor usually takes amongst all investors. For example, if *1,000* investors had invested in a campaign, and investor A was the *10th* contributor, her relative position in the campaign is expressed as *1%*, as she was amongst the first percent of investors to contribute. This value is averaged across all invested campaigns for a given investor. Investors are incentivized to join campaigns as early as possible, with many ICOs heavily discounting their tokens during the early stages to attract investors. The *Relative position* hence enables us to draw more nuanced insights from investment *Timing* as compared to *Primary share*.

Last, we measure the *Reselling activity* of investors, as operationalized by *three* variables: the share of purchased tokens subsequently sold on crypto-exchanges (*Share sold*), the number of sales on exchanges (*Sales number*), and the total amount realized during sales transactions (*Sales amount*). This dimension is of particular interest, as it is unique to ICOs. Prior research pointed to the introduction of secondary markets as an opportunity to overcome information asymmetries and has called for research to analyze its effectiveness (Turan, 2015; Mochkabadi and Volkmann, 2018). In that sense, ICOs represent a novel environment to explore the effects of introducing secondary market trading to a crowdfunding environment. *Figure 4.4* summarizes all variables employed to profile ICO investors in our cluster analysis.

FIGURE 4.4: Cluster variables: characterizing investment behavior of ICO investors



4.5.3 Descriptive statistics

Table 4.1 shows the descriptive statistics for our clustering dataset of *1.4* mn investors in *1,628* ICOs. The mean investor has invested *2,290* US Dollar in *five* ICOs or *460* US Dollar per ICO, on average. The median investment is lower at *320* US Dollar per campaign. This indicates that heterogeneity persists in the dataset, despite the exclusion of one-time and high-frequency investors. *Diversification* takes a mean value of *4.76*. The range, not observable from the table, is *[0,13.8]*. While we observe investors' tendency to diversify, our initial results indicate varying diversification among investors. Furthermore, the average investor buys on exchanges 75% of the time, while only participating in fundraising in every fourth investment. This indicates that the majority of investments are carried out in secondary markets. Whether this is deliberate or if investors could not secure tokens during fundraising due to high demand and a limited supply of tokens remains unclear. The mean *Relative position* is by definition set to around *50%* and does not require further attention in this analysis. Last, the *three* variables measuring *Reselling activity* point to significant heterogeneity among investors. While the average investor takes part in *four* sales transactions (moderately active on crypto-exchanges), selling for *1,258* US Dollar in total and ∼*315* US Dollar per transactions, the median investor does not take part in any sales transactions.

In preparation for our cluster analysis, it is necessary to make some adjustments to our dataset. As *column 5* and *6* indicate, all variables violate the assumptions of normality. We log all clustering variables, as shown in *table 4.2*. The transformed variables approximately meet the assumptions of normality. In a final data processing step, we standardize each variable by calculating the mean deviation, expressed in units of standard deviation, the *z-scores*. This step is important for cluster analysis to make variables that are measured in different units comparable. *Table 4.3* shows the correlation matrix for all variables. The majority of variables exhibit moderate correlation, indicating that each variable measures a distinct investment behavior aspect. Two variables, namely *Sales number* and *Sales amount*, exhibit high positively correlations. One could consider combining the variables using principal component analysis (*PCA*). As each variable has some distinct explanatory power, we decide on another approach. We include all variables in our cluster analysis and assess the sensitivity of results to changes in this *two* variables subsequently.

TABLE 4.1: Summary statistics of clustering data sample

Variable	Mean	Median	STD	Skew	Kurtosis
Invest campaigns	5	3	6.72	7.52	117.55
Invest amount	2,290	952	11,423	62.76	> 500
Diversification	4.76	4.50	2.28	1.11	5.46
Primary share	25%	17%	0.29	0.92	2.98
Relative position	49%	49%	0.20	0.13	2.51
Share sold	28%	0%	0.41	0.98	2.14
Sales number	4	0	564	> 500	> 500
Sales amount	1,258	0	153,715	> 500	> 500

Note: Above table reports the summary statistics for 1.4 mn investors in 1,628 ICOs as employed in our cluster analysis. Amounts are measured in US Dollar and represent the total amount invested/ sold. Relative position is a mean value of relative positions assumed by a given investor across invested campaigns.

TABLE 4.2: Summary statistics of logged clustering data sample

Variable	Mean	Median	STD	Skew	Kurtosis
Invest campaigns	1.41	1.10	0.77	1.16	4.01
Invest amount	6.82	7.09	1.68	-0.42	2.89
Diversification	1.65	1.65	0.42	-0.67	5.48
Primary share	0.17	0.09	0.19	0.80	2.45
Relative position	0.39	0.39	0.14	-0.03	2.43
Share sold	0.26	0.03	0.30	0.51	1.45
Sales number	0.84	0.69	0.96	1.21	5.04
Sales amount	3.61	4.20	3.44	0.11	1.35

Note: Above table reports the summary statistics for 1.4 mn investors in 1,628 ICOs after they have been logarithmized. Amounts are measured in US Dollar and represent the total amount invested/ sold. Relative position is a mean value of relative positions assumed by a given investor across invested campaigns.

TABLE 4.3: Correlation of clustering variables

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Invest campaigns	1							
Invest amount	0.60	1						
Diversification	0.06	0.02	1					
Primary share	0.05	(0.08)	(0.05)	1				
Relative position	(0.08)	(0.11)	0.01	(0.32)	1			
Share sold	0.04	0.12	(0.04)	0.02	(0.14)	1		
Sales number	0.40	0.41	(0.01)	0.01	(0.20)	0.72	1	
Sales amount	0.28	0.35	(0.02)	0.03	(0.20)	0.81	0.91	1

Note: Above table above shows the Pearson correlation matrix for clustering variables and is based on the final clustering input dataset for 1.4 mn investors and 26 mn transactions after transformation (logarithmized and standardized). Brackets represent negative values.

4.5.4 Cluster methodology

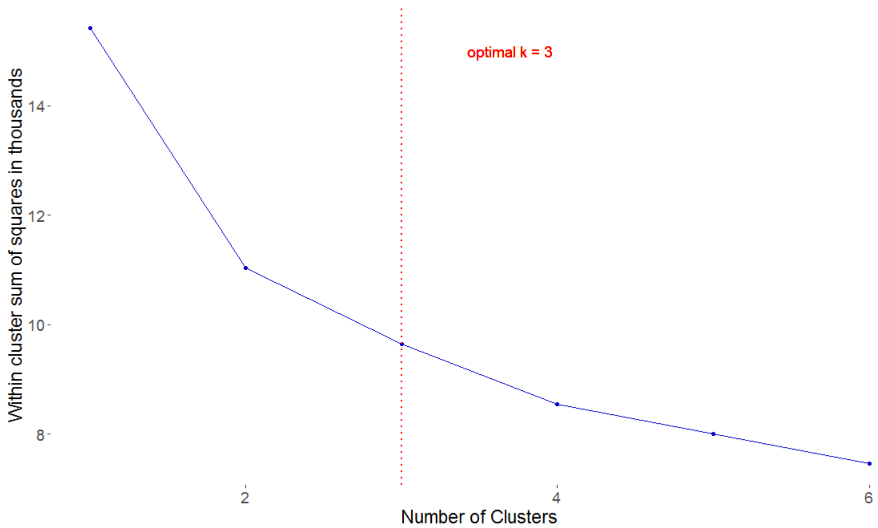
We employ k-means clustering to identify groups of investors with similar investment patterns. K-means clustering is an unsupervised machine learning algorithm that, based on a quantitative comparison of observation similarities, detects natural groupings in data. Clusters are formed so that the squared error between observations in a given cluster and its centroid - the weighted average mean of all cluster observations - is minimized for all clusters (Berkhin, 2002; Jain, 2010; Hartigan and Wong, 2012). While k-means clustering is a powerful technique in general, several characteristics make its implementation challenging.

First, due to its reliance on centroids, k-means clustering is sensitive to outliers. It is of particular importance that the normality assumption on the underlying dataset is not violated (Datanovia, 2019). To account for this, we perform normality tests and data transformation, as discussed. We also employ sensitivity analyses using an alternative algorithm that is less sensitive to outliers to test our results' robustness.

Second, the k-means algorithm normally initiates at a random initial cluster result and incrementally optimizes the objective function from there. Therefore, the algorithm may converge to a local minimum, leading to a suboptimal solution. A common remedy is to run several initializations on a given number of clusters and choose the global minimum obtained across all solutions. In our analysis we initialize *50* times.

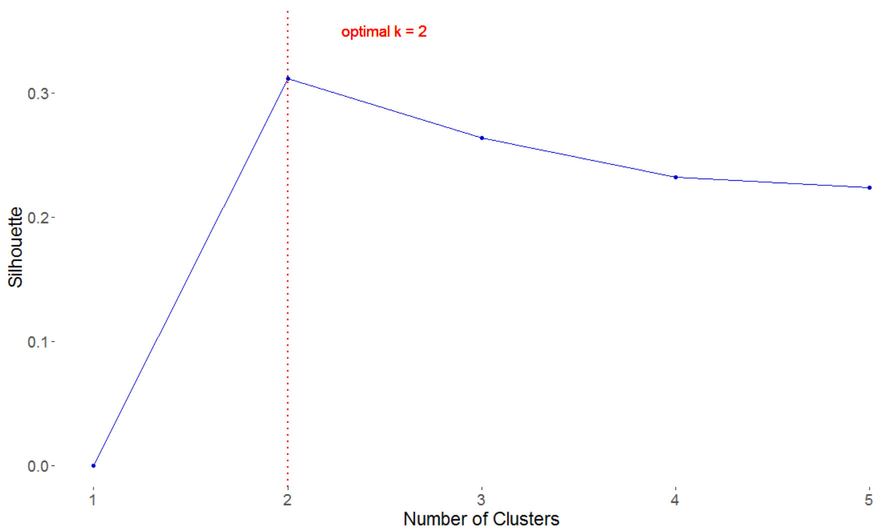
Third, k-means clustering requires the number of clusters (k) to be determined in advance. To aid decision-making, we perform *three* common heuristic analyses: *WCSSE*-, *silhouette*- and *dissimilarity* comparison across possible solutions. *WCSSE* measures the within-cluster sum of squares, the objective function of k-means clustering. The optimal number of clusters is the solution for which adding another cluster that does significantly lower *WCSSE*, the so-called elbow (Datanovia, 2019). As *figure 4.5* indicates, in our dataset this occurs around adding the third cluster.

FIGURE 4.5: Determining the optimal number of clusters: WCSSE analysis



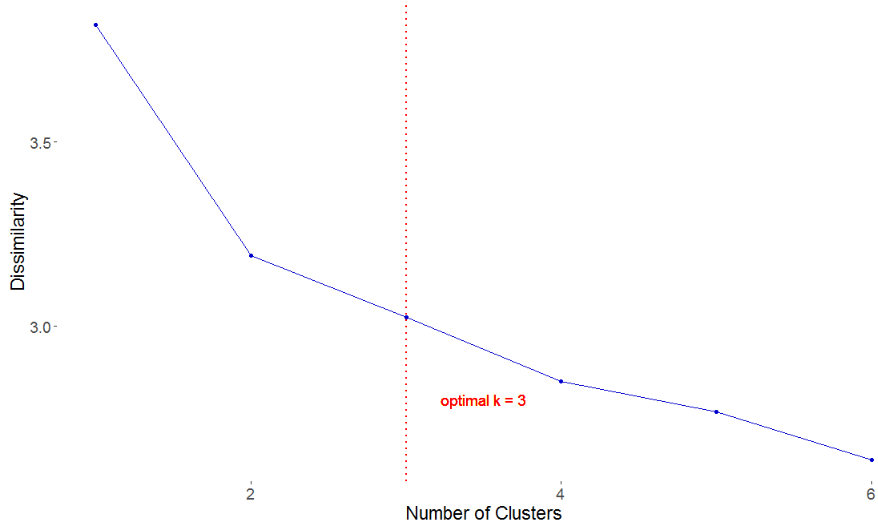
Note: The figure above shows the within cluster sum of squares (WCSSE) for k-means clustering solutions with 1-6 clusters. Solutions with lower WCSSE are desirable. Employing the "elbow principal" the optimal k should be chosen so that adding another cluster does not significantly reduce WCSSE. The analysis suggests 3 is the optimal number of clusters for our dataset.

FIGURE 4.6: Determining the optimal number of clusters: Silhouette comparison



Note: The above figure visualizes the results from a silhouette comparison for k-means clustering solutions with 1-5 clusters. The silhouette score measures the within cluster homogeneity of a given solution by measuring the mean distance of observations within a cluster. A higher average silhouette score [-1,1], indicates better clustering solutions. The analysis suggests a 2-cluster solution to be optimal.

FIGURE 4.7: Determining the optimal number of clusters: Dissimilarity analysis



Note: The above figure summarizes the results from a dissimilarity analysis for clustering solutions with 1-6 clusters. The dissimilarity criteria measures the average intra-cluster dissimilarity across clusters for a given solution. Higher dissimilarities indicate better clustering solutions. The analysis suggests 3 to be the optimal number of clusters for our dataset.

According to the elbow method, the optimal number of clusters is *three*.

The silhouette score measures within-cluster homogeneity by calculating the mean distance of observations within a cluster. Higher average silhouette scores indicate better clustering solutions (Rousseeuw, 1987). *Figure 4.6* summarizes the results of the silhouette analysis. The average silhouette score is highest for a *two*-cluster solution and declines after that. Therefore, our analysis would guide us to the *two*-cluster solution (Datanovia, 2019).

Last, the dissimilarity criteria measure the average intra-cluster dissimilarity across clusters for a given solution. Higher dissimilarities indicate better clustering solutions, given that we want distinct clusters. As shown in *figure 4.7*, cluster dissimilarity declines with an increasing number of clusters (Datanovia, 2019). Applying the elbow method here indicates that *three* is the optimal number of clusters. We have cross-validated our findings with an automated solution provided by the *R-package Nb-Clust (Charrad et al., 2014)*, which calculates 23 criteria for evaluating the optimal number of clusters. While the results are omitted for brevity, *seven* criteria suggest a *three* clusters, while another *seven* criteria suggest *two* clusters. Together with our previous indications, we decide to pursue our analysis with *three* clusters.

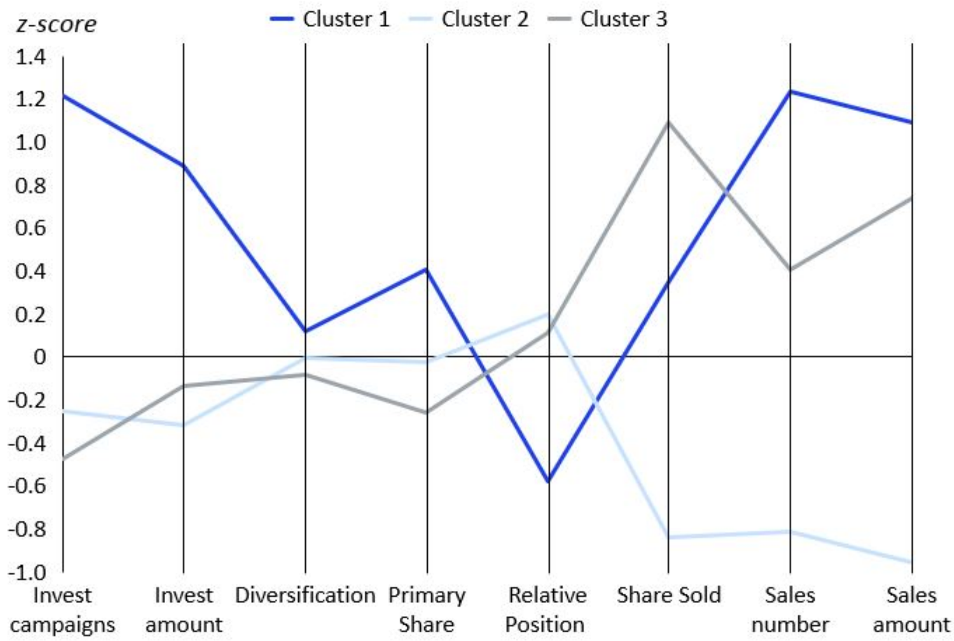
4.5.5 Cluster results

The results of our k-means cluster analysis are shown in *figure 4.8* and *table 4.4*. The values for each cluster represent the *centroid*, the weighted average value taken by investors within the clusters. Let us understand each of their patterns individually, along the *four* behavioral dimensions: *Invested stake, Diversification, Timing* and *Reselling activity*.

Cluster 3 invests in comparatively few ICOs, with a median of *three* campaigns, but invests more than the median investor (*952* US Dollar) if they do (*1,009* US Dollar). These individuals invest rather late, based on the lowest value for *Primary share* and second highest values for mean *Relative position*. Of all clusters, *Cluster 3* sells the highest share of purchased tokens subsequently. *Table 4.4* shows that a typical *Cluster 3* investor resells *100%* of their tokens. They also have the lower *Diversification*. Our findings lead us to interpret *Cluster 3* as opportunistic investors who scan the market for investment opportunities with the primary goal to realize profits through reselling on secondary markets quickly. Their relatively late investment leads us to believe that they are insecure about which ICOs represent good investment opportunities, and therefore wait for an established market opinion. This could also be interpreted as trend-chasing behavior, which in turn may lead us to believe that, beyond the financial motivation, *Cluster 3* investors also have a fear of missing out on a historic opportunity.

Cluster 2 has the lowest *Invested stake*, with a total investment amount of *743* US Dollar and on average ∼*250* US Dollar per ICO, even though they invest in as many ICOs as *Cluster 3*. *Cluster 2* investors purchase slightly earlier than *Cluster 3* investors (higher *Primary share*, however also slightly higher *Relative position*).

FIGURE 4.8: Cluster results A (z-scores)



Note: This figure shows cluster results based on k-means clustering of 1.4 million investors based on their investment behavior. Cluster 1 contains 315,938 investors, 695,080 and 433,948 investors were placed in Cluster 2 and 3 respectively.

TABLE 4.4: Cluster results B (values)

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Invest campaigns	10	3	3	1
Invest amount	4,653	743	1,009	189
Diversification	4.5	4.1	4.0	0
Primary share	30%	0%	0%	0%
Relative position	40%	50%	50%	50%
Share sold	40%	0%	100%	0%
Sales number	6	0	2	0
Sales amount	1,822	0	598	0

Note: The table above summarizes median cluster values obtained by a k-means algorithm with 3 clusters. Values are untransformed, interpretable in their original unit. Amounts are measured in US Dollar. Cluster 1 contains 315,938 investors, 695,080 and 433,948 investors were placed in Cluster 2 and 3 respectively, the one-time investor group (Cluster 4) contains 3.5 million observations.

These investor clusters are most differentiated in their *Reselling activity*. While *Cluster 3* investors heavily trade on exchanges, *Cluster 2* sell few to none of their acquired tokens (lowest *Share sold*). We interpret *Cluster 2* investors as value investors. They systematically scan the environment for high-quality ICOs that they believe in and purchase them for their long-term intrinsic value.
Last, *Cluster 1* is heavily invested in the ICO industry with *10* campaigns and on average *465* US Dollar investment per campaign. This investor group seems to invest the most capital and spread their risk most widely amongst ICOs, as indicated by the comparatively high *Diversification*. Interestingly, *Cluster 1* investors are earliest to invest in campaigns (highest *Primary share* and lowest *Relative position)*. In contrast to *Cluster 3*, they retain a larger share of purchased tokens (*Share sold*), yet also selling *40%* of purchased tokens. We interpret *Cluster 1* as crypto enthusiasts who are well-informed market participants (invest early and often). They invest heavily compared to other investors and are assumed to be willing to take risks who trust in their analysis (instead of waiting for a market opinion to evolve). We suspect their investment motive to be a combination of financial incentives (active trading on exchanges), value investing (retention and high investment amount), as well as early adopter excitement about taking part in the development of new technology (overall market activity).
Table 4.4 summarizes the findings represented as median local unit values per variable and cluster rather than *z-scores*. Here, we have also included the median values for the investor group previously excluded from the dataset for clustering, such as one-time investors (*Cluster 4*). Suppose we were to view these investors as a homogeneous group. In that case, they tend to buy and hold their rather low investments of *189* US Dollar in a single campaign (no *Reselling activity*) and, similar to *Cluster 3*, tend to enter the stage rather late.

4.5.6 Sensitivity analyses

We perform several sensitivity analyses to test our findings' robustness; the results are summarized in *table 4.5*. First, as discussed in section 4.5.4, due to being based on *centroids* (mean values), k-means clustering is sensitive to outliers. We have accounted for this by logging and standardizing our data. Nevertheless, testing the sensitivity of our results using a method less prone to outliers helps evaluate our results' robustness.

To do this, we employ k-medoids clustering. In contrast to k-means clustering, the k-medoids algorithm chooses an actual observation as a cluster center (instead of the weighted mean). The cluster *medoid* minimizes the difference with all other observations in that cluster, representing a typical investor in that cluster. Therefore, a k-medoids clustering algorithm is less sensitive to outliers (Arora and Varshney, 2016; Berkhin, 2002; Velmurugan and Santhanam, 2010).

Computation complexity is significantly higher for k-medoids algorithms. Due to the large size of our dataset, we switch to the more efficient Clara algorithm[5], drawing multiple samples from the dataset and retaining the best solution across samples. We draw *1,000* samples, each containing *15,000* observations.

As shown in *section B* of *table 4.5* the results are similar to those obtained from k-means algorithm. One interesting difference is observed: in contrast to k-means results, the k-medoids algorithm arrives at lower investment and sales amounts for *Cluster 1*, as compared to our main analysis (*A*). This indicates a particularly high dispersion of total investment and sales amount for investors in this cluster. With regards to the remaining variables and cluster classification, the results remain stable. We conclude that our initial analysis is robust concerning data outliers.

We perform an additional sensitivity analysis to test whether our main results are biased by multicollinearity. As discussed in *section 4.5.3*, some variables in our dataset exhibit high positive correlations. Such a high correlation introduces

[5]Clara refers to clustering large applications

TABLE 4.5: Sensitivity of cluster results

Analysis	Cluster	Invest campaigns	Invest amount	Diversi-fication	Primary share	Relative position	Share sold	Sales number	Sales amount
(A)	1	10	4,653	4.5	26%	35%	44%	6	1,822
Main	2	3	743	4.1	0%	50%	0%	0	0
analysis	3	3	1,009	4.0	0%	52%	100%	2	598
(B)	1	8	3,546	4.4	29%	34%	48%	5	1,342
k-Medoids	2	3	733	4.1	0%	52%	0%	0	0
	3	2	970	4.1	0%	55%	100%	2	581
(C)	1	9	4,932	4.5	22%	37%	46%	6	1,861
PCA	2	3	728	4.1	0%	52%	0%	0	0
	3	3	941	4.0	0%	49%	99%	2	538

Note: The above tables compares our main results with the results of 2 alternative estimation approaches. The results from analysis (B) are obtained employing a Clara (sample based) k-medoids algorithm instead of k-means clustering. Analysis (C) uses principal component analysis (PCA) to condense variables to 5 components. All amount related variables are measured in US Dollar.

multicollinearity and may consequently bias our results. We test for multicollinearity biases by performing *principal component analysis (PCA)* to transform our clustering dataset. *PCA* is a dimensionality reduction procedure, which condenses correlated input variables to a set of uncorrelated *principal components (PCs)* (Jolliffe et al., 2016; Wold et al., 1987; Abdi and Williams, 2010). We find that *five PCs* account for *89%* of data variation. We hence include the first *five PCs* in our clustering analysis and apply k-means clustering with three clusters. The results are summarized in *table 4.5*. Again, our results are almost identical to those of our main analysis. We conclude that our results are robust with regards to multicollinearity. To further qualify investor clusters, we perform a second set of analyses to understand investor behavior by studying how environmental cues influence investment decisions.

4.6 Regression analysis

4.6.1 Regression data sample and variables

To study investor decision-making, we switch the unit of analysis. We employ the transaction dataset, as described in *section 4.4.1* and enrich it with the cluster affiliation of each investor, as obtained from our cluster analysis in section 4.5. The resulting dataset consists of *4.9* mn investors in *1,628* ICOs who conducted *26* mn transactions. We further enrich our data with environmental cues investors may consider when making investment decisions.

First, we add social media sentiment. *Thomson Reuters' Refinitiv* has been so kind to provide us with data from their *MarketPsych* cryptocurrency index (*TRMI*). The *TRMI* measures general market sentiment as well as a large number of sentiment-dimensions for the largest *100* cryptocurrencies based on their market capitalization. Please refer to *section 3.2.2* for a detailed description of the dataset. We select sentiment $Sentiment_{t-2}$, and *four* sentiment-dimensions: $Uncertainty_{t-2}$, $Emotion_{t-2}$,

$Innovation_{t-2}$ and $Anonymity_{t-2}$ for our regression analysis. $Uncertainty_{t-2}$, $Innovation_{t-2}$ and $Anonymity_{t-2}$ are one-sided indices, measured on a scale *[0,1]*. $Sentiment_{t-2}$ and $Emotion_{t-2}$ are two-sided, subtracting negative references from positive ones, the resulting indices take the range *[-1,1]*.

All indices are aggregated to a weighted average of the *100* largest cryptocurrencies by market capitalization. Moreover, each index is lagged by *two* weeks to capture the influence of sentiment at the time of decision formation. The preliminary analysis provides confidence for our lag choice, showing that a two week lag has a larger effect on investment decisions than smaller lags (one week lag or real-time).

We analyze the correlation among all *five* sentiment indices to ensure each variable measures a distinct behavioral aspect. As *table 4.6* shows most variable pairs low correlations. *Emotion* and *Uncertainty* exhibit moderate correlation, therefore we should be careful to combine both variables in the same model. Apart from this, we conclude that our choice of sentiment indices does not risk introducing multicollinearity biases to our results.

Second, we control for the attractiveness of the cryptocurrency investment environment by including weekly trading volume (*Log ETH* vol_{t-1}) and price (*Log ETH* $price_{t-1}$) of *Ether*, the second-largest cryptocurrency by market capitalization (coinmarketcap, 2021). We choose to focus on *Ether*, as our dataset is based on *Ethereum*

TABLE 4.6: Correlation table of TRMI sentiment indices

Variable	Sentiment	Uncertainty	Emotion	Innovation	Anonymity
Sentiment	1				
Uncertainty	0.14	1			
Emotion	0.30	(0.45)	1		
Innovation	0.20	0.16	0.04	1	
Anonymity	0.28	0.09	0.17	0.12	1

Note: Above table above shows the Pearson correlation matrix for all 5 TRMI sentiment indices employed in our regression analyses. Correlations are based on value-weighted averages of the top 100 cryptocurrencies and are measured over our observation period of 183 weeks between 2015 and year-end 2018.

blockchain. Therefore, the *Ethereum* currency (*Ether*) is likely to affect investment decisions relatively more than other cryptocurrencies (e.g., *Bitcoin*, at the time of writing the largest cryptocurrency (coinmarketcap, 2021)). Nevertheless, we have tested both *Bitcoin* and *Ether*. The results suggest minimal differences in our regression results. As the market prices of both currencies are highly correlated (*0.9*), we decide to include only *Ether*.

Last, we control for *active campaigns*, which measures the number of ICO campaigns open for fundraising in a given week, to account for the opportunity set of ICOs open to investors at the time of investing. A larger opportunity set will increase the likelihood of investing regardless of environmental cues observed by investors and, as such, needs to be controlled for. *Table 4.7* summarizes all regression variables.

Our enriched dataset enables us to investigate how investors employ environmental cues to form investment decisions in the face of high uncertainty and whether these cues are perceived differently by investor clusters.

4.6.2 Regression methodology

We set-up the regression to take advantage of our data's panel structure, given that we have information on the investor and timing for each transaction. We track the investment decisions for each investor over the entire observation period of *183* weeks between 2015-18 and construct the independent variable to equal to *one* if a given investor has invested in a given week and *zero* otherwise.

We sample *9,000* investors from our dataset of *4.9* mn investors, as our panel dataset would become too large otherwise[6]. To improve the ratio of events (has invested) to non-events (has not invested), we start to observe a given investor *12* weeks prior to their first investment. This choice is based on the assumption that investors will likely only start observing the market shortly before their first

[6]730 mn observations given 183 weeks for each of the 4.9 mn investors.

TABLE 4.7: Summary of regression variables

Variable	Description	Measurement	Source
Dependent variable			
Invest	Specifies whether a given investor has invested in a given week	Binary	ethplorer.io
Independent variables			
Cluster	Categorical variable specifying the cluster affiliation of each investor. Cluster 4 (*the crowd*) is specified as base level.	Dummy	Own analysis
Sentiment$_{t-2} \mid _{t-3}$	Overall positive references, net of negative references, lagged by 2∣3 weeks.	Index [-1,1]	TRMI*
Uncertainty$_{t-2}$	Uncertainty and confusion, lagged by 2 weeks	Index [0,1].	TRMI*
Emotion$_{t-2}$	Emotional sentiments, net of all factual and topical references, lagged by 2 weeks.	Index [-1,1]	TRMI*
Innovation$_{t-2}$	Innovativeness	Index [0,1]	TRMI*
Anonymity$_{t-2}$	User anonymity in transactions	Index [0,1]	TRMI*
Control variables			
Quarter	Specified quarter and year for each week to account for systemic changes in the market environment.	Dummy	Own analysis
Log ETH vol$_{t-1}$	Logarithmized trading volume of Ether cryptocurrency.	Continuous	yahoo finance
Log ETH price$_{t-1}$	Logarithmized market price of Ether cryptocurrency.	Continuous	yahoo finance
Active campaigns	Number of ICOs open for fundraising in a given week. Approximates investor opportunity set.	Integer	Own analysis

*Note: *TRMI stands for Thomson Reuters MarketPysch Index. Accordingly, all concerning variable definitions are based on definition provided by the TRMI. Calculations are based on the top 100 cryptocurrencies (adjusted periodically) by market capitalization (based on coinmarketcap.com). All sentiment indices represent weekly value weighted indices and are based on social media content.*

investment. We follow each investor until the end of our observation period (year-end 2018). This approach is motivated by the existence of secondary markets. Even if a given investor does not actively invest for a time period, they may still decide to sell their tokens on exchanges and are likely to observe market developments to reevaluate their investments. We setup a logit regression as follows:

$$invest_{i,t} = \alpha + \beta_1 S_{t-2} + \sum_{k=1}^{3} \beta_{2,k} C_{i,k} + \sum_{k=1}^{3} \beta_{3,k} S_{t-2} C_{i,k} + \gamma_t + \epsilon$$

Hence, we estimate the likelihood that a given investor (*i*) invests in time period (*t*) given the sentiment (*S*) in time period *t-2*. Also, we account for investor-specific factors by including cluster affiliation *(C)* of each investor. Finally, we add an interaction term between sentiment and cluster affiliation to investigate whether cluster affiliation moderates the effect of sentiment on a given investment decision. In other words, whether investor clusters interpret sentiment differently as a function of their cluster affiliation. γ_t is a vector of time-variant control variables. We control for *Quarter* to account for systemic changes in the crypto market, ICO market activity (*Active campaigns*) and the general cryptocurrency environment $Log\ ETH\ vol_{t-1}$ and $Log\ ETH\ price_{t-1}$ (detailed in the previous section).
Our analysis is based on a panel dataset, following individual investors over time. As error terms are likely not independent over time and per investor series, the assumption of independent standard errors is likely violated. We follow standard practice for dealing with clustered standard errors and calculate robust standard errors as a remedy (e.g., Wick and Ihl, 2019).

4.6.3 Regression results and sensitivity analyses

The regression results are summarized in *table 4.8*. *Model 1-4* show the results of our main analyses with robust standard errors. *Model 1-3* include variables sequentially, enabling us to estimate the effect of including each variable in the model. *Model 4* includes both lagged market sentiment and cluster affiliation as well as interaction effects. *Model 5* varies the lag choice of our sentiment variable. *Model 6* and *7* present alternative estimation approaches to test the sensitivity of results. Specifically, *model 6* does not correct for heteroskedasticity through robust standard errors and *model 7* is a random-effects model to test whether investor fixed effects bias our results.

Our findings indicate that higher positive market sentiment ($Sentiment_{t-2}$) increases the likelihood to invest. This is in line with prior ICO research by Momtaz (2018b) and Lee et al. (2019). Both find that ICO investors employ environmental - in addition to campaign cues - to form investment decisions.

The coefficients of our cluster dummies in *model 4* indicate that the overall likelihood to invest varies with cluster affiliation - which is not surprising, as the number of investments was used as an input variable to our cluster analysis. The dummy coefficients indicate how cluster affiliation changes the expected investment behavior of an investor compared to the *crowd* (i.e., one-time investors).

The coefficients of our interaction terms reveal that sentiment influences investors differently depending on their cluster affiliation. Surprisingly, *Cluster 1* investors seem to be influenced most by sentiment, while *the crowd (Cluster 4)* is least influenced by it. This finding is unexpected and seemingly contradicts findings from similar contexts. For example, Kim and Viswanathan (2018) as well as Wick and Ihl (2019) find evidence that novice investor exhibit stronger herding behavior than other investors. Similarly, Greenwood and Nagel (2009) find that trend-chasing behavior is stronger for inexperienced fund managers.

On the contrary, our findings are in line with Greenwood and Nagel (2009), who find that experienced (professional) investors chase trends on the grounds of rational speculation. They find evidence for coordinated selling efforts by institutional investors. Another possible explanation is that *Cluster 1* investors are market makers. We measure sentiment based on social media (user-generated content).

TABLE 4.8: Regression results: The influence of sentiment on investment decisions of ICO investors

	Dependent variable: Invest (binary)						
						invest	
			coefficient test			*logistic*	*coefficient test*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cluster 1 (c1)		1.715***	1.718***	1.704***	1.714***	1.704***	0.100***
		(0.038)	(0.038)	(0.038)	(0.038)	(0.017)	(0.007)
Cluster 2 (c2)		0.752***	0.756***	0.743***	0.749***	0.743***	0.026***
		(0.031)	(0.031)	(0.031)	(0.031)	(0.019)	(0.003)
Cluster 3 (c3)		0.446***	0.450***	0.442***	0.446***	0.442***	0.013***
		(0.032)	(0.032)	(0.031)	(0.032)	(0.024)	(0.002)
Sentiment$_{t-2}$	2.809***		2.791***	−2.048***		−2.048***	−0.128*
	(0.393)		(0.404)	(0.653)		(0.579)	(0.070)
Sentiment$_{t-3}$					−4.922***		
					(0.637)		
Sentiment$_{t-2}$*c1				9.544***		9.544***	0.865***
				(0.955)		(0.716)	(0.235)
Sentiment$_{t-2}$*c2				6.155***		6.155***	0.258***
				(0.913)		(0.776)	(0.087)
Sentiment$_{t-2}$*c3				4.554***		4.554***	0.146***
				(1.138)		(1.014)	(0.056)
Sentiment$_{t-3}$*c1					7.734***		
					(0.921)		
Sentiment$_{t-3}$*c2					6.074***		
					(0.893)		
Sentiment$_{t-3}$*c3					4.785***		
					(1.083)		
Constant	−8.905***	−10.561***	−10.512***	−10.486***	−10.559***	−10.486	0.018
	(0.585)	(0.750)	(0.746)	(0.686)	(0.685)	(42.864)	(0.013)
Quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log ETH volume$_t - 1$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log ETH price$_t - 1$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Active campaigns	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations						543,684	
Log Likelihood						-85,032.470	
Akaike Inf. Crit.						170,108.900	

Note: Above table summarizes the regression results testing different models for approximating the effect of sentiment on the decision of individual investors to invest in a given week. t-x specifies the time lag of sentiment to a given investment decision. For example t-2 specifies that sentiment was measures 2 weeks prior to the investment decision.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Cluster 1 investors are the most active market participants, therefore, they might coin the market opinion through their social media contributions. We perform *two* additional analyses to develop a more nuanced understanding of individual cluster decision-making.

First, we analyze whether or not *Cluster 1* investors are indeed market makers. If this hypothesis is true, we should be able to observe the time-varying effects of sentiment on different investor groups. That is, we would expect *Cluster 1* to react upon changes in market sentiment first, and other investor groups to follow, i.e., show lagged reactions. We repeat our analysis from *model 4*, this time lagging sentiment by *three* weeks, as compared to *two* weeks previously.

The results are shown in *model 5* of *table 4.8*. Our findings indicate that, while the general effect remains largest for *Cluster 1*, the influence of market sentiment on *Cluster 2* and *3* investment decisions increases or remains stable, while that of *Cluster 1* decreases. Thus, *Cluster 2* and *3* investors indeed have stronger but more lagged reactions than initially indicated by our main analysis (*model 4*). This finding yields support for our hypothesis that *Cluster 1* investors are market makers and other investor clusters follow the market opinion. Alternatively, as active market participants, *Cluster 1* investors might be able to scent market sentiment earlier by following market developments more closely and consequently react upon it earlier.

Second, we investigate how investment decisions are influenced by different sentiment-dimensions by including *four* sentiment sub-dimensions: $Uncertainty_{t-2}$, $Emotion_{t-2}$, $Innovation_{t-2}$ and $Anonymity_{t-2}$. Our regression set-up remains unchanged otherwise. The results are summarized in *table 4.9*. *Model 9* shows the main effects of all sentiment dimensions, while controlling for effects from general sentiment. We find that $Innovation_{t-2}$ and $Anonymity_{t-2}$ both motivate market participants to invest, just like general sentiment. In contrast, $Uncertainty_{t-2}$ and $Emotion_{t-2}$ reduce the number of investments. These findings are in line with

TABLE 4.9: Regression results: The influence of sentiment sub-dimensions on investment decisions of ICO investors

	Dependent variable: Invest (binary)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cluster 1 (c1)		1.689***		1.758***		1.742***		1.721***	1.724***
		(0.018)		(0.020)		(0.018)		(0.018)	(0.021)
Cluster 2 (c2)		0.748***		0.769***		0.772***		0.766***	0.755***
		(0.019)		(0.021)		(0.019)		(0.019)	(0.022)
Cluster 3 (c3)		0.441***		0.395***		0.464***		0.469***	0.388***
		(0.025)		(0.029)		(0.024)		(0.024)	(0.031)
Uncertainty_{t-2}	−16.680***	−27.831***							−26.374***
	(2.152)	(3.187)							(3.978)
Emotion_{t-2}			−9.449***	−10.405***					−5.716***
			(0.738)	(0.898)					(0.958)
Innovation_{t-2}					98.265***	−77.251***			−46.580
					(14.389)	(21.969)			(30.587)
Anonymity_{t-2}							26.815***	52.175***	119.074***
							(8.152)	(11.790)	(12.882)
Uncertainty_{t-2}*c1		23.581***							6.815
		(3.964)							(5.044)
Uncertainty_{t-2}*c2		8.048*							−10.291*
		(4.286)							(5.547)
Uncertainty_{t-2}*c3		9.825*							−9.989
		(5.529)							(7.253)
Emotion_{t-2}*c1				2.919***					−0.899
				(0.783)					(0.939)
Emotion_{t-2}*c2				1.073					−3.675***
				(0.886)					(1.072)
Emotion_{t-2}*c3				−4.002***					−8.241***
				(1.198)					(1.499)
Innovation_{t-2}*c1						300.416***			170.574***
						(25.986)			(37.851)
Innovation_{t-2}*c2						252.790***			270.793***
						(26.657)			(41.555)
Innovation_{t-2}*c3						170.456***			224.105***
						(33.977)			(52.576)
Anonymity_{t-2}*c1								−2.938	−93.845***
								(15.026)	(17.095)
Anonymity_{t-2}*c2								−42.151***	−90.797***
								(15.676)	(18.221)
Anonymity_{t-2}*c3								−96.346***	−127.291***
								(20.992)	(24.428)
Constant	−9.034***	−10.619***	−9.638***	−11.140***	−8.846***	−10.475***	−8.870***	−10.443***	−10.835
	(0.226)	(0.133)	(0.336)	(0.217)	(0.211)	(0.261)	(0.258)	(0.171)	
Quarter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log ETH volume_{t-1}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log ETH price_{t-1}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Active campaigns	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sentiment_{t-2}	No	No	No	No	No	No	No	No	Yes

Note: Above table shows the regression results testing the effect of sentiment-dimensions on ICO investors' investment decisions. All dimensions are measured with a time lag of 2 weeks.
**p<0.1; **p<0.05; ***p<0.01*

our expectations and common sense. Yet, cluster interactions reveal interesting differences between investor groups, allowing us to derive more nuanced cluster profiles based on combining our cluster and regression results.

Cluster 1 can be described as *Crypto Enthusiasts*. Our cluster analysis shows these investors are the most active investor group regarding the number of investments and amount invested. They also invest early and diversely and actively manage their ICO portfolio, selling off *40%* of their initially purchased tokens on exchanges on average. Our regression analysis reveals that they are more likely to invest when the overall market sentiment is positive. This contradicts our initial interpretation stating that *Cluster 1* investors trust their analysis and are likely to be influenced relatively less by market opinions. However, *model 5* in *table 4.8* supports the market maker-hypothesis. These first movers set the overall market tone, which is in line with our initial interpretation. Further supporting this interpretation, *Cluster 1* investors are strongly influenced by $Innovation_{t-2}$. *Crypto Enthusiasts* also show a tendency to be thrill-seeking. They are more likely to invest when $Uncertainty_{t-2}$ is high and are not influenced by $Anonymity_{t-2}$ issues, which may either be interpreted as over-confidence or an attempt to move against the market to buy at low prices and sell at high prices.

In contrast to *Cluster 1*, *Cluster 2* investors are less likely to invest under high $Uncertainty_{t-2}$ and if the market is driven by $Emotion_{t-2}$. This leads us to believe that *Cluster 2* investors are more risk-averse than *Crypto Enthusiasts*. Furthermore, although their investment decisions are influenced by market sentiment, they are less sensitive to it and show a time-lagged response compared to *Cluster 1*. This finding is in line with our initial interpretation stating that *Cluster 2* investors scan the environment to identify ICOs with high intrinsic value. This may lead them to focus on ICO attributes than on environmental cues like market sentiment. Alternatively, *Cluster 2* may be less observant of market developments, potentially due to their smaller investment amount. The timing of their investments, which is roughly average across all investments, leaves us inconclusive about both possible interpretations. *Cluster 2* investors are unlikely to sell tokens in secondary markets, at least during our observation period. Consequently, these investors either consciously or unconsciously follow a buy-and-hold investment strategy, which may lead them to be less sensitive to short-term market developments than an actively trading investor such as *Cluster 1* or *Cluster 3*. In the light of this additional evidence, we refine our initial interpretation to *Passive Investor* (from value investor), as our findings do not provide clear evidence for the investment motives of *Cluster 2*.

Last, we refine *Cluster 3* investors as *Risk-avoiding Rationalists*. They take comparatively long to form investment decisions (they invest relatively late based on our cluster findings and show lagged reactions to sentiment), confirming our cluster results' initial interpretation. They either carefully analyze campaigns before investing or deliberately wait for an established market opinion (in the form of lead investors or market opinion) before investing. This leads us to believe that *Risk-avoiding Rationalists* avoid *uncertainty*, which is also supported by our analysis shown in *table 4.9* and the findings that they avoid more risky markets environments ($Anonymity_{t-2}$). Moreover, they are more likely to invest when market sentiment is positive, supporting our earlier hypothesis that they wait for an established market opinion before investing, potentially to reduce risk. We describe *Cluster 3* as rationalists as they strongly shy away from markets that are dominated by emotional discussions instead of fact-driven debates ($Emotion_{t-2}$). One observation is puzzling: the high sell-off rate of *100%*. Our initial interpretation of this finding was that *Cluster 3* investors purchase tokens to quickly resell them at a profit to other investors, which led us to term them *opportunistic*. The additional evidence from our regression analyses leads us to an alternative interpretation. The rapid sales could indicate high responsiveness to increasing market risk. As *Risk-avoiding Rationalists* feel uncomfortable in high-risk environments, they react quickly to negative market developments by selling off all tokens of a given ICO. Further analysis is needed to validate this hypothesis.

Model 9 combines all sentiment dimensions. While most effects remain unchanged, the effects from *emotions* and *uncertainty* become (partially) insignificant. This is explained by the relatively high correlation (*-0.5*), which causes multicollinearity issues when including both variables. The correlations of all other dimensions are rather low, *0.2* on average, but always below *0.3*.

4.7 Discussion and limitations

4.7.1 Discussion

Our findings suggest that the crowd in ICOs exhibits strong heterogeneity in investment behavior, likely rooted in different investment motives and varying investor backgrounds and expertise. This finding is in line with findings for the crowdfunding market. In particular, Li et al. (2016) and Wallmeroth (2019) find evidence for strong heterogeneity in US and German crowdfunding campaigns.

Our findings are also in line with the findings of Li et al. (2016), who cluster crowdfunding investors. They identify *four* clusters based investment behavior using *seven* variables: percentage of reward-based projects invested, project size, number of investments, comment activity on the platform, campaign popularity [7], diversification of investments [8] and the number of self-initiated projects. Trend backers target popular projects. Altruistic backers go for non-reward based projects, and active backers identify high-quality projects and the crowd, who is yet undecided of the strategy to follow. Their clusters can broadly be matched with our results for the ICOs. Li et al. (2016)'s trend followers match with our *Risk-avoiding Rationalists*, who invest late in projects with selection likely based on external validation from

[7] Expressed as the number of backers per campaign.
[8] Category diversity of invested campaigns.

other investors. Their altruistic investors may be comparable with our *Passive investors*, who select projects they believe in and go in for the long-run. Active backers can be compared to our *Crypto Enthusiast*, who are heavily invested in ICOs, likely investing more resources to identify high-quality projects. Last, Li et al. (2016)'s crowd are our one-time investors. They may either be indecisive on their strategy or invest for reputational reasons - wanting to be part of the innovative crowd and be able to say they took part in it, or for fear of missing out. Consequently, our findings suggest that a similar crowd is attracted by crowdfunding and ICOs. Nevertheless, our analysis also extends Li et al. (2016)'s findings by *two* dimensions: reselling behavior and sensitivity to market sentiment in investment decisions. Consequently, our findings broaden the understanding of investor heterogeneity in crowdfunding environments.

Our findings also provide insights on investor dynamics in the ICOs. An apparent pecking order exists among investors. *Crypto Enthusiasts* and *Passive Investors* tend to invest early, especially *Crypto Enthusiasts* who often seem to be the first to arrive. *Risk avoiding Rationalists* tend to be at the receiving end of the tokens sold by *Crypto Enthusiasts*, usually purchasing tokens either at late stages of fundraising (low *relative position*) or on exchanges (*primary share*). In turn, it is likely that, on average, one-time investors are at the receiving end of tokens sold by *Risk avoiding Rationalist*[9]. Two alternative explanations exist: First, the later entering groups could not invest earlier because they were not aware of the ICO or could not secure a spot during fundraising (where supply is limited and popular campaigns are often oversubscribed). Second, these investors consciously decided to wait, possibly to gain more insights into the market opinion on a given campaign. During fundraising, the market opinion on a campaign is accessible via reading social media conversations or approximated by the amount raised at a given point (often published). In contrast to some of the crowdfunding platforms for which

[9]The relative position of the one-time group is slightly higher than that of *Risk-avoiding Rationalist*.

herding effects were observed (e.g., *Kickstarter*), ICO investors cannot directly observe transactions made by other individuals nor research the background and experience of lead investors. Market sentiment in such a scenario becomes visible in token prices once the token is traded. In this reasoning, our findings align with the findings by Kim and Viswanathan (2018), who find evidence for herding effects in equity crowdfunding campaigns. Specifically, less experienced investors observe and follow expert investors.

In our context, *Crypto Enthusiasts* may be regarded as expert investors, judging by their relative investment experience. In contrast, *Risk-avoiding Rationalists* and one-time investors can be viewed as novice investors. Thus, expert investors in ICOs seem to attract fellowship by novice investors. We find initial evidence for this hypothesis from our regression analysis. Indeed, while *Crypto Enthusiasts* seem to be most reactive to sentiment, other investors show time-lagged reactions to sentiment and seem to follow the dominant market opinion, as set by alleged expert investors. Still, more research is needed to yield further insight into what drives the investment of different investor groups and dynamics. Specifically, further research could look at factors influencing the decision-making of different investor groups.

The ICO market is dominated by inexperienced investors. Our data shows that *70%* of investors have invested a single time during our observation period. This is in line with environments in which bubbles have developed in the past, as researched by Binglin et al. (2017) at the example of the Baosteel Call Warrant bubble. According to Binglin et al. (2017), the dominance of novice investors and substantial investor heterogeneity create a favorable environment for bubbles. We find that experienced investors (our *Crypto Enthusiasts*) do not, as expected, move against the market (i.e., invest when sentiment is low) but instead actively move with the market (invest when sentiment is high). Yet, they do so earlier than inexperienced investors. While this finding seems counter-intuitive initially, it is in line with the finding of Greenwood and Nagel (2009). They research investor behavior during the dot-com bubble and find that experienced investors move with rather than against the market, thereby actively driving up prices and collectively selling to inexperienced investors. While we have not investigated token price developments, this behavior aligns with our ICO market findings. Experienced investors buy in rising (high sentiment) markets and then sell to inexperienced investors, as observed by our timing variables. Our finding may therefore be interpreted as initial evidence for rational herding by some market participants, i.e., the *Crypto Enthusiasts*. This may, in turn, mean that *Cluster 1* is not only dominated by trend-chasing *Crypto Enthusiasts*, but also by strategic investors driven by financial motives.

4.7.2 Limitations

Our research is not without limitations. First, we regard an investor as the sum of all transactions performed by a given blockchain ID. While this approach serves as a good approximation, some investors may possess more than one wallet. In some cases, one investor may have been treated as several investors in the analysis, leading to a bias (as we cannot analyze that investor's full investment portfolio). Furthermore, some IDs may also represent exchanges carrying out transactions on behalf of investors. As explained before, we mitigated this risk by excluding all IDs which have carried out more than *3,500* investment transactions. Still, we cannot rule out that some exchange accounts remain in our dataset. Second, larger deals with professional investors may be carried out privately (not via the blockchain) and are not visible in our dataset. Thus, our dataset is likely to be biased towards non-professional investors. Third, our data allows inferences about the investors' ICO portfolio only. We cannot observe their complete investment portfolio, including other asset classes. As the investigation of some aspects of investment strategy (such as diversification) requires full portfolio visibility, our dataset does not allow a holistic view of investment behavior. Fourth, our analysis focused on *Ethereum*-based ICOs. While we are confident that we capture a critical mass of ICOs during the observation period, excluded ICOs and their investors may systematically differ from those in our sample.

Despite these limitations, we feel that our research has some strengths, primarily due to the novel large-scale investor-based dataset, which allows us to contribute *two* perspectives. It allows us to explore investor heterogeneity in the ICO market, providing a basis for future research. Additionally, by analyzing investors' secondary market behavior, we can generate insights into the effects of secondary market trading on crowdfunding projects.

4.8 Conclusion

Our study is the first of its kind to explore investor heterogeneity in ICOs. We use a novel dataset of *26* mn investments by *4.9* mn investors, employ cluster analysis to detect behavioral investor clusters, and link it with market sentiment data.

We find that investment is highly concentrated, with *70%* of contributors having invested a low amount in a single ICO. Our findings suggest that the remaining *30%* of contributors can be classified into *three* behavioral clusters.

Crypto Enthusiasts are well-informed market participants who closely follow the market and identify high-quality ICOs early. They actively manage their ICO portfolio, diversifying across campaigns, and selectively selling tokens where a suitable opportunity arises. *Passive investors* selectively invest in campaigns they believe in and follow a buy-and-hold investment strategy. They spread risk across campaigns and follow the market opinion but do not closely trace market developments. Last, *Risk-avoiding Rationalists* carefully select a few campaigns but commit a high relative amount to them. They tend to shy away from risk and sell-off their investments if risks arise or to realize high profits.

We further studied how environmental cues influence investors in the form of social media sentiment. We find that positive market sentiment increases the likelihood to invest across investor groups. However, investor groups differ with regards to sensitivity to and timing of sentiment reactions. Experienced investors such as *Crypto-Enthusiasts* are most sensitive to positive market sentiment and react to it quickly. Less experienced investors (such as *Passive Investors*) tend to follow with a significant time lag and are in general less sensitive to sentiment.

Our findings highlight the importance of disentangling the crowd in crowdfunding by exemplifying the crowds' heterogeneity in the example of ICOs. Our research contributes to the existing literature by making the first attempt to capture and define this heterogeneity, developing a foundation for future research to advance our current understanding of investor heterogeneity.

Future research is needed to understand the drivers of investment decisions of heterogeneous investor groups. Advancements are also needed to understand the implications for new ventures. That is, does the attraction of certain investor groups contribute more to venture success? This could be particularly important for ICO projects, many of which are platform businesses where an active community of early-adopters is critical. If such a venture was to attract mainly financially-motivated or even speculation-driven investors, funding success might become negatively correlated with venture success.

From a practical perspective, implications arise for ventures and regulators. The market activity of ICOs has drastically declined in 2019 after an increasing number of fraud scandals became public and more regulatory authorities took actions against ICOs (Crowdfundinsider, 2019). Nonetheless, blockchain technology's potential for revolutionizing the funding process of early ventures has yet to be determined. To further advance this blockchain use case, it is critical for regulators and ventures alike to understand their audiences to develop an ICO market trustworthy and supportive of venture success.

Chapter 5

Sheep or Shepherd? Rational herding in ICOs and the role of experience

5.1 Abstract

Can blockchain-based start-up financing work? We tackle this question from an investor-angle by asking if ICO crowds effectively channel capital to prospering ventures. We investigate whether ICO investors are vulnerable to herding, irrational behavior in which investors blindly follow others, thereby ignoring their own judgment. We ask whether herding in ICOs may also be rational, meaning investors employing social cues to extend, not replace, their analysis. This may be particularly valuable in ICOs, where information asymmetry exists, and observing others may cascade private information. We show that ICO crowds herd, attributable primarily to irrational behavior. Nevertheless, our analyses point to significant value from disentangling the crowd. Novice investors irrationally follow the general market opinion, while experienced investors strategically employ social cues to aid their analysis and hence herd rationally. We extend the existing theory by showing that ICO herding may be both rational and irrational; and by pointing to the moderating role of experience.

5.2 Introduction

Can blockchain-based venture financing (ICOs[1]) work? A growing body of research attempts to answer this question. In our opinion, the answer will depend on a fundamental question: are crowds capable of channeling capital to prospering ventures. Put differently are ICOs the *wisdom*- or *madness* of the crowd at work? There has been considerable debate over this topic and *two* schools of thought have emerged. One school of thought believes that *"crowds bring up tulip crazes, subprime meltdowns, the Kitty Genovese scandal, Salem witch trials, and other tragedies."* (Isenberg, 2012, para 11). Supporters of this view argue that crowds are susceptible to numerous behavioral biases and are easily manipulated. One widely documented bias is herding, in which individuals blindly follow the general group opinion, thereby ignoring their own judgment (Mollick and Nanda, 2016).

Crowdfunding environments, particularly ICOs, may be particularly vulnerable to herding for *three* reasons. First, crowdfunding is characterized by outcome uncertainty, as early-ventures are inherently difficult to value. Indeed, according to Prosser (2016), about *90%* of start-ups fail. ICOs, which depend on the value of blockchain technology, arguably possess even higher levels of uncertainty. Second, information asymmetry exists in the ICOs, as acquiring reliable information on venture quality is prohibitively costly for retail investors with insignificant individual contributions. Third, other investors' behavior is observable, making social influence possible (Yum et al., 2012).

Consequently, Isenberg (2012) argues that placing a crowd in an early-venture financing context can only end in tragedy. Indeed, evidence from comparative contexts supports this hypothesis. Simonsohn and Ariely (2008) study buyers' behavior on *eBay* and find that they rush into auctions with a high number of previous bids, thereby ignoring salient factors, such as low starting prices, as

[1]The term Initial Coin Offering describes the process of raising capital through the blockchain in the form of cryptocurrency tokens. Tokens are comparable to stock market shares. Nevertheless, they do not necessarily represent ownership in a firm.

the reason for the higher number of bids. Binglin et al. (2017) show that crowd dynamics played a crucial role in the development of the dot-com bubble.

In contrast, the second school of thought believes that crowds form higher quality decisions by bringing together diverse perspectives, capabilities and collectively reducing outcome biases. Evidence from the crowd-sourced encyclopedia *Wikipedia* supports this hypothesis (Greenstein and Zhu, 2014). Herding may result from rational observational learning instead of irrational mimicry, at least in some contexts. Herding may help reduce information asymmetry in contexts where market participants possess varying information. For herding to be rational, investors must not blindly follow others but instead use the inferences drawn from observing others to aid their analysis. Zhang and Liu (2012) find evidence for such rational herding in US microloan markets.

A growing body of research attempts to determine which circumstances determine whether crowd decision-making is superior or inferior to that of experts. Yum et al. (2012) argue that the nature of the underlying problem determines the effectiveness of the decision-making body: *"(...) collective intelligence from crowds is superior in handling probabilistic problems"* (Yum et al., 2012, p.475). Following this line of argumentation, ICOs represent the ideal context for superior crowd decision-making. It is inherently difficult to make such inferences in the context of ICOs, as it may take years before it is possible to evaluate the quality of the decision outcome (i.e., venture success). As such, most research has shifted to studying the existence of herding effects. The majority of herding research finds evidence for pronounced herding in crowdfunding contexts (Wick and Ihl, 2019; Kuppuswamy and Bayus, 2018; Ajaz and Kumar, 2018; Gurdgiev and O'Loughlin, 2020). Research on herding behavior in ICOs is scarce and the question of whether herding may be rational remains entirely unanswered to this date.

Our research aims to fill this gap by exploring herding effects in ICOs. We also attempt to shed light on the underlying processes, investigating whether potential herding is rational or irrational.

Our research builds on a unique panel dataset, tracking the funding progress and exchange-trading activity of *1,274* ICOs between 2015 and year-end 2018. Our dataset is a novel re-combination of *17* mn blockchain transactions by *3.6* million ICO investors; ICO campaign and venture attributes from *nine* ICO platforms; and daily cryptocurrency sentiment, as provided by TRMI[2]. We find evidence for pronounced herding in ICOs, which seems to be primarily irrational. Our findings suggest that experience moderates herding. While all investors herd, the underlying behavioral processes depend on the experience of individual investors. Novice investors herd irrationally, imitating the behavior of other market participants. This effect disappears with increasing experience. Instead, experienced investors strategically employ the cues collected from observing others to adjust their behavior accordingly.

Our findings contribute to the existing literature in *two* main ways. First, our study is the first large-scale panel research to provide evidence for herding in ICOs. Second, we extend the current understanding of crowd dynamics by shedding light on the moderating role of investor experience in determining whether herding behavior is (ir)rational.

The remainder of this paper is structured as follows. The first section discusses the theoretical background of herding and experiential learning in settings comparable to ICOs. The second section elaborates on the dataset and methodology employed in our analysis. The third section discusses our results, starting with main herding effects and then elaborating on the moderating role of experience and campaign phase. We end this paper with a discussion to place our findings in a broader context and point to limitations and future research directions.

[2]Thomson Reuters MarketPsych Indices.

5.3 Theoretical background

5.3.1 Herding

The democratization of venture financing is at the heart of ICOs as an innovation and shifts decision-making power from a few (*the experts*) to many (*potentially everybody*). This shift may enlarge the overall capital pool and potentially lead to a superior outcome by channeling capital to higher-quality start-ups.

Research on venture capitalist decision-making shows that even experts cannot escape natural human biases. Malmström et al. (2020) show that female founders, especially confident ones, are less likely to strike a deal settle the deal. Franke et al. (2006) find evidence for homophily biases and Cumming and Dai (2010) provide evidence for local biases in venture firms. Hence, leaving funding decisions to a few (*the experts*) may result in promising ventures being precluded from the market (Mollick, 2014).

It is argued that crowds, in some settings, may arrive at superior decisions, driven by greater diversity leveraged during decision-making - the so-called *wisdom of the crowds* (Mollick, 2014). For example, a study by Greenstein and Zhu (2014) shows that, even though *Wikipedia* articles are more biased when first published, they end up less biased than expert content due to the larger number of edits by the crowd. Crowds as well are susceptible to behavioral biases that may lead to suboptimal decisions: e.g., group conformity biases such as groupthink (Janis, 1991) or overreliance on experienced group members - a phenomenon more generally referred to as *herding* (Wilson and Schooler, 1991). Herding has been widely documented in group-based investment settings, such as stock markets (e.g., Neupane et al., 2014; Binglin et al., 2017; Greenwood and Nagel, 2009), crowdfunding (e.g., Wick and Ihl, 2019; Vismara, 2018), or microloan lending (e.g., Zhang and Liu, 2012; Lin et al., 2015; Yum et al., 2012).

We define herding as sequential decision-making: individual investors observe others' behavior to incorporate the observation in their decision-making. In its conventional definition herding is seen as an irrational bias in which individuals blindly mimic others' behavior, neglecting their judgment (Zhang and Liu, 2012; Ajaz and Kumar, 2018; Chen and Lin, 2014). Advocates of this perspective on herding argue that group-based investment decisions are doomed to fail as the *madness of the crowds* will inevitably lead to inferior decisions, as demonstrated by multiple historical disasters caused by crowd malfunctioning, e.g., medieval witch hunts (Mackay, 2008). Several studies find evidence for irrational behavior in group-based (investment) settings. Simonsohn and Ariely (2008) find that bidders on *eBay* tend to rush into campaigns with many bids, ignoring that the number of bids results from a low starting price. Avery and Zemsky (1998) assert the role of irrational herding in the development of stock market bubbles. Lin et al. (2013) and Herzenstein et al. (2011) find evidence for irrational herding in the Chinese microloan market by showing that investors follow the crowd. Yet, the crowd does not collectively pick higher-quality loans, i.e., resulting in lower default rates. Another strand of literature argues that herding may also be rational (Zhang and Liu, 2012; Yum et al., 2012). In particular, contexts in which some market participants possess private information create situations in which observing others' decisions provides additional informational value. Herding may become strategic and better described as *observational learning*. Zhang and Liu (2012) find evidence for rational herding in US microloan-lending by showing that investors' interpretation of loan attributes depends on the observed herding momentum. In light of positive herding momentum for a particular loan, investors will value loan attributes that they initially interpreted adversely. Hence, investors employ social cues to amend their analysis yet do not ignore loan attributes altogether, as hypothesized under the irrational herding hypothesis. This understanding of herding as cascading of private information is rooted in studies by Welch (1992) and Bikhchandani et al. (1992).

5.3.1.1 Herding in Crowdfunding

While herding has been studied in many group-based settings, crowdfunding is of particular interest due to its close resemblance to ICOs. *Three* characteristics make crowdfunding susceptible to herding. First, the uncertainty surrounding early-venture financing is particularly high given the absence of any track records. Second, information asymmetry exists between ventures and investors, as ventures have the incentive to window-dress in order to maximize their valuation. The cost of performing due diligence to resolve such asymmetries is prohibitively high for retail investors. Isenberg (2012) states that a particular VC firm regularly spends *50,000* US Dollar to perform diligent analysis of a single ventures' prospects. Performing due diligence comes at different costs for investors depending on their background. Such significant expenditures could only be justified by large investment amounts. Third, the behavior of investors is observable, enabling social influence among them. Crowdfunding platforms often display information on the cumulative amount funded and individual investments. It is in line with the expectation that an increasing number of studies find evidence for herding in crowdfunding environments (e.g., Vismara, 2018; Wick and Ihl, 2019; Kim and Viswanathan, 2018; Kuppuswamy and Bayus, 2015; Li et al., 2016).
Crowdfunding research has progressed to study the processes underlying herding. For example, Li et al. (2016), and Kim and Viswanathan (2018) show that investors single out and follow experts in the crowd. Research points to the value of disentangling the crowd to investigate how individual investors' herding behavior differs. For example, Wick and Ihl (2019) show that herding differs, depending on the individual investors' background. They show that experts actively avoid campaigns that attract a large number of novice investors. Hence, crowdfunding research suggests that herding may be interpreted as a mechanism for cascading information in the face of high information asymmetry.

5.3.1.2 Herding in ICOs

The investment setting in ICOs is comparable to that of crowdfunding. In consequence, we expect that ICOs too are susceptible to herding. Indeed, the buzz around blockchain technology may attract many inexperienced investors and ventures with variable quality to the largely unregulated market. In consequence, the challenge to identify high-quality ventures is elevated in ICOs. We expect herding to be even more pronounced in this environment.

While research in this field is just evolving, several studies have investigated herding in the context of ICOs (Kallinterakis and Wang, 2019; Gurdgiev and O'Loughlin, 2020; Lee et al., 2019). Kallinterakis and Wang (2019) as well as Gurdgiev and O'Loughlin (2020) analyze crypto-exchange data and find evidence for sentiment-induced herding in secondary markets for ICO tokens. Lee et al. (2019) study campaign-level data and provide initial evidence for the existence of herding in ICO campaigns. They show how positive analyst opinions trigger large first-day subscriptions, which in turn generate higher subsequent subscriptions. Their findings provide evidence that in ICOs as well, investors aim to identify credible social signals to aid their own decisions.

While herding has been documented in the context of ICOs, to our knowledge, no attempt has been made to disentangle the crowd to study how herding mechanism may differ in this highly heterogeneous crowd depending on investor background. Moreover, herding in ICOs has been generally classified as being irrational, spurred by buzz. Within this work, we attempt to tackle both research gaps. While crowd-based investment environment research provides a basis for heterogeneity of herding processes among investors, the conditions under which herding may be regarded as rational are much less understood to date. While several factors may play a role, evidence leads us to turn to experiential learning as a critical behavioral determinant of investor irrationality.

5.3.2 The role of experience

Two schools of thought, advocating the rational and irrational herding hypothesis, seemingly contradict each other. Indeed, evidence for both has been provided for the very same context, namely microloan-lending (Zhang and Liu, 2012; Herzenstein et al., 2011; Chen and Lin, 2014). While many factors may play a role in shaping the existence and magnitude of herding momentum (for example, cultural context (Chen and Lin, 2014)), research increasingly points to experience as a moderating factor (Griffin et al., 2011; Greenwood and Nagel, 2009; Boreiko and Risteski, 2020; Binglin et al., 2017).

Several studies of stock market bubbles relate the development of bubbles to investor experience. For example, Greenwood and Nagel (2009) find that young mutual fund managers are more likely to herd after trends than more experienced investors. Likewise, Binglin et al. (2017) show that bubbles are characterized by a shift in investor crowd towards inexperienced investors. The continuous inflow of novice investors' prevents overall market learning, leading to potentially dangerous market dynamics, i.e., irrational upwards price spirals with subsequent price-correcting market crashes.

Boreiko and Risteski (2020) are the first to study the role of investor experience in ICOs. They analyze the investment behavior of *400,000* investors and investigate whether reimbursement learning - the improvement of strategies through experience - is observable in ICOs. Their findings suggest that ICO investors do not learn through experience. However, investment size moderates this effect. Expert investors tend to learn through experience, while retail investors do not. Within this study, we aim to detail the role of experience as a moderator for herding behavior in ICOs. We investigate whether experiential learning moderates irrational herding in particular. We employ, in contrast to Boreiko and Risteski (2020), a large scale panel dataset covering both the fundraising and secondary market stage of ICOs.

5.4 Data and methodology

5.4.1 Data sample

Our analyses require an accurate specification of the transaction day. We exclude *354* ICOs with potentially delayed transaction visibility. The reason for such delay is simple, while most ICOs process investments directly on the blockchain, some collect registrations during fundraising and issue all tokens at once after the campaign. In such cases, while the order of investments is preserved, transaction timestamps do not accurately reflect investment timing. We identify potentially time-distorted campaigns by plotting the cumulative investments against time and excluding campaigns with a high concentration of bids share in a short time interval (a few hours). A detailed description of this cleaning step can be found in *section 3.1.5*. Our final input dataset consists of *17* million transactions by *3.6* million investors in *1,274* ICOs.

Based on this dataset, we calculate the daily progress of each ICO campaign, in terms of the number of investments and funding amount, during the fundraising phase (the ICO, also referred to in this paper as *primary phase*) as well as trading on crypto-exchanges (*secondary phase*). We enrich our data by linking it with *two* additional datasets, time-dependent social media sentiment and time-invariant ICO attributes. We will briefly elaborate on each dataset in the following. For a detailed description on the construction of each dataset please refer to *section 3.2*. A detailed description of the variables will be presented in the next section.

First, we add a set of *seven* ICO attributes that will allow us to measure how investors employ campaign characteristics to form investment decisions. As described in *section 3.2.1*, the choice of our *three* campaign attributes is based on prior ICO research. The *four* venture attributes are selected on the basis of venture capitalist decision-making criteria, as conceptualized by Baum and Silverman (2004). Baum and Silverman (2004) argues that venture capitalists evaluate *3* types of capital possessed by a venture: human, intellectual and alliance capital, which we approximate by *GitHub starts* (intellectual capital) and *CEO experience* (human and alliance capital). Additionally, we approximate risk by including *Blockchain industry* and *Project status*. Second, we measure daily social media sentiment at the time of decision-formation, as detailed in *section 3.2.2*.

Due to the low data availability of campaign attributes, our dataset is reduced to *58,250* observations, based on the investments of *1.8* million investors in *213* ICOs between 2017-18. We include a total of *7.1* million transactions, thereof *75%* exchange-trading and the remaining *25%* during fundraising.

We also aim to study the role of investor experience and create *two* additional data-subsets, one for novice and experienced investors each. While both sets represent a sub-sample of the full set, the reverse is not true. Consequently, they do not add up to the full dataset. Rather we categorize experience dynamically at the moment of a given investment decision. This, in turn, means the same investor can belong to both subgroups, transitioning from a novice to an experienced investor. With this approach, we aim to capture the effects of learning. The novice investor dataset is based on a total of *1.2* million investments by the same number of investors. The experienced dataset is based on *3.4* million transactions by *320,000* investors.

5.4.2 Variables

5.4.2.1 Dependent variables

We employ *three* dependent variables in our analyses. Our main analysis employs *Log amount raised*$_t$ as the dependent variable, measuring the amount raised by ICO on day (t). This is matched with the cumulative amount raised as an independent variable, detailed in the next section, to measure herding. Following the arguments of Zhang and Liu (2012), we employ *Log amount raised*$_t$ rather than the number of investments, as the former reflects both the decision to invest and the outcome of the investor-specific capital allocation decision between competing ICOs. The

amount raised is often displayed on ICO platforms or venture websites, while the number of investments is usually much less visible. Investors would have to analyze blockchain transactions in order to obtain such information. It follows that the *Log amount raised*$_t$ is expected to be the primary source of social influence among investors. We log the amount raised to correct for the non-normality observed in the variables' distribution.
The second analysis disentangles investors, investigating whether herding differs between novice and experienced investors. We define novice investors as all one-time investors and the first investment of all serial investors. Experienced investors have made at least *nine* investments during the time under investigation, representing the *95*th quantile of investments per individual in our sample. Importantly, this classification is independent of campaigns. We only count the first investment per ICO and disregard follow-up investments in the same campaign. We employ *Log novice investments*$_t$ and *Log experienced investments*$_t$ as dependent variables in this analysis, measuring the number of investments received from novice and experienced investors on a given day by an ICO. The decision to switch from the amount to the number of investments is based on systematically different contribution patterns of novice versus experienced investors, making the latter a more meaningful measure for this analysis.

5.4.2.2 Independent variables

We employ *nine* independent variables, *two* herding variables and *seven* ICO attributes, in our analyses. Each are detailed in the following and summarized in *table 6.1*. *Log total amount raised*$_{t-1}$ is our primary herding variable. It is defined as the cumulative amount raised by a given ICO until the previous day of a given investment decision (dependent variable). As described, this metric is visible for most campaigns on ICO platforms and venture websites. Accordingly, we argue that it is the primary source of social influence among investors.

TABLE 5.1: Summary of regression variables

Variable	Description	Measurement
Dependent variables		
Log amount raised$_t$	Amount raised per campaign on day *t*	Continuous
Log novice investments$_t$	Number of investments by novice investors, first time investors, per campaign on day *t*	Continuous
Log experienced investments$_t$	Number of investments by experienced investors, at least 9 prior investments up to day *t*, per campaign on day *t*	Continuous
Independent variables		
Log total amount raised$_{t-1}$	Cumulative amount raised per campaign on day *t-1*.	Continuous
Sentiment$_{t-2}$	Overall positive references, net of negative references of social media content of the top 100 cryptocurrencies weighted by market capitalization on day *t-2*.	Index [-1,1]
Pre ICO	Specifies whether an ICO has carried out a preliminary token sale before the actual ICO.	Dummy
Token price	Price of ICO tokens during fundraising in US Dollar.	Continuous
Accept fiat	Tokens can be purchased directly in exchange for fiat currency (legal tender) during the ICO.	Dummy
Blockchain industry	Specifies whether the ventures core business is related to blockchain technology.	Dummy
Project status	Founding stage of the venture: Just an idea (base level), MVP*, Beta version or Working product.	Categorical
CEO experience	Years of CEO experience	Continuous
GitHub stars	Measure of popularity of a project on GitHub, a widely known and used open-source coding repository (Github, 2020).	Integer
Control variables		
Campaign	ICO campaign to account for systematic differences between campaign fundraising processes	Categorical
Day-of-week	Weekday on day *t* to account for weekday seasonality in investments	Dummy
Day-of-listing	Age of ICO campaign on day (t), measured in days from ICO-start date	Dummy
Quarter	Year and year-quarter on day *t* to account for systematic changes in the market environment	Dummy
Duration	Length of ICO campaign in days	Integer
Phase	Primary (ICO, base level) versus secondary market trading (crypto-exchanges)	Dummy

**Minimum viable product*

We also measure sentiment-induced herding by $Sentiment_{t-2}$, defined as social media sentiment *two* days prior the investment under investigation. The variable is constructed as a value-weighted index of the top *100* cryptocurrencies. Please refer to *section 3.2.2* detailing the dataset and index calculation. $Sentiment_{t-2}$ is, in contrast to *Log total amount raised*$_{t-1}$, not campaign-specific. Rather, it measures market-wide sentiment, hence, hype-induced irrational herding, as it does not reveal any additional campaign-specific information.

The second set of independent variables characterizes the investment target. We include *seven* variables, *three* campaign and *four* venture attributes. The campaign attributes are *Pre ICO*, *Token price* and *Accept fiat*. *Pre ICO* specifies whether a venture conducted a prior token sale before the actual ICO. Pre-sales are used for demand-testing and fine-tuning the ICO design. They are expected to be interpreted as a positive signal by investors. *Token price* specifies the US Dollar price one token was sold for during the ICO. If the *Token Price* varied during the ICO, we calculate the average price throughout fundraising. Higher token prices may serve as a signal for quality[3]. We specifically want to investigate whether investors systematically choose lower-priced ICOs. As lower token prices might be driven by (large) discounts, often given out by ICOs during fundraising, they might indicate lower investor commitment, as investors might be attracted by the potential arbitrage opportunity rather than intrinsic token value. Last, we include *Accept fiat*, specifying whether legal tender (fiat) was an accepted medium of payment during fundraising. Accepting fiat is generally interpreted as a negative signal by investors, making it easier for potential scammers to cash-out.

We also include *four* venture attributes, which characterize the start-up behind the ICO rather than the campaign set-up. First, *Blockchain industry* is a dummy variable taking the value of *one* if a ventures core business is related to blockchain

[3]While investors may interpret a higher token price as a quality signal, technically, it is not. The valuation of a venture depends on both the number of tokens and the token price. An ICO may choose to issue more tokens at a lower price to issue the same amount of capital.

technology. It serves as a proxy for technology focus and risk, as blockchain technology's market value is still uncertain. Second, we include *Project status*, which specifies the development phase of the start-up and is classified along *four* stages: *Just an idea*, *MVP*[4], *Beta version* or *Working product*. In the absence of a historical track record, the availability of test versions of the product may serve as a credible quality signal. This variable also approximates investors' risk appetite, as ventures in earlier stages have a higher risk attached to them.
Last, we include *CEO experience* and *GitHub stars*, serving as proxies for human/ alliance and intellectual capital, respectively, as detailed in *section 3.2.1*. *CEO experience* measures the years of experience as CEO. It captures the effects of both human and alliance capital. More experienced CEOs have a *proven* track record, hence possess human capital. Experienced executives are more likely to have established network ties. *GitHub stars* measures the number of stars on the coding repository *GitHub* and approximates the popularity of a ventures code among fellow developers. We argue that codes with a higher number of *GitHub stars* indicate higher intellectual capital, as it expresses project excitement and visibility within the developer community.

5.4.2.3 Control variables

We include several control variables, enabling us to isolate herding effects. First, we include *Campaign* fixed effects to control for systematic differences among the *213* ICOs in our sample. We also include *Day-of-week* and *Quarter* fixed effects to account for time-dependent trends. For example, the investment activity may be higher on Sundays, as investors use weekends to restructure their investment portfolio. ICO activity has also increased rapidly during our observation period. We need to control for systematic differences in periods, as approximated by the *Quarter*-control, which also includes the specific year.

[4]Minimum viable product.

Next, we control for *Day-of-listing*, measuring the campaign age in days from the ICO-start date to account for lifecycle specific effects. We also control for campaign *Duration*. Some campaigns have raised funds within *two* days, while others have spread their fundraising over an entire year. This campaign design choice affects investor dynamics, and we need to control for its effect. Last, we account for the *Phase*, specifying whether an ICO is actively raising funds or trading on exchanges on a given day.

5.4.3 Empirical approach

Our empirical strategy is based on the work of Zhang and Liu (2012) and can be described in *three* steps. First, we isolate herding momentum by analyzing how an individuals' investment decision on a given day (*Log amount raised*$_t$) is influenced by the collective investment decisions of all previous investors (*Log total amount raised*$_{t-1}$). Next, we isolate the effects of time-invariant ICO attributes on the amount raised per day. We have selected *seven* ICO attributes, as described in the previous section. This approach allows us to measure which ICO attributes investors consider in their decision-making and how they are interpreted.

In the third step, we investigate the behavioral processes underlying (potential) herding. In particular, we distinguish between rational and irrational herding by introducing interaction terms between the primary herding variable (*Log total amount raised*$_{t-1}$) and each ICO attribute. Irrational herding occurs when investors blindly mimic the behavior of others, ignoring their analysis of ICO attributes. Positive herding momentum increases the investment likelihood of investors without altering the interpretation of ICO attributes. Hence, the interaction terms are expected to be insignificant under the irrational herding assumption.

In contrast, under the rational herding assumption, investors employ collective decisions of others as an additional decision cue and interpret campaign attributes in the light of these additional information. Consequently, rational herding should alter their interpretation of campaign attributes in the presence of positive herding momentum and be reflected in significant interactions terms with signs *opposing* the base effects. Our regression equation is set up as follows:

$$log(y_{i,t}) = \alpha + \beta_1 log(Y_{i,t-1}) + \beta_2 \boldsymbol{C_i} + \beta_3 Y_{i,t-1}\boldsymbol{C_i} + \beta_4 \boldsymbol{X_{i,t}} + \epsilon i$$

where $log(y_{i,t})$ is the *Log amount raised*$_t$ by campaign i on day t, $log(Y_{i,t-1})$ is the *Log total amount raised*$_{t-1}$ by campaign i up to day t, $\boldsymbol{C_i}$ is a vector of time-invariant attributes of the ICO project i and $\boldsymbol{X_{i,t}}$ is a vector of control variables, as described in the previous section.

We extend Zhang and Liu (2012)'s approach by another herding variable, particularly relevant to ICOs. *Sentiment*$_{t-2}$ measures general crypto market attitude and is a measure of hype-induced irrational herding, as it does not reveal any additional campaign-specific information and should not alter the interpretation of ICO attributes. In particular, we are interested to explore the effect of sentiment on the herding momentum and its dynamics with campaign-specific herding (i.e., complement or substitute). The regression equation becomes:

$$log(y_{i,t}) = \alpha + \beta_1 log(Y_{i,t-1}) + \beta_2 S_{t-2} + \beta_3 \boldsymbol{C_i} + \beta_4 Y_{i,t-1}\boldsymbol{C_i} + \beta_5 \boldsymbol{X_{i,t}} + \epsilon_i$$

where S_{t-2} captures sentiment-induced herding. While sentiment should not alter the interpretation of ICO attributes (hence we do not include interaction terms), we test for this in the sensitivity analyses.

5.5 Results

5.5.1 Descriptive results

The descriptive results for the *213* ICOs included in our regression are shown in *table 5.2*. *Fundraising statistics (Primary)* and *Exchange-trading (Secondary)* summarize the campaign level results during the ICO and trading on crypto-exchanges, respectively. The average ICO in our sample raised *2.2* million US Dollar. The large discrepancy to the medium value of ∼*650* thousand US Dollar indicates high heterogeneity of the amounts raised by ICOs. Indeed, the largest ICO in the sample has raised *80* million US Dollar. We report the amount raised for the fundraising phase only, as the transaction values in the secondary phase are not meaningful since they represent values from circulating existing tokens rather than representing funding of the venture.

The average ICO has attracted 5,169 investors during fundraising and *10,696* investors on exchanges. In line with the funding amount, large discrepancies between the mean and median values indicate high heterogeneity between ICOs. This implies that most market participants purchase tokens on exchanges rather than in the ICO. This observation may be attributable to either of the *three* factors. First, the supply of tokens during fundraising is restricted. Excess demand during that phase may force investors to purchase tokens on exchanges[5]. Second, investors may only become aware of the ICO at later stages, possibly related to increasing buzz around cryptocurrencies. Third, investors may deliberately decide to wait until tokens trade on exchanges to collect additional information on the intrinsic value of tokens, expressed in the market price. ICOs attract, on average, more novice than experienced investors, but experienced investors contribute more capital than novices. Both findings are in line with comparable

[5]Once ICO (fundraising) investors choose to sell tokens.

TABLE 5.2: Descriptive results of ICO campaigns and market sentiment

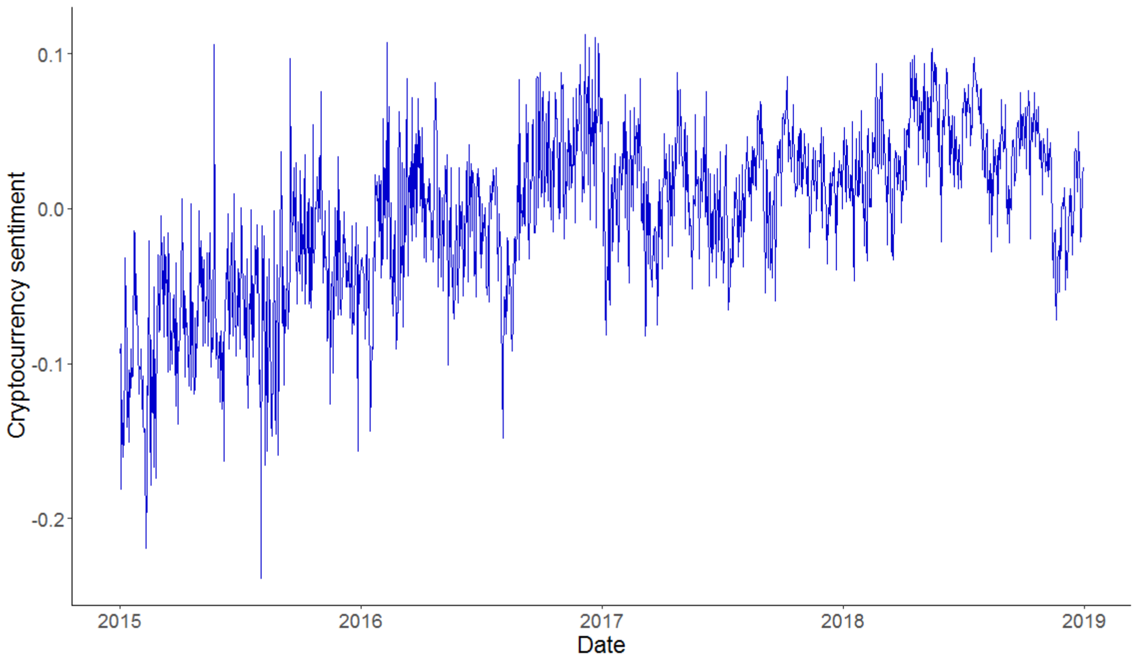
Variable	Mean	Median	STD	Min	Max
Fundraising statistics (Primary)					
Total amount raised (US Dollar)	2,243,469	650,956	6,832,786	301	80,114,822
Novice amount	694,769	146,518	2,057,109	38	18,035,600
Experienced amount	795,342	155,005	4,126,946	181	55,743,584
Number of investors	5,169	1,866	11,389	13	92,943
Novice investors	1,849	430	4,706	2	34,288
Experienced investors	1,645	333	4,267	2	33,724
Exchange-trading (Secondary)					
Number of investors	10,696	5,541	19,053	1	229,094
Novice investors	4,103	1,516	6,483	1	49,777
Experienced investors	3,032	2,043	3,755	1	36,857
ICO attributes					
Pre ICO	0.35	0	0.48	0	1
Token price	1.32	0.14	5.43	0	52.46
Accept fiat	0.10	0	0.30	0	1
Blockchain industry	0.16	0	0.37	0	1
Project stage: MVP	0.18	0	0.39	0	1
Project stage: Beta version	0.24	0	0.43	0	1
Project stage: Working product	0.51	1	0.50	0	1
CEO experience	4.65	3	5.30	0	40
GitHub stars	78	10	205	0	1,275
Duration	37	31	26	1	190
Sentiment	-0.00	0.00	0.05	-0.24	0.11

Note: Above table shows the summary statistics of ICO attributes and fundraising statistics for the 213 ICOs analysed in our regression. Novice investors are defined as all first-time investors, experienced investors have invested at least nine times at the time of investment under investigation.

contexts, e.g., crowdfunding (Wallmeroth, 2019), and for this reason in line with our expectation. The contributions of experienced investors seem to be more varied than those of novice investors, as indicated by the higher standard deviation. Depending on the ICO, either a different group of experienced investors is attracted, or their contribution pattern systematically differs based on ICO attributes. This effect will be further explored in the next section.

Last, we analyse the distribution of cryptocurrency social media sentiment during our observation period, as shown in the last line of *table 5.2* and *figure 5.1*. We switch the unit of analysis to investigate sentiment, i.e., no longer analyzing campaign-level data. The average sentiment over our observation period is slightly negative but almost zero. This indicates that periods of positive and negative sentiment are approximately balanced. As observable in *figure 5.1*, social media sentiment is volatile with positive and negative episodes and punctual spikes. We observe a trend as the general level of sentiment steadily increases during the observed period. This supports our model specification, as described in *section 5.4.3*, to account for time-dependent trends. The lower half of *table 5.2* provides initial insights on ICO attributes, a few observations stand out. The share of ventures with a blockchain industry focus is rather low with *16%*. Therefore, most ventures attempting to raise capital via the blockchain, at least in our dataset, use it *solely* for the purpose of fundraising. Also, the status of the average project is surprisingly advanced, with *51%* of ventures claiming to have a functional product in place and 75% to have at least a *Beta version* available. Moreover, the

FIGURE 5.1: Daily cryptocurrency social media sentiment



Note: Above figure shows the social media sentiment over our observation period, as kindly provided by Thomson Reuters in the form of their TRMI cryptocurrency index. Sentiment is calculated as the weighted-average sentiment of the 100 largest cryptocurrencies by market capitalization (obtained from coinmarketcap.com).

average CEO has significant prior management experience with a mean value of *4.65* years.

Other findings are more intuitive to understand. For example, *35%* of ICOs have conducted a *Pre ICO*, meaning they sold a restricted number of tokens before the actual token sale. Such preliminary token sales have become common practice to test demand and fine-tune the actual ICO design. *Table 5.2* also indicates that *Token price* during the fundraising phase is quite variable. While the average price was *1.3* US Dollar per token, most ventures sold their token for much lower prices, with a median of only *14* cents. Without considering the number of tokens issued, we cannot draw any inferences about the underlying company valuation. It does, nevertheless, reveal different pricing strategies by ICOs. Only *10%* of campaigns *Accept fiat* currencies, legal tender issued by governments. Thus, the vast majority of tokens were purchasable solely against cryptocurrencies. The average venture in our sample had *78* stars of *GitHub*, which is equivalent to followers, most of whom are programmers themselves. The number of *GitHub stars* may serve as a proxy for code-quality, project excitement, and visibility in the coding network. Last, the average ICO ended after *37* days.

Table 5.3 shows the correlation of all independent variables included in the regression. No variable pair exhibits a high correlation. Based on this, we are confident that our independent variable composition does not cause multicollinearity issues.

5.5.2 Main results

The main results of our regression analyses are shown in *table 5.4*. *Column 1* shows the main herding results. The coefficients of both herding variables, *Log total amount raised*$_{t-1}$ and *Sentiment*$_{t-2}$ are positive and significant, with *Sentiment*$_{t-2}$ dominating. Therefore, our results indicate that a positive herding momentum is triggered by both a higher amount of prior funding and positive market sentiment.

TABLE 5.3: Correlation of regression variables

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log amount raised	1									
Pre ICO	0.01	1								
Token price	(0.15)	0.11	1							
Accept fiat	0.00	0.22	(0.03)	1						
Blockchain industry	(0.01)	(0.05)	0.01	(0.06)	1					
Project stage: MVP	0.03	0.01	0.03	0.05	0.13	1				
Project stage: Beta version	0.09	(0.08)	(0.10)	0.00	(0.10)	(0.27)	1			
Project stage: Working product	(0.13)	0.08	0.06	(0.02)	(0.01)	(0.48)	(0.58)	1		
CEO experience	0.10	(0.03)	0.01	0.27	0.03	0.00	(0.04)	0.07	1	
GitHub stars	(0.11)	(0.17)	(0.07)	(0.03)	0.21	0.07	(0.05)	0.02	(0.02)	1

Note: Above table above shows the Pearson correlation matrix for all independent variables used in our regression for the set of 213 ICOs. We exclude Sentiment$_{t-2}$ for the purpose of this analysis, as it is a time-dependent variable, while this analysis is based on the campaign-level (a single observation per ICO). Log amount raised refers to the total amount raised by an ICO over the fundraising and exchange-trading phase. As such, it replicates the primary herding variable employed in the regression analysis, but differs from the fundraising variable shown in table 5.2. Brackets indicate negative correlation coefficients.

This finding is in line with the findings by Wick and Ihl (2019) for crowdfunding, and initial herding research for ICOs by Lee et al. (2019). *Column 2* shows the decision influence of our ICO attributes. All variables significantly affect the *Log amount raised*$_t$. ICOs that conducted a *Pre ICO* attract more capital. Hence a *Pre ICO* is interpreted as a quality signal by investors. Furthermore, both the number of *GitHub stars* and years of *CEO experience* lead to higher funding. This indicates that ICO investors, much like venture capital firms, evaluate intellectual and human/alliance capital signals positively in ICOs.

ICOs with higher token prices attract less capital than other ventures. It follows that lower-priced ICOs compensate for the lower funding amount per token by attracting more investors. *Accept fiat* also negatively affects *Log amount raised*$_t$. While this seems counterintuitive at first, it is quite intuitive. Dodgy businesses have used ICOs as an opportunity to scam investors in the light of media buzz and low regulation. An example is *Pincoin*, which turned out to be a professionally run Ponzi scheme. According to Techcrunch.com (2018, para 4), the company aimed at building an *"online collaborative consumption platform for the global community"*. They hosted several conferences and registered the company as a legal entity in Vietnam. After the ICO ended, managers fled the country with *600* million US Dollar, leaving behind fooled investors. While it is not always easy to distinguish scams from legitimate ICOs, awareness and fear of scams has grown. It turns out that *Accept fiat* makes it easier for ICOs to cash-out and is thus interpreted as a negative signal. Ventures that operate in the *Blockchain industry* attract less capital than ventures operating in other industries. This finding is surprising in the sense that blockchain technology, although far from understood by many, has triggered hype, often leading to over-valuation. A possible explanation is that this represents a risk-discount. Arguably, the risk of blockchain-related businesses is high, as blockchain technology's underlying value is uncertain.

Last, investors interpret the existence of an *MVP* as a positive signal. However,

TABLE 5.4: Regression results: Herding effects in ICOs

	Dependent variable:		
	Log amount raised$_t$		
	(1)	(2)	(3)
Log total amount raised$_{t-1}$	0.257*** (0.005)	0.257*** (0.005)	0.188*** (0.017)
Sentiment$_{t-2}$	0.550*** (0.204)	0.550*** (0.204)	0.527*** (0.203)
Pre ICO		1.374*** (0.082)	1.218*** (0.147)
Token price		−3.681*** (0.134)	−3.528*** (0.134)
Accept fiat		−54.178*** (1.941)	−53.339*** (1.958)
Blockchain industry		−11.910*** (0.401)	−12.489*** (0.427)
Project status: MVP		15.966*** (0.763)	13.869*** (0.818)
Project status: Beta version		−4.510*** (0.144)	−5.870*** (0.321)
Project status: Working product		0.059 (0.104)	−1.595*** (0.302)
CEO experience		0.120*** (0.012)	0.172*** (0.016)
GitHub stars		0.033*** (0.001)	0.036*** (0.001)
Log total amount raised$_{t-1}$*Pre ICO			0.009 (0.009)
Log total amount raised$_{t-1}$*Token price			−0.007*** (0.001)
Log total amount raised$_{t-1}$*Accept fiat			−0.006 (0.015)
Log total amount raised$_{t-1}$*Blockchain industry			0.053*** (0.011)
Log total amount raised$_{t-1}$*Project status: MVP			0.116*** (0.019)
Log total amount raised$_{t-1}$*Project status: Beta version			0.090*** (0.018)
Log total amount raised$_{t-1}$*Project status: Working product			0.105*** (0.018)
Log total amount raised$_{t-1}$*CEO experience			−0.004*** (0.001)
Log total amount raised$_{t-1}$*GitHub stars			−0.0002*** (0.00002)
Constant	0.965*** (0.213)	3.269*** (0.177)	4.526*** (0.319)
Campaign fixed effects	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes
Day-of-listing fixed effects	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes
Control: Campaign duration	No	Yes	Yes
Control: Secondary market	Yes	Yes	Yes
Observations	58,250	58,250	58,250
Adjusted R^2	0.568	0.568	0.569

Note: Above table shows the OLS regression results analyzing herding in the investment decisions of 1.8 million individual investors in 213 ICOs. Sentiment is based on social media content and was kindly provided by Thomson Reuters MarketPsych in the form of their TRMI cryptocurrency index.
$^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

more advanced project status does not seem to be valued by investors. Indeed, the existence of a *Beta version* is even viewed as a negative signal.

Model 3 includes interactions terms between the primary herding variable, *Log total amount raised*$_{t-1}$, and each ICO attribute. As described in the previous section, these interaction terms shed light on the behavioral processes underlying herding. As a reminder, herding is conventionally understood as an irrational process. Encouraged by the findings of Zhang and Liu (2012), we ask whether herding can also be a rational process in which investors strategically incorporate the behavior of other investors in their decision-making. Rational herding should be reflected in the interpretation of campaign attributes depending on the herding momentum, as reflected in *Log total amount raised*$_{t-1}$.

Our results provide evidence for such rational herding, with *four* attributes switching signs in the light of positive herding momentum. The binary variable *Blockchain industry* has a negative main effect on funding amount but turns positive when interacted with *Log total amount raised*$_{t-1}$. A venture operating in the *Blockchain industry* is interpreted as a positive signal if the venture has received high prior funding, hence validation from other investors. Similarly, both *CEO experience* and *GitHub stars* have a positive main effect and turn negative in the light of positive herding momentum. As such, credible signals for human/affiliation and intellectual capital become less important. Interestingly, *Beta version* also switches signs and is interpreted as a positive signal if the campaign has received more funding, and *Working product* becomes significant. Thus, while investors prefer earlier stage ventures, later-stage ventures become interesting if the crowd seems to think they are a good investment opportunity.

5.5.3 The role of investor experience

Next, we set out to investigate how investor experience affects herding behavior. The regression results are shown in *table 5.5*. *Model 1-3* are based on *novice investors* and *model 4-6* replicate these analyses for *experienced investors*. Disentangling investors reveals significant differences in the decision-making behavior of both groups. Comparing the main herding effects in *column 1* and *4*, we observe that while both investor groups herd, they do so based on different cues. *Novice investors* herd after social media (*Sentiment*$_{t-1}$) and observe the behavior of other investors (*Log total amount raised*$_{t-1}$). In contrast, experienced investors are unimpressed with the general market opinion, yet they follow the general investment crowd. This difference is interesting, as it may serve as evidence that investors are drawn into the market by the buzz and then decide on a campaign, rather than the other way around.

TABLE 5.5: Regression results: Investor experience and herding effects in ICO

	Dependent variable:					
	Log novice investments$_t$			log experienced investments$_t$		
	(1)	(2)	(3)	(4)	(5)	(6)
Log total amount raised$_{t-1}$	0.146*** (0.003)	0.146*** (0.003)	0.128*** (0.012)	0.195*** (0.004)	0.195*** (0.004)	0.157*** (0.015)
Sentiment$_{t-2}$	0.605*** (0.144)	0.605*** (0.144)	0.583*** (0.144)	−0.227 (0.176)	−0.227 (0.176)	−0.232 (0.176)
Pre ICO		0.777*** (0.065)	0.566*** (0.109)		1.500*** (0.079)	1.461*** (0.133)
Token price		0.038 (0.063)	0.087 (0.064)		−0.209*** (0.077)	−0.157** (0.078)
Accept fiat		−308.370*** (10.612)	−300.308*** (10.619)		−163.221*** (12.968)	−157.761*** (12.992)
Blockchain industry		−75.590*** (2.571)	−73.651*** (2.575)		−40.910*** (3.142)	−39.530*** (3.151)
Project status: MVP		−3.299*** (0.422)	−4.074*** (0.472)		−3.038*** (0.516)	−4.065*** (0.577)
Project status: Beta version		−2.159*** (0.089)	−2.858*** (0.221)		−2.697*** (0.109)	−3.745*** (0.271)
Project status: Working product		−0.026 (0.071)	−0.872*** (0.212)		−1.047*** (0.086)	−2.015*** (0.259)
CEO experience		0.237*** (0.013)	0.314*** (0.015)		0.218*** (0.015)	0.224*** (0.018)
GitHub stars		0.286*** (0.010)	0.280*** (0.010)		0.151*** (0.012)	0.148*** (0.012)
Log total amount raised$_{t-1}$*Pre ICO			0.013** (0.006)			0.001 (0.008)
Log total amount raised$_{t-1}$*Token price			−0.005*** (0.001)			−0.005*** (0.001)
Log total amount raised$_{t-1}$*Accept fiat			0.004 (0.010)			−0.046*** (0.013)
Log total amount raised$_{t-1}$*Blockchain industry			−0.0004 (0.008)			−0.022** (0.009)
Log total amount raised$_{t-1}$*Project status: MVP			0.062*** (0.013)			0.077*** (0.016)
Log total amount raised$_{t-1}$*Project status: Beta version			0.053*** (0.013)			0.073*** (0.016)
Log total amount raised$_{t-1}$*Project status: Working product			0.058*** (0.013)			0.063*** (0.015)
Log total amount raised$_{t-1}$*CEO experience			−0.007*** (0.001)			−0.001 (0.001)
Log total amount raised$_{t-1}$*GitHub stars			−0.0001*** (0.00002)			−0.0001*** (0.00002)
Constant	−3.329*** (0.150)	−4.018*** (0.177)	−3.542*** (0.259)	−3.088*** (0.184)	−2.995*** (0.216)	−2.227*** (0.316)
Campaign fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-listing fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control: Duration of campaign	No	Yes	Yes	No	Yes	Yes
Control: Secondary market	Yes	Yes	Yes	Yes	Yes	Yes
Observations	58,250	58,250	58,250	58,250	58,250	58,250
Adjusted R^2	0.573	0.573	0.575	0.487	0.487	0.488

Note: Above table shows the OLS regression results analysing herding effects in the investment decisions of 1.8 million individual investors in 213 ICOs. Sentiment is based on social media content and was kindly provided by Thomson Reuters MarketPsych in the form of their TRMI cryptocurrency index. Column 1-3 report the results for novice investors (first-time investors). Column 4-6 replicate the analyses for experienced investors, who have invested at least 9 times at the time of the investment under investigation.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Model 2 and *5* compare the main effects of ICO attributes and show that both groups incorporate ICO attributes in their decision-making. Interestingly, the crowds seem to interpret some cues differently. For example, experienced investors prefer investing in low-priced ICOs, while novice investor decisions are not influenced by the *Token price*. This could indicate that experienced investors seek arbitrage opportunities to exploit by scanning the market for highly discounted tokens. Ventures with *Working product* are avoided by experienced investors, while the effect is insignificant for novices. This may be interpreted as a relative preference for early-stage investments by experienced investors and a higher risk appetite. All other campaign attributes are interpreted similarly by both groups.

Model 3 and *6* again include interaction terms to help us distinguish rational herding. Surprisingly, we find evidence for rational herding for both crowds. *CEO experience*[6], *GitHub stars* and *Project status* change signs in light of positive herding momentum for both groups. *Token price* switches signs only for experienced investors. Consequently, we find evidence for experiential learning, as experienced investors' investment decisions are no longer influenced by market-wide sentiment. Nevertheless, we find that rational herding processes in the form of observational learning are in place even for novice market participants.

5.5.4 The role of campaign phase

Herding effects may differ significantly between the primary (fundraising) and secondary (exchange-trading) ICO phase. We expect the influence of ICO attributes to diminish over time, as new information about the venture becomes available and the venture progresses. We expect the influence of ICO attributes to remain high in our results. Our analyses in *chapter 4*, *section 4.4.2* show that

[6]Although not significant in the case of experienced investors.

TABLE 5.6: Regression results: Herding effects during the primary stage (ICO)

	Dependent variable:					
	Log amount raised$_t$		Log novice investments$_t$		Log experienced investments$_t$	
	(1)	(2)	(3)	(4)	(5)	(6)
Log total amount raised$_{t-1}$	0.049*** (0.010)	0.026 (0.034)	0.038*** (0.007)	0.008 (0.024)	0.018** (0.008)	0.056** (0.027)
Sentiment$_{t-2}$	0.168 (0.854)	0.180 (0.853)	0.829 (0.619)	0.789 (0.616)	−0.404 (0.690)	−0.451 (0.690)
Pre ICO	−0.999*** (0.359)	−1.229*** (0.430)	−0.662*** (0.252)	−0.449 (0.301)	−0.861*** (0.281)	−0.984*** (0.337)
Token price	−1.018*** (0.157)	−0.993*** (0.158)	−0.717*** (0.105)	−0.666*** (0.106)	−0.697*** (0.117)	−0.713*** (0.119)
Accept fiat	−75.343 (48.841)	−61.614 (49.053)	−160.044*** (35.511)	−143.676*** (35.516)	−0.473 (39.540)	−4.807 (39.771)
Blockchain industry	−18.986 (11.715)	−15.756 (11.768)	−39.866*** (8.520)	−35.839*** (8.523)	−2.065 (9.487)	−3.063 (9.544)
Project status: MVP	1.659** (0.769)	0.374 (0.870)	1.620*** (0.468)	−0.003 (0.551)	1.050** (0.521)	1.122* (0.617)
Project status: Beta version	0.836 (0.882)	1.522 (0.961)	0.528 (0.423)	0.755 (0.506)	0.735 (0.470)	1.594*** (0.567)
Project status: Working product	1.157 (0.789)	0.829 (0.854)	1.262*** (0.386)	0.683 (0.466)	0.125 (0.429)	0.424 (0.522)
CEO experience	0.274*** (0.071)	0.259*** (0.075)	0.210*** (0.052)	0.202*** (0.055)	0.162*** (0.058)	0.197*** (0.061)
GitHub stars	0.070 (0.045)	0.058 (0.045)	0.148*** (0.033)	0.134*** (0.033)	0.001 (0.036)	0.005 (0.037)
Log total amount raised$_{t-1}$*Pre ICO		0.022 (0.020)		−0.018 (0.014)		0.015 (0.016)
Log total amount raised$_{t-1}$*Token price		−0.0001 (0.002)		−0.002 (0.001)		0.001 (0.002)
Log total amount raised$_{t-1}$*Accept fiat		−0.031 (0.032)		0.001 (0.023)		0.009 (0.026)
Log total amount raised$_{t-1}$*Blockchain industry		−0.001 (0.028)		−0.006 (0.021)		−0.001 (0.023)
Log total amount raised$_{t-1}$*Project status: MVP		0.111*** (0.039)		0.148*** (0.028)		−0.007 (0.031)
Log total amount raised$_{t-1}$*Project status: Beta version		−0.072* (0.038)		−0.022 (0.028)		−0.086*** (0.031)
Log total amount raised$_{t-1}$*Project status: Working product		0.017 (0.035)		0.043* (0.025)		−0.031 (0.028)
Log total amount raised$_{t-1}$*CEO experience		0.001 (0.002)		−0.0003 (0.001)		−0.003* (0.001)
Log total amount raised$_{t-1}$*GitHub stars		−0.00004 (0.00005)		−0.0001** (0.00003)		0.00001 (0.00004)
Constant	5.646*** (1.038)	6.091*** (1.093)	−2.476*** (0.681)	−1.786** (0.723)	−0.233 (0.758)	−0.661 (0.810)
Campaign fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-listing fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control: Duration of campaign	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,298	5,298	5,298	5,298	5,298	5,298
Adjusted R^2	0.522	0.525	0.562	0.567	0.606	0.607

Note: Above table shows the OLS regression results analysing herding effects during the primary ICO phase, i.e., during the fundraising. Column 1 and 2 show the results for all investors, Column 3 and 4 show results for novice (first-time) investors and Column 5 and 6 replicate the results for experienced investors (they have invested at least 9 times).

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE 5.7: Regression results: Herding effects during the secondary stage (crypto-exchanges)

	Dependent variable:					
	Log amount raised$_t$		Log novice investments$_t$		Log experienced investments$_t$	
	(1)	(2)	(3)	(4)	(5)	(6)
Log total amount raised$_{t-1}$	0.353*** (0.007)	0.090** (0.040)	0.190*** (0.005)	0.042 (0.028)	0.273*** (0.006)	0.045 (0.035)
Sentiment$_{t-2}$	0.634*** (0.198)	0.635*** (0.197)	0.667*** (0.139)	0.658*** (0.139)	−0.212 (0.170)	−0.208 (0.170)
Pre ICO	1.451*** (0.088)	0.678*** (0.210)	0.733*** (0.070)	−0.136 (0.151)	1.637*** (0.085)	0.481*** (0.185)
Token price	−3.880*** (0.140)	−3.569*** (0.143)	−0.053 (0.064)	−0.039 (0.067)	−0.353*** (0.078)	−0.283*** (0.082)
Accept fiat	−55.119*** (2.045)	−53.324*** (2.082)	−295.454*** (11.132)	−284.109*** (11.151)	−173.355*** (13.612)	−157.485*** (13.652)
Blockchain industry	−12.868*** (0.431)	−12.182*** (0.487)	−72.480*** (2.693)	−68.478*** (2.702)	−42.962*** (3.293)	−38.400*** (3.309)
Project status: MVP	18.046*** (0.787)	12.494*** (1.052)	−2.162*** (0.446)	−4.579*** (0.664)	−1.450*** (0.545)	−4.400*** (0.813)
Project status: Beta version	−3.969*** (0.158)	−8.937*** (0.698)	−1.848*** (0.097)	−5.213*** (0.488)	−2.233*** (0.118)	−6.617*** (0.598)
Project status: Working product	0.479*** (0.101)	−4.666*** (0.689)	0.185*** (0.071)	−2.710*** (0.485)	−0.703*** (0.087)	−4.837*** (0.594)
CEO experience	0.070*** (0.013)	0.156*** (0.021)	0.197*** (0.014)	0.332*** (0.018)	0.179*** (0.017)	0.189*** (0.022)
GitHub stars	0.032*** (0.001)	0.035*** (0.001)	0.273*** (0.010)	0.262*** (0.010)	0.159*** (0.013)	0.147*** (0.013)
Log total amount raised$_{t-1}$*Pre ICO		0.050*** (0.013)		0.058*** (0.009)		0.075*** (0.011)
Log total amount raised$_{t-1}$*Token price		−0.015*** (0.002)		−0.005*** (0.001)		−0.013*** (0.001)
Log total amount raised$_{t-1}$*Accept fiat		0.037* (0.021)		0.071*** (0.015)		−0.031* (0.018)
Log total amount raised$_{t-1}$*Blockchain industry		−0.027* (0.017)		−0.082*** (0.012)		−0.069*** (0.014)
Log total amount raised$_{t-1}$*Project status: MVP		0.297*** (0.042)		0.161*** (0.030)		0.207*** (0.037)
Log total amount raised$_{t-1}$*Project status: Beta version		0.320*** (0.041)		0.225*** (0.029)		0.286*** (0.036)
Log total amount raised$_{t-1}$*Project status: Working product		0.315*** (0.041)		0.179*** (0.029)		0.254*** (0.036)
Log total amount raised$_{t-1}$*CEO experience		−0.007*** (0.001)		−0.012*** (0.001)		−0.003*** (0.001)
Log total amount raised$_{t-1}$*GitHub stars		−0.0002*** (0.00003)		−0.00001 (0.00002)		−0.0002*** (0.00003)
Constant	1.347*** (0.220)	5.827*** (0.700)	−4.724*** (0.200)	−2.140*** (0.509)	−4.353*** (0.244)	−0.479 (0.623)
Campaign fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-listing fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control: Duration of campaign	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,952	52,952	52,952	52,952	52,952	52,952
Adjusted R^2	0.619	0.621	0.615	0.619	0.533	0.536

Note: Above table shows the OLS regression results analysing herding effects during the secondary ICO phase, i.e., while the ICO tokens are trading on crypto-exchanges. Column 1 and 2 show the results for all investors, Column 3 and 4 show results for novice (first-time) investors and Column 5 and 6 replicate the results for experienced investors (they have invested at least 9 times)

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

trading activity is concentrated around the ICO phase, indicating that most secondary market investments occur shortly after the ICO. Synthesized with our findings in *table 5.2*, showing that most participants invest in secondary markets, we hypothesize that many market participants invest shortly after the exchange-listing of a particular ICO. In consequence, we expect these investors' decisions to remain influenced by ICO attributes. The social influence might be increased during exchange-trading, as crowd behavior is directly observable in trading volume and market price. Last, systematically different investors may self-select to prefer investments in a particular ICO phase.

The regression results investigating the primary stage (ICO) are summarized in *table 5.6*. *Table 5.7* replicates the analyses for the secondary phase (crypto-exchanges). Our results show that investors herd in both phases; yet, herding processes differ by phase. Primary phase herding is induced solely by *Log total amount raised*$_{t-1}$, while secondary phase herding is also induced by *Sentiment*$_{t-2}$. Our analyses further reveal interesting nuances in the behavior of novice and experienced investors.

Results for the primary phase substantially differ from those in *section 5.5.3*. Specifically, we find that during fundraising experienced investors show weak herding based on *Log total amount raised*$_{t-1}$ (significant only at 5% level). Instead, their decision-formation is primarily based on business fundamentals (ICO attributes). It follows that we also find weak evidence for rational herding in the interaction terms. In contrast, novice investors herd in both phases. However, primary stage herding is not induced by *Sentiment*$_{t-2}$, compared to the findings in the previous section. Novice investor herding during the primary phase seems to be exclusively irrational, given that the interpretations of ICO attributes show no sign of switches in the light of positive herding momentum. In conclusion, we find evidence for experiential learning during the primary phase. While novice investors herd irrationally, herding effects weaken with experience. This implies that experience moderates irrational herding.

In contrast, the secondary stage (exchange-trading) results resemble our previous findings from *section 5.5.3*. Therefore, we find evidence for pronounced herding during the secondary phase, induced by $Sentiment_{t-2}$ and $Log\ total\ amount\ raised_{t-1}$ for novice investors, and $Log\ total\ amount\ raised_{t-1}$ for experienced investors. Both groups show rational herding tendencies, revealed by significant interaction terms between the primary herding variable and ICO attributes. Consequently, we find evidence for the moderating role of experience in the secondary stage and experienced investor herding is primarily rationally. In contrast, novice investor herding indicates mixed effects from rational and irrational herding.

Our results also suggest that primary and secondary stage investors interpret ICO attributes slightly differently. *Pre ICO* is interpreted negatively during the primary stage and positive during exchange-trading; the opposite effect is observed for *Project status: MVP*. Indeed, for the overall crowd, independent of experience level, ICO attributes seem to become more important during the secondary stage. We see *two* possible explanations. First, large discounts during fundraising make investors less sensitive to ICO attributes, causing them to select instead tokens based on (possible) arbitrage opportunities in the form of underpriced tokens. Second, systematically different investors may self-select to invest in both phases. In summary, our findings suggest that investors become more susceptible to herding once a token is trading on crypto-exchanges, especially sentiment-induced herding. In both phases, we find evidence for experience as a moderating factor, causing investor behavior to become less irrational.

5.5.5 Sensitivity of results

We perform several sensitivity analyses to test our results' stability. In our first set of analyses, we test how our results are affected by employing $Sentiment_{t-2}$ as the primary herding variable. As discussed, sentiment measures the market's attitude towards cryptocurrencies. In consequence, sentiment attracts investors to the ICO market, as our previous analyses show. It should not alter an investors' interpretation of ICO attributes, as it does not reveal additional information relevant about venture quality. Thus, if investors behaved rationally, we would not expect to see significant interaction effects between $Sentiment_{t-2}$ and ICO attributes. The main results are shown in *table 5.8*. A single variable, *GitHub stars* switches sign in light of positive herding momentum, providing weak evidence for irrational, sentiment-induced behavior. *Project status MVP* and *CEO experience* increase in importance with the increasing sentiment, indicating that quality signals become more important in times of positive sentiment, possibly because more participants *flood* the market and investors become more careful.

Table 5.9 replicates the analysis for novice and experienced investors. Again, we find weak evidence for irrational herding effects based on $Sentiment_{t-2}$. For experienced investors, the signs of the *three* project status variables switch, indicating that with positive market sentiment, venture quality signals become important (e.g., some form of a functioning product).

Last, we investigate the herding behavior of the medium experience investor crowd (*Log medium experience investments*$_t$), which we excluded in our experience analyses, defined as investors with more than one but less than *nine* prior investments. The results are summarized in *table 5.10*. We find that sentiment-induced herding dominates versus amount-based herding, similar to our finding for novice investors. Nevertheless, the magnitude of sentiment-induced herding is smaller for the medium investment crowd, while amount-based herding increases. We also find evidence for rational herding in the form of alternating interpretation of campaign cues with positive herding momentum (based on *Log total amount raised*$_{t-1}$) for *four* variables: *Blockchain industry*, *Project status: MVP*, *product status: Working product* and *GitHub stars*. Sentiment-induced alternations of ICO attributes are found for *Token price* and *GitHub stars*.

TABLE 5.8: Sensitivity analysis: Sentiment-based herding in ICOs

	Dependent variable:		
	Log amount raised$_t$		
	(1)	(2)	(3)
Log total amount raised$_{t-1}$	0.257*** (0.005)	0.257*** (0.005)	0.258*** (0.005)
Sentiment$_{t-2}$	0.550*** (0.204)	0.550*** (0.204)	−1.333 (0.845)
Pre ICO		1.374*** (0.082)	1.745*** (0.093)
Token price		−3.681*** (0.134)	0.152* (0.089)
Accecpt fiat		−54.178*** (1.941)	−361.483*** (14.985)
Blockchain industry		−11.910*** (0.401)	−89.048*** (3.630)
Project status: MVP		15.966*** (0.763)	−6.943*** (0.598)
Project status: Beta version		−4.510*** (0.144)	−3.818*** (0.130)
Project status: Working product		0.059 (0.104)	−0.922*** (0.104)
CEO experience		0.120*** (0.012)	0.351*** (0.018)
GitHub stars		0.033*** (0.001)	0.337*** (0.014)
Sentiment$_{t-2}$*Pre ICO			0.564 (0.403)
Sentiment$_{t-2}$*Token price			0.014 (0.052)
Sentiment$_{t-2}$*Accept fiat			−0.034 (0.650)
Sentiment$_{t-2}$*Blockchain industry			0.624 (0.531)
Sentiment$_{t-2}$*Project status: MVP			2.368** (0.931)
Sentiment$_{t-2}$*Project status: Beta version			0.550 (0.895)
Sentiment$_{t-2}$*Project status: Working product			1.190 (0.874)
Sentiment$_{t-2}$*CEO experience			0.136*** (0.037)
Sentiment$_{t-2}$*GitHub stars			−0.003*** (0.001)
Constant	0.965*** (0.213)	3.269*** (0.177)	0.552** (0.252)
Campaign fixed effects	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes
Day-of-listing fixed effects	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes
Control: Duration of campaign	No	Yes	Yes
Control: Secondary market	Yes	Yes	Yes
Observations	58,250	58,250	58,250
Adjusted R^2	0.568	0.568	0.568

Note: Above table reports the OLS regression results analysing social media sentiment based herding of investors in ICO campaigns (213) between 2015 and 2018. Sentiment is based on social media content and was kindly provided by Thomson Reuters MarketPsych in the form of their TRMI cryptocurrency index.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE 5.9: Sensitivity analysis: Investor experience and sentiment-based herding in ICOs

	Dependent variable:					
	Log novice investments$_t$			log experienced investments$_t$		
	(1)	(2)	(3)	(4)	(5)	(6)
Log total amount raised$_{t-1}$	0.146*** (0.003)	0.146*** (0.003)	0.146*** (0.003)	0.195*** (0.004)	0.195*** (0.004)	0.196*** (0.004)
Sentiment$_{t-2}$	0.605*** (0.144)	0.605*** (0.144)	0.196 (0.599)	−0.227 (0.176)	−0.227 (0.176)	−3.491*** (0.731)
Pre ICO		0.777*** (0.065)	0.753*** (0.066)		1.500*** (0.079)	1.496*** (0.080)
Token price		0.038 (0.063)	0.037 (0.063)		−0.209*** (0.077)	−0.215*** (0.077)
Accecpt fiat		−308.370*** (10.612)	−308.397*** (10.614)		−163.221*** (12.968)	−162.195*** (12.961)
Blockchain industry		−75.590*** (2.571)	−75.539*** (2.571)		−40.910*** (3.142)	−40.668*** (3.140)
Project status: MVP		−3.299*** (0.422)	−3.330*** (0.423)		−3.038*** (0.516)	−3.077*** (0.517)
Project status: Beta version		−2.159*** (0.089)	−2.155*** (0.092)		−2.697*** (0.109)	−2.747*** (0.112)
Project status: Working product		−0.026 (0.071)	−0.032 (0.074)		−1.047*** (0.086)	−1.121*** (0.090)
CEO experience		0.237*** (0.013)	0.237*** (0.013)		0.218*** (0.015)	0.205*** (0.016)
GitHub stars		0.286*** (0.010)	0.286*** (0.010)		0.151*** (0.012)	0.150*** (0.012)
Sentiment$_{t-2}$*Pre ICO			0.654** (0.285)			0.137 (0.349)
Sentiment$_{t-2}$*Token price			0.017 (0.037)			0.026 (0.045)
Sentiment$_{t-2}$*Accept fiat			0.245 (0.461)			−1.454*** (0.563)
Sentiment$_{t-2}$*Blockchain industry			−0.964** (0.376)			0.311 (0.459)
Sentiment$_{t-2}$*Project status: MVP			1.148* (0.659)			3.316*** (0.805)
Sentiment$_{t-2}$*Project status: Beta version			−0.165 (0.634)			1.628** (0.774)
Sentiment$_{t-2}$*Project status: Working product			0.153 (0.619)			2.018*** (0.756)
Sentiment$_{t-2}$*CEO experience			0.018 (0.026)			0.289*** (0.032)
Sentiment$_{t-2}$*GitHub stars			−0.001 (0.001)			−0.002* (0.001)
Constant	−3.329*** (0.150)	−4.018*** (0.177)	−4.012*** (0.178)	−3.088*** (0.184)	−2.995*** (0.216)	−2.876*** (0.218)
Campaign fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-listing fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control: Duration of campaign	No	Yes	Yes	No	Yes	Yes
Control: Secondary market	Yes	Yes	Yes	Yes	Yes	Yes
Observations	58,250	58,250	58,250	58,250	58,250	58,250
Adjusted R^2	0.573	0.573	0.573	0.487	0.487	0.488

Note: Above table reports the OLS regression results analysing social media sentiment effects on the investment decisions of investors with different prior ICO investment experience. Column 1-3 report the results for novice investors (first-time investors). Column 4-6 replicate the analyses for experienced investors, who have invested at least nine times at the time of the concerning investment.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE 5.10: Sensitivity analysis: Extended investor experience and herding in ICOs

	Dependent variable:		
	Log medium experience investments$_t$		
	(1)	(2)	(3)
Log total amount raised$_{t-1}$	0.161*** (0.004)	0.121*** (0.013)	0.162*** (0.004)
Sentiment$_{t-2}$	0.380** (0.157)	0.369** (0.156)	−0.090 (0.649)
Pre ICO	0.966*** (0.071)	0.887*** (0.118)	0.941*** (0.071)
Token price	−0.314*** (0.068)	−0.250*** (0.069)	−0.325*** (0.068)
Accept fiat	−290.055*** (11.512)	−288.197*** (11.533)	−288.714*** (11.507)
Blockchain industry	−71.469*** (2.789)	−71.480*** (2.797)	−71.136*** (2.787)
Project status: MVP	−2.132*** (0.458)	−3.283*** (0.512)	−2.091*** (0.459)
Project status: Beta version	−2.243*** (0.096)	−3.016*** (0.240)	−2.192*** (0.100)
Project status: Working product	−0.005 (0.077)	−0.767*** (0.230)	0.023 (0.080)
CEO experience	0.284*** (0.014)	0.291*** (0.016)	0.277*** (0.014)
GitHub stars	0.268*** (0.011)	0.268*** (0.011)	0.267*** (0.011)
Log total amount raised$_{t-1}$*Pre ICO		0.005 (0.007)	
Log total amount raised$_{t-1}$*Token price		−0.005*** (0.001)	
Log total amount raised$_{t-1}$*Accept fiat		−0.032*** (0.011)	
Log total amount raised$_{t-1}$*Blockchain industry		0.026*** (0.008)	
Log total amount raised$_{t-1}$*Project status: MVP		0.082*** (0.014)	
Log total amount raised$_{t-1}$*Project status: Beta version		0.050*** (0.014)	
Log total amount raised$_{t-1}$*Project status: Working product		0.048*** (0.014)	
Log total amount raised$_{t-1}$*CEO experience		−0.001 (0.001)	
Log total amount raised$_{t-1}$*GitHub stars		−0.0001*** (0.00002)	
Sentiment$_{t-2}$*Pre ICO			0.653** (0.309)
Sentiment$_{t-2}$*Token price			0.138*** (0.040)
Sentiment$_{t-2}$*Accept fiat			−0.249 (0.499)
Sentiment$_{t-2}$*Blockchain industry			0.385 (0.408)
Sentiment$_{t-2}$*Project status: MVP			1.099 (0.715)
Sentiment$_{t-2}$*Project status: Beta version			−1.134* (0.687)
Sentiment$_{t-2}$*Project status: Working product			−0.720 (0.671)
Sentiment$_{t-2}$*CEO experience			0.161*** (0.029)
Sentiment$_{t-2}$*GitHub stars			−0.003*** (0.001)
Constant	−3.776*** (0.192)	−3.021*** (0.281)	−3.761*** (0.193)
Campaign fixed effects	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes
Day-of-listing fixed effects	Yes	Yes	Yes
Quarter fixed effects	Yes	Yes	Yes
Control: Duration of campaign	Yes	Yes	Yes
Control: Secondary market	Yes	Yes	Yes
Observations	58,250	58,250	58,250
Adjusted R^2	0.569	0.569	0.569

Note: Above table reports the OLS regression results analysing herding effects based on investment decisions by other investors and social media sentiment on investment decisions of individual investors in 213 ICOs. Medium experience investments are classified as investments by investors who have invested at least twice but less than nine times prior to the concerning investment (in ICOs).
**p<0.1; **p<0.05; ***p<0.01*

5.6 Discussion

Our findings are consistent with a range of studies documenting herding effects in group-based investment settings (Kallinterakis and Wang, 2019; Gurdgiev and O'Loughlin, 2020; Devenow and Welch, 1996; Simonsohn and Ariely, 2008; Ajaz and Kumar, 2018; Zhang and Liu, 2012; Yum et al., 2012; Chen and Lin, 2014). We contribute a perspective on Initial Coin Offerings, showing that investors are also susceptible to herding in an ICO context. Our findings are also in line with initial research of herding in ICOs, as evidenced by, e.g., Lee et al. (2019), who shows that first-day ICO-subscriptions are a strong predictor for follow-up subscriptions. Nevertheless, our study contributes *four* novel perspectives.

First, we show that herding in ICOs is partially explained by rational behavior, as investors utilize others' behavior as an additional cue in their decision-making. This finding is in line with the findings by Zhang and Liu (2012) for microloan-lending and opposes advocates of the universal irrationality of herding, e.g., Ajaz and Kumar (2018). Therefore, our findings yield support for observational learning theory in the context of ICOs, which states that investors learn through experience and adjust their behavior accordingly (Devenow and Welch, 1996).

Second, we extend the understanding of the factors moderating irrational herding by studying the role of experience. Our findings suggest that inexperienced investors are more susceptible to irrational herding, as shown in the example of sentiment-induced herding. These findings are in line with research on stock market bubbles, e.g., Griffin et al. (2011) and Greenwood and Nagel (2009). We show that inexperienced investors dominate the ICO market. While not explicitly researched in this study, our findings point to worrisome parallels between the dynamics observed during the rise and bursting of stock market bubbles in the past and the current state ICO markets. During stock market bubbles, inexperienced investors with a stronger tendency towards trend-chasing behavior also dominated the market and were deliberately exploited by informed market partic-

ipants (Greenwood and Nagel, 2009; Binglin et al., 2017; Griffin et al., 2011).
Third, we show that herding differs by ICO phase. Our findings support the findings by, e.g., Gurdgiev and O'Loughlin (2020) and Kallinterakis and Wang (2019), showing that crypto investors irrationally herd with general market sentiment. We show that this tendency is restricted to secondary markets, as our findings provide no evidence for sentiment-induced herding during fundraising. Our analyses also enable us to draw an initial conclusion on the effects of introducing secondary markets in crowdfunding contexts, as called upon by crowdfunding research (Turan, 2015; Mochkabadi and Volkmann, 2018). We find evidence for rational herding among investors in secondary markets, as investors tend to amend their ICO attribute interpretation in the light of positive herding momentum. Our findings point to the value of secondary markets as a mechanism for cascading information. Nevertheless, our findings also pinpoint the flipside of the coin. Secondary markets also increase the general tendency to herd; particularly that of irrational, sentiment-induced herding by inexperienced investors.
Last, our findings highlight the importance of disentangling the crowd. We show that inexperienced investors dominate the ICO market. This would falsely lead us to conclude that market participants herd irrationally if we did not disentangle the crowd. Such differentiation of investors in high heterogeneity contexts has previously led to important insights into the functioning of crowdfunding. Abrams (2017), for example, shows that early and late investors differ substantially in their decision-making processes. Our findings build on this and show that, depending on their experience, investors interpret ICO attributes differently. Our findings highlight the potential fallacies of applying signaling theory in high heterogeneity contexts, a common approach in the early stages of ICO investor research (Amsden and Schweizer, 2018; Cerchiello, 2018; Fisch, 2019; Lee et al., 2019). Similarly, Wick and Ihl (2019) show that herding depends on the nature of the observed subgroup, distinguishing investors based on their contribution pattern as a proxy for investor

background. Their findings show that experts purposefully avoid novice crowds, while novice and expert investors single out and follow investments by perceived experts. Their findings align with our finding, suggesting that behavioral processes underlying observed herding differs between investors, depending on their background.

While attempts to differentiate subgroups of ICO investors are scarce, some significant progress has been made. For example, Boreiko and Risteski (2020) study the learning effects of ICO investors, thereby recognizing the importance of disentangling the crowd. Their findings suggest that, while the average investor does not learn through experience, expert investors[7] do learn over time. This finding is in line with the conclusion we draw from our analyses. Even as the overall crowd herds irrationally, some market participants herd rationally depending on the context (phase) and background (experience). We believe that our findings are complementary to those of Boreiko and Risteski (2020). While we find that learning reduces irrationality, Boreiko and Risteski (2020) asserts that learning effects may differ between subgroups, which is very well possible but not researched in our study. They also choose to differentiate investors by *expertise* rather than *experience*, making their findings not directly comparable to ours. Still, both studies conclude that high heterogeneity exists and differentiating investors is vital for understanding their behavior.

5.6.1 Limitations

Our research is not without limitations. First, our dataset does not allow us to link investor decisions to realized outcomes, i.e., venture success. This would enable us to draw higher quality inferences about the rationality of herding. Given the topic's recent development and the length of time required to build a company, most of the projects in our data have not yet come to realize their

[7] Defined as investors with large financial commitments to a given campaign.

potential. As such, no inferences can be drawn about the factual quality of investor decisions. We could have approximated success as fundraising success. We believe that fundraising success is not a good approximation for venture quality in an environment characterized by high uncertainty. Second, the patchiness of our ICO attribute data forced us to select a small number of ICO attributes for our analyses. While we feel that these variables allow us to extract valuable insights, we understand that the picture drawn here is far from complete. We mitigated the statistical biases by controlling for unobserved campaign heterogeneity and hence feel comfortable that we are not distorting any results. Third, our blockchain transaction data had missing values and inaccurate transaction timestamps due to technical imperfections. This forced us to exclude observations of *431* ICOs and *5.8* million investors in our analyses. Even though we have no reason to believe so, we cannot rule out the possibility that these ICOs systematically differ from those observed, thus biasing our results. Last, the sentiment measure used is not campaign-specific. Replicating the analyses with campaign-specific sentiment would allow for a more detailed understanding of decision-weights at a campaign-level.

5.7 Conclusion

Our study is the first of its kind to investigate the rationality of herding in ICOs. We base our analyses on a novel large-scale panel dataset, tracking the fundraising process and of *213* ICOs based on *17* million blockchain transactions, enriched with ICO attributes and social media sentiment. We find evidence for pronounced herding in ICOs, which is in line with initial research for ICOs (e.g., Lee et al., 2019) and comparable group-based investment settings, e.g., crowdfunding (Vismara, 2018) or microloan-lending (Zhang and Liu, 2012).

We further investigate whether herding in the context of ICOs may serve as a mechanism for cascading information in the face of information asymmetries. Under this assumption, herding may be regarded as rational behavior and better described as the result of observational learning. Our results show that ICO herding is primarily irrational. Nevertheless, our findings also point to the value of disentangling herding effects by showing that the rationality of ICO herding depends upon investor experience and the observed fundraising phase.

Specifically, we show that irrational herding is moderated by investor experience. Novice investors herd after a general market hype and momentum. This tendency disappears as investors gain experience and they learn to observe others to aid their decision-making strategically. Our findings also indicate that herding is more pronounced on crypto-exchanges (as compared to fundraising). We find evidence for an increased tendency to herd rationally across investor groups, indicating the value of secondary markets as an opportunity to resolve information asymmetry. However, we also show that inexperienced investors become more susceptible to sentiment-induced, irrational herding. Hence, our findings empirically contribute to the understanding of herding in ICOs by investigating *three* novel perspectives: the moderating role of experience, the rationality of herding in ICOs, and the role of secondary markets as a mechanism to resolve information asymmetries.

Our research holds practical implications for policymakers. Our findings highlight the potentially dangerous dynamics in ICO crowds by pointing to inexperienced investors' predominance and irrationality. In the past, such situations have led to dramatic economic crises (Binglin et al., 2017). Therefore, our findings represent a call to introduce regulatory mechanisms to protect the investors' wealth, mostly inexperienced ones.

Chapter 6

Wisdom of the crowd in ICOs? The decision-quality of ICO crowds

6.1 Abstract

Considerable academic debate exists about the quality of crowd-based decision-making. The context of early-venture financing is particularly challenging, as the quality of the decision outcome is less clear. This study investigates the quality of crowd-based investment decisions in the context of blockchain-based crowdfunding (ICOs). We ask whether ICO crowds channel capital to prospering ventures and if expert investors possess superior decision-making capabilities. We analyze the crowd composition in *198* ICOs based on *17* million blockchain transactions and link them to ICO post-campaign success and ICO attributes. We find that *45%* of ICOs have failed within our observation period, and the average ICO earns *-19%* return. It follows that ICOs seem to have a higher failure rate than ventures in comparative contexts and destroy wealth, on average. We also show that general crowds do not succeed in channelling capital to prospering ventures. Yet, we find weak evidence that experienced crowds tend to form better investment decisions than novice investors in selecting higher-performing ICOs.

6.2 Introduction

According to Prosser (2016), *90%* of ventures fail and less than *50%* make it to their *5th* year (Chernev, 2020). From an investor perspective, early-venture financing is a risky business. Yet, ventures that do succeed reward investors with high payoffs. For example, the investors of *Ethereum*, a customizable blockchain infrastructure, earned *4,520x* their investment within *three* years (Merre, 2020).

Indeed, an entire industry has formed around start-up financing, namely venture capital. VC firms bundle the capital of investors and purchase equity in start-ups. Ultimately, venture capital firms seek to resell their share at a profit, usually through an Initial Public Offering (IPO). They go through a lengthy process to select high-prospect firms from the crowd of capital seeking ventures. Research has termed this selection process the *scouting* role of venture capital firms. It has been argued that VC-*scouting* is reflected in lower failure rates of VC-backed start-ups (Gage, 2012).

A recent innovation, Initial Coin Offerings (ICOs), may have the potential turn this industry upside down. In ICOs, ventures raise capital from many contributors via the blockchain. ICOs may close a funding gap left by professional investors. Nevertheless, early-stage ventures are risky, as shown by their high failure rate. They are also inherently difficult to value without track records and functional products in place. Retail investors may lack the capability and resources to perform diligent investment analysis in this high-uncertainty context. The question arises whether crowds can channel capital to prospering ventures, i.e., resume the *scouting* role otherwise performed by VCs.

Due to the ICO phenomenon's recency and the inherent difficulty in objectively evaluating the quality of decision-making in the context of early-venture financing, little is known about the *scouting* ability of ICO crowds. In crowdfunding, a context comparable to ICOs, considerable debate has arisen on crowd decision-quality effectiveness. For example, Mollick (2014) argues that crowd decision-making, in some settings, may be superior to expert decision-making, driven by greater diversity, so-called *wisdom of the crowds*. Others, like Isenberg (2012), argue that crowd-based start-up funding is doomed to fail. In his line of argumentation, the crowd consists of *"unwitting consumers who simply cannot know enough about the highly risky ventures or the highly complex venture investing process to make informed investment decisions"* (Isenberg, 2012, para 1).

In the context of ICOs, the majority of research has focused on fundraising instead of venture success (e.g., Amsden and Schweizer, 2018; Cordova et al., 2015; Cerchiello, 2018). Little is known about the longer-term post-campaign success of ICOs altogether. The majority of research in this young field has focused on the short-term token return to investigate, e.g., the determinants of ICO underpricing (e.g., Benedetti and Kostovetsky, 2018; Lyandres et al., 2019). Nevertheless, ICOs represent a particularly interesting environment to study the effectiveness of crowd-based investment decisions. ICOs are assumed to attract more financially motivated contributors than crowdfunding, as most tokens are traded on exchanges, making them liquid and subject to speculation. Moreover, considerable hype has surged around ICOs, leading to many novices entering the market. Consequently, ICOs bring together a diverse crowd in background and motivation.

In this research, we aim to investigate crowd decision-making effectiveness in the context of ICOs. We study whether crowds select ventures with higher post hoc performance and expert crowds' ability to make higher-quality decisions than novices. Our analysis is based on *17* million blockchain transactions, which we employ to catalyze ICO crowd composition of *198* ICOs. We link the data with ICO attributes and post-campaign success. Our findings show that *45%* of ICOs in our sample failed within *two* years, and most ICOs have generated high negative abnormal returns. We also show that ICOs attract systematically different crowds, yet, the crowds' composition has little influence on venture success.

The remainder of this paper is structured as follows. The first section elaborates on the theoretical background, discussing the role of ownership in VC-backed and crowdfunded ventures. The second section introduces the data sample and further elaborates on the methodology employed in our analyses. The third section is divided into *two* complementary analyses. First, we investigate the stability of crowd composition over an ICOs' life cycle. Second, we investigate the influence of crowd composition on success. We conclude this paper with a discussion, placing our findings in a broader context of related research and subsequently elaborate on our research limitations.

6.3 Theoretical background

6.3.1 Venture capital and venture success

Research indicates that failure rates are lower for VC-backed start-ups, *75%* (in comparison to *90%* across all start-ups, independent of their source of financing) according to Chernev (2020). The lower failure rates may be attributed to *three* factors. First, VC-backed start-ups have already been able to secure the capital required to nurture their business, taking them a step ahead of other capital seeking firms. Second, VCs take on a *scouting* role in selecting high-prospect start-ups, hence the set of VC-backed ventures consists of *higher quality* start-ups. Third, beyond the provision of capital, VCs provide start-ups with resources that boost their success chances, such as (industry) expertise, (management) experience, reputation or network connections, the so-called *coaching* role of VCs.

Research from the *scouting* lens has focused on investigating how VCs form investment decisions (Streletzki and Schulte, 2013; Schefczyk and Gerpott, 2001; Fried and Hisrich, 1994; Zacharakis and Shepherd, 2001). While VCs are usually quite secretive about their selection criteria, Baum and Silverman (2004) argue that they rely on *three* types of capital possessed by a venture: human, intellectual, and alliance capital. *Human* capital means the founder team capabilities, experience, and composition. High *Intellectual* capital exists if a venture possesses factors (e.g., technologies) that are innovative and difficult to replicate. *Alliance* capital means the extent and quality of a ventures network connections (e.g., research).

Interestingly, most research assumes that venture capital firms possess superior decision-making abilities and can detect high potential start-ups from the opportunity set. Research has rarely explored whether VCs do pick more successful start-ups (Baum and Silverman, 2004). With the advancement of venture capital research, doubts about the absolute superiority of VC-decisions have surfaced. Research has shown, for example, that VCs tend to over-rely on venture-team performance (Schefczyk and Gerpott, 2001; Baum and Silverman, 2004), are subject to natural behavioral biases such homophily (Cumming and Dai, 2010; Franke et al., 2006; Malmström et al., 2020) and suffer from over-confidence (Zacharakis and Shepherd, 2001).

The *coaching* lens research investigates if VCs improve venture performance beyond what would be expected without external support (Carnahan et al., 2010; Baum and Silverman, 2004; Megginson and Weiss, 1991). For example, Megginson and Weiss (1991) show that VC-backed start-ups tend to realize higher IPO prices than ventures with other financing sources. Edelmann (2002) finds that active management support by VCs increases firm performance. Hence, research suggests that VCs improve start-up performance through *coaching*. In contrast to VCs, crowds do usually not get involved in ventures. Most ICOs do not even sell equity in ICOs. Instead, they pre-sell future products in the form of utility tokens. Ventures may tap crowd wisdom to co-develop products, fostering innovativeness (Ritter, 1991; Hervé, 2016). Such processes work well if the crowd consists of future customers interested in using the firms' offerings (Hervé, 2016). In ICOs, the buzz and existence of secondary markets arguably lead to a predominance of financially rather than intrinsically motivated contributors. Hence, we argue that ICO crowds do not *coach* ventures.

6.3.2 Crowd investment and venture success

While measuring venture success has attracted significant attention from researchers, the majority of work has focused on campaign success - the ability of ventures to attract capital from the crowd (e.g., Amsden and Schweizer, 2018; Cordova et al., 2015; Cerchiello, 2018). The findings indicate that ICO investors employ decision cues comparable to VCs (Mollick, 2013). Nevertheless, they also show that ICO crowds irrationally herd after market sentiment (Ajaz and Kumar, 2018). Hence, mixed evidence for crowd decision-making effectiveness exists. Fundraising success may also be weakly correlated with venture success, allowing limited inferences to crowd decision quality.

The under-representation of post-campaign venture success research is attributable to *two* factors. The phenomenon's recency makes long-term performance tracking difficult, and the outcome's quality is less clear. Indeed, with the ICO activity peaking in 2018, most ICO-backed ventures are younger than *two* years. However, success may take years before it is measurable in conventional performance metrics (e.g., profit). Many ICO tokens are traded on secondary markets. In consequence, while for ICOs as well, the recency makes it difficult to measure performance objectively, it may be approximated by post-campaign trading return. Such performance data is not available for conventional crowdfunding.

Recent research in ICOs approaches venture success by investor return (Fisch, 2019; Benedetti and Kostovetsky, 2018; Lyandres et al., 2019). The results point to varying ICO performance: while the average ICO has earned a considerable return on secondary markets, most ICOs generate negative returns. On this grounds, it has been argued that the ICO market has developed into a market for lemons. In the absence of regulation and the presence of considerable buzz, low-quality firms self-select to pursue ICOs. High-quality ventures, under this hypothesis, avoid ICOs for reputational reasons. Research has also started to explore the effectiveness of ICO investor decision-making. Fisch and Momtaz (2020) study how institutional investor presence in an ICO campaign affects long-term success. They find that ICOs with institutional investment generate higher returns than other ICOs and attribute this finding to institutional investor *coaching*. The general crowds' decision-quality has not yet been investigated in the context of ICOs. We turn to the context of convectional crowdfunding for related work.

Mollick and Nanda (2016) were among the first to investigate crowdfunding post-campaign success by analyzing the success of *2,105* projects on the crowdfunding platform *Kickstarter*. They find that *90%* of successfully funded start-ups were still operating in 2014. Even though their observation period is relatively short, this result of impressive in comparison to regular VC-backed ventures estimated survival rates of *only 75%* Chernev (2020). Mollick and Nanda (2016) also investigate the quality of expert decision-making at the example of art crowdfunding projects and find no evidence for superior expert decisions. Greenstein and Zhu (2014) find that crowd-sourced content is less biased than expert content, attributable to crowd diversity leveraged during decision-making. Opponents to the hypothesis of effective crowd decision-making exist. In his opinion paper, Isenberg (2012) asserts that retail investors, hence crowds, cannot make wise investment decisions in crowdfunding. He explains that professional investors have significant experience and spend *50,000* US Dollar or more on a single due diligence to assess a ventures quality. In contrast, retail investors must rely on their (incomplete) set of (difficult to interpret) information. It is not surprising that they follow other investors, whom they assume to have superior knowledge, leading to irrational herding behavior of the overall crowd (Ajaz and Kumar, 2018; Simonsohn and Ariely, 2008).

The ICO context provides an ideal environment for studying crowd decision-making. Investment decisions are recorded on the blockchain in real-time and readily attributable to campaigns and investors. Furthermore, as discussed, ICOs bring together a diverse crowd, consequently providing decision-diversity. In this work, we study crowd decision effectiveness in ICOs, which, to our knowledge, has not been investigated before. We leverage blockchain data to classify investors based on their experience, enabling us to investigate whether some sub-crowds (experienced investors) form superior decisions to the general crowd. We measure post-campaign success based on returns to investors to link investor decisions with decision outcome.

6.4 Data and methodology

6.4.1 Data sample

Our analysis is based on *17* million blockchain transactions by *1,274* ICOs between 2015 and year-end 2018, which we have extracted from *ethplorer.io*. We refer to *section 3.1* for a detailed discussion of the dataset construction and cleaning steps. We enrich our dataset with campaign and venture attributes from *nine* ICO platforms, as detailed in *section 3.2.1*, reducing our set to *213* ICOs. A distinction between the fundraising and exchange-trading phase, which is crucial for our analyses, was not possible for 15 ICOs, reducing our set to 198 ICOs between 2017 and 2018. Last, the sample set was enriched with post-campaign success variables, as explained in the next section.

6.4.2 Variables

6.4.2.1 Dependent variables

Crowd composition In our first analysis, we analyze ICO crowd composition and employ *two* dependent variables: *Log novice share (Exchange)* and *Log experienced share (Exchange)*. Both variables are defined as the number of novices and experienced investors over the total number of investors per ICO during trading on crypto-exchanges. As crowd composition is constantly changing on crypto-exchanges, we need to define a cut-off date for our calculations, set to *31-12-2018*. Novice investors are defined as all first-time investors. Experienced investors are all investors who have carried out at least *nine* investments before their first investment in a given ICO[1]. We select share instead of the total number to measure differences in composition rather than the number of investors attracted by an ICO. Both variables are measured against fundraising phase pendants, as detailed in the next section.

ICO post-campaign success Our second analysis requires us to measure ICO post-campaign success, which is challenging. As quite precisely worded by (Mollick and Nanda, 2016, p.4), the *"nature of a correct answer is less clear"*, as compared to other contexts, for example *Wikipedia*. *Wikipedia* is about facts, start-up success, on the contrary, may take years to materialize, and even if a start-up has successfully acquired (many) customers, they may still be unable to operate profitably or generate revenue (Mollick and Nanda, 2016). Most ventures in our sample have not yet reached a point where success could be measured in financial performance, such as revenue or profit. We proxy start-up success by *two* variables with overlapping yet distinct explanatory power.

We measure *Survival* on the basis of *Twitter* activity, thereby following the approach of Benedetti and Kostovetsky (2018) and Fisch (2019). *90%* of ICOs in our sample have reported a *Twitter* link, making this variable widely available. Research shows that *Twitter* postings affect ICO campaign success (Fisch, 2019; Benedetti and Kostovetsky, 2018; Lee et al., 2019). Hence, most firms extensively post on *Twitter*. A start-up is defined as active if they have tweeted at least once in the past *six* months; otherwise, we assume it has ceased to exist.

Second, we approximate post-campaign success by the *buy-and-hold abnormal returns* (hereafter referred to as *BHAR*). *BHAR* measures the financial return to an investor who bought a token upon listing and held it for *six* months, adjusted

[1]Representing the 95^{th} quantile of investors by investment frequency in our data sample.

for the return of the market during the same holding period. *BHAR* has become a standard measure of long-term performance in IPO research (e.g. Ritter, 1991). It is particularly valuable in ICOs, where conventional performance measures are unavailable. We follow the example of prior ICO research to approximate post-campaign performance by *BHAR* (e.g., Benedetti and Kostovetsky, 2018; Fisch and Momtaz, 2020). In accordance with prior literature, we shorten the common holding period of *three* years to *six* months, as longer-term holding periods are usually unavailable in the context of ICOs. If we let i denote a specific campaign, P a tokens' trading price and *Mcap* the market capitalization, *BHAR* is defined as:

$$BHAR_i = \frac{P_{i,t=180} - P_{i,t=1}}{P_{i,t=1}} - \sum_{j=1 j \neq i}^{n} \frac{Mcap_{j,t=180}}{\sum_{j=1}^{n} Mcap_{j,t=180}} \times \frac{P_{j,t=180} - P_{j,t=1}}{P_{j,t=1}}$$

where $P_{i,t=1}$ is the listing price of a given ICO token and $P_{i,t=180}$ is the price after *six* months. We adjust the return of an individual ICO by the overall return of the market during the same period. The *market* consists of all tokens included in our data sample. We acknowledge that the true market portfolio is comprised of more tokens, of course. Nevertheless, it is reasonable to assume that our data sample replicates any general market trend, serving as a decent proxy for the overall market return, for the purpose of this analysis. We obtain market capitalization and price information from *coinmarketcap.com*. Of the *198* ICOs *49* are not listed on exchanges, reducing our sample size to *149* ICOs for all analyses employing BHAR as dependent variable.

6.4.2.2 Independent variables

Investor attraction We measure ICO crowd composition by *three* variables. *Log novice share (ICO)* and *Log experienced share (ICO)* measure the share of first-time and experienced investors (at least *nine* investments) over the total number of investors during the fundraising (ICO) phase. The *two* variables enable us to measure whether ICOs attract different investors crowds and if sub-crowds systematically select ICOs with higher post hoc performance. *Raised amount (ICO)* measures the US Dollar amount an ICO raised during fundraising and is included in all analyses. The total amount raised during fundraising affects post-campaign success independent of crowd decision-making capabilities, and we need to control for its effect.

ICO attributes We also measure ICO attributes to estimate their effect on ICO post-campaign success. Specifically, we include *seven* variables, *three* campaign and *four* venture attributes. The campaign attributes are *Pre ICO*, *Token price* and *Accept fiat*. *Pre ICO* specifies whether a venture conducted a prior token sale before the actual ICO. Pre-sales are used for demand testing and fine-tuning the ICO design. *Token price* specifies the US Dollar price one token was sold for during the ICO. If the *Token Price* varied during the ICO, we calculate the average price throughout fundraising. Last, *Accept fiat* specifies whether legal tender (fiat) was an accepted medium of payment during fundraising. Accepting fiat is generally interpreted as a negative signal by investors, making it easier for potential scammers to cash-out; our analyses enable us to understand whether this is the case.
We also include *four* venture attributes, characterizing the firm behind the ICO. *Blockchain industry* is a dummy variable taking the value of *one* if a ventures core business is related to blockchain technology. It serves as a proxy for technology focus and risk, as blockchain technology's market value is still uncertain. *Project status* specifies the development phase of the start-up and is classified along *four* stages: *Just an idea*, *MVP*[2], *Beta version* or *Working product*. Early-stage ventures are generally associated with a higher risk of failure, as their early-stage introduces higher outcome uncertainty. We also include *CEO experience* and *GitHub stars*, serving as proxies for human/ alliance and intellectual capital, respectively, as detailed in *section 3.2.1*. *CEO experience* measures the years of experience as CEO.

[2]Minimum viable product.

It captures the effects of both human and alliance capital. More experienced CEOs have a *proven* track record, therefore represent higher human capital. Experienced executives are also more likely to have established network ties. *GitHub stars* measures the number of stars on the coding repository *GitHub* and approximates the popularity of a ventures code among fellow developers. Accordingly, we argue that codes with a higher number of *GitHub stars* indicate higher intellectual capital, as it expresses project excitement and visibility within the developer community. Please revisit *section 3.2.1* and *section 5.4.2.2* for a detailed discussion of ICO attribute variable selection and interpretation.

6.4.2.3 Control variables

We control for *Duration*, *Quarter* and *Age* in all analyses to isolate the effects of investor decisions on post-campaign success. ICO activity has increased rapidly during our observation period. For this reasons, we need to control for systematic differences in periods, as approximated by the *Quarter*, which also includes the specific year. We also control for *Age*, measuring the campaign days since the ICO-start, accounting for life cycle-specific effects. Last, we control for ICO *Duration*. Some campaigns have raised funds within *two* days, while others have spread their fundraising over an entire year. This campaign design-choice affects investor dynamics, and we need to control for its effect. *Table 6.1* summarizes the variables employed in our analysis.

6.4.3 Empirical approach

Stability of crowd composition In our first analysis, we investigate whether ICOs attract systematically different investor crowds and if these crowds remain stable thorough their life cycle. This analysis may also help us understand, whether expert investors systematically purchase certain types of ICO tokens during primary stages and coordinate to sell to inexperienced investors (at higher prices) on

TABLE 6.1: Summary of regression variables

Variable	Description	Measurement
Dependent variables		
Log novice share (Exchange)	Share of first-time investors attracted by a venture while trading on crypto-exchanges.	Continuous
Log experienced share (Exchange)	Share of experienced investors (have invested at least 9 times) attracted by a venture while trading on crypto-exchanges.	Continuous
Survival	Specifies whether a venture is active based on *Twitter* activity. An ICO is defined as active if it has tweeted at least once within the past 6 months.	Binary
Log BHAR	ROI* to an investor who purchased a token upon listing and held it for 6 months, adjusted by the overall market return.	Continuous
Independent variables		
Log novice share (ICO)	Share of novice investors attracted by a venture during the fundraising phase (ICO).	Continuous
Log experienced share (ICO)	Share of experienced investors attracted by a venture during the fundraising phase (ICO).	Continuous
Raised amount (ICO)	Amount raised by a venture during the fundraising phase (ICO) in US Dollar.	Continuous
Pre ICO	Specifies whether an ICO has carried out a preliminary token sale before the actual ICO.	Dummy
Token price	Price of ICO tokens during fundraising in US Dollar.	Continuous
Accept fiat	Tokens can be purchased directly in exchange for fiat currency (legal tender) in the ICO.	Dummy
Blockchain industry	Specifies whether the ventures core business is related to blockchain technology.	Dummy
Project status	Founding stage of the venture: Just an idea (base level), MVP**, Beta version or Working product.	Categorical
CEO experience	Years of CEO experience	Continuous
GitHub stars	Measure of popularity of a project on GitHub, a widely known and used open-source coding repository (Github, 2020).	Integer
Control variables		
Duration	Fundraising (ICO) duration in days	Integer
Quarter	Year and quarter of ICO-start	Dummy
Age	Days since ICO-start	Integer

Note: The cut-off date for the calculation of Twitter activity was set to 12-08-2020. The number of investors on crypto-exchanges is based on transactions up to 31-12-2018. All shares are calculated over the total number of investors in the respective phase (ICO versus exchange).
**Return on investment.*
***Minimum viable product*

crypto-exchanges. Our analyses is inspired by Greenwood and Nagel (2009)'s research on the dynamics during stock market bubbles. We set up an OLS regression which looks as follows:

$$S_{i,Novice|Expert} = \alpha + \beta_1 s_{i,ICO,Novice} + \beta_2 s_{i,ICO,Expert} + \beta_3 y_i + \beta_4 \boldsymbol{C_i} + \beta_5 \boldsymbol{X_{i,t}} + \epsilon_i$$

where $S_{i,Novice|Expert}$ is the share of *Novice* | *Expert* investors attracted on crypto-exchanges (each analyzed in a separate regression), $s_{i,ICO,Novice}$ and $s_{i,ICO,Expert}$ measure the share of *Novice* | *Expert* investors during the fundraising phase (ICO), y_i is the total amount raised during fundraising (ICO). $\boldsymbol{C_i}$ is a vector of a *seven* ICO attributes. We control for *Duration*, *Quarter* and *Age*, represented by the vector $\boldsymbol{X_{i,t}}$ to account for the general market environment during the time of listing, as detailed in *section 6.4.2*.

ICO post-campaign success and crowd composition Our main analysis investigates whether experts make superior investment decisions. We also study how ICO attributes affect post-campaign success. As described in *section 6.4.2*, post-campaign success is measured by *Survival* and *BHAR*. We approach crowd composition by *Log novice share (ICO)* and *Log experienced share (ICO)*. We restrict our analysis to ICO investors, excluding investors on crypto-exchanges, as this would create a circular dependence biasing our results. Information on initial return and survival become available to investors on crypto-exchanges. Hence, their selection of tokens will be influenced by these information. In consequence, *Survival* and *Log BHAR* may influence the selection of investors. Our regression equation looks as follows:

$$VS_{proxy,i} = \alpha + \beta_1 s_{Novice,i} + \beta_2 s_{Expert,i} + \beta_3 y_i + \beta_4 \boldsymbol{C_i} + \beta_5 \boldsymbol{X_{i,t}} + \epsilon_i$$

where $VS_{proxy,i}$ describes the success proxy (*Survival* or *BHAR*), y_i is the capital raised during the ICO, $s_{Novice,i}$ and $s_{Expert,i}$ are the share of *Novice* | *Expert* investors respectively. Again, $\boldsymbol{C_i}$ and $\boldsymbol{X_i}$ are vectors of ICO campaign and control variables respectively.

6.5 Results

6.5.1 Descriptive results

The descriptive results for the *198* ICOs included in our analyses are shown in *table 6.2* and *table 6.3*. *Table 6.2* is divided along (post) campaign success, investor attraction and ICO attributes. We slightly deviate from the variable order in our regression analyses for the purpose of clarity. We have also not yet logged variables to make them intuitively understood.

The average ICO in our sample has raised *2* million US Dollar, yet, the median ICO has only collected *650* thousand US Dollar. This indicates that the high funding amounts are concentrated to a small number of mega ICOs, such as EOS[3]. Moreover, *55%* of the ventures in our sample are still active, based on their *Twitter* activity (*Survival*). While the failure rate of *45%* seems very high, it needs to be placed in context. According to research by Kotashev (2019), *90%* of all start-ups fail, *30%* have failed after their second year in business and *50%* have failed after their *5th* year. The picture looks similar for crowdfunding, the UK crowdfunding platform *Seedrs* analysed about *250* crowdfunding campaigns between *2012-16* and find that *40%* have failed (Prosser, 2016). While these findings from comparative contexts put the ICO survival rate in perspective, our findings suggest that the failure rate of ICOs is higher than that in comparable contexts.

The average *BHAR* is -19%, with an even higher median value of -26%, which

[3]EOS is a software company that has conducted the largest ICO in history, collecting *4* billion US Dollar during a year-long campaign that started in 2017 (Suberg, 2018).

TABLE 6.2: Descriptive results

Variable	Mean	Median	STD	Min	Max
(Post-) campaign success					
Survival	0.55	1	0.50	0	1
BHAR	-19%	-26%	183%	-678%	1,096%
Raised amount (ICO)	2,187,357	650,956	6,807,513	301	80,114,822
Number of investors	15,026	9,444	21,898	14	229,113
Investors (ICO)	5,109	1,866	11,380	13	92,943
Novice share	37%	32%	25%	1%	100%
Experienced share	32%	29%	21%	1%	89%
Investors (Exchange)	10,773	5,595	18,615	1	229,094
Novice share	33%	32%	20%	0%	100%
Experienced share	37%	34%	19%	6%	100%
ICO attributes					
Pre ICO	0.35	0	0.48	0	1
Token price	1.21	0.14	5.07	0	52.46
Accept fiat	0.09	0	0.29	0	1
Blockchain industry	0.16	0	0.36	0	1
Project status: MVP	0.19	0	0.39	0	1
Project status: Beta version	0.23	0	0.42	0	1
Project status: Working product	0.51	1	0.50	0	1
CEO experience	4.49	3	5.15	0	40
GitHub stars	80	10	211	0	1, 275
Duration	36	31	23	1	133

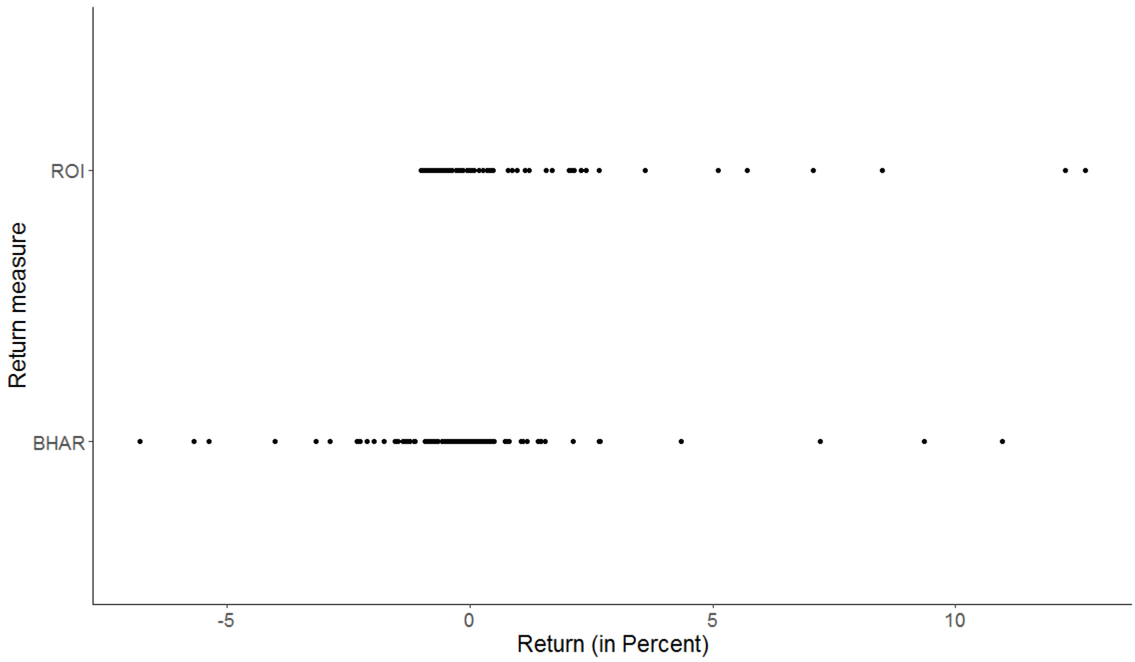
Note: Above table shows the descriptive statistics of all regression variables and the data sample of 198 ICOs. Amounts represent US Dollar amounts.

indicates that ICOs have destroyed overall wealth. *Figure 6.1* shows the *BHAR* and *ROI* distribution in our data sample. *ROI* is the raw return to investors over the *six* months period, and, in contrast to *BHAR*, not adjusted for the market return. The average *ROI* in our sample is *-1%* with a median of *-67%*. *Figure 6.1* shows how the majority of tokens have generated negative returns. A few extremely high performing tokens dominate the market capitalization, boosting mean values. While some investors have been fortunate enough to earn *1,000% (!)* return on their investments, the majority of investors has lost close to their total contribution.

The middle third of *table 5.2* shows the distribution of investors across the ICOs in our sample. The average ICO has attracted *15,026* investors, *72%* of them on crypto-exchanges. It follows that the majority of investors buy tokens on secondary markets. We do not know whether this is due to the restricted token supply during fundraising or a deliberate choice investors make. Again, we find that mean and median values deviate strongly, indicating high heterogeneity of individual campaigns' investor attraction ability. The share of novice and experienced investors remains relatively stable over both phases (fundraising versus crypto-exchanges). Interestingly, the share of both investor groups is, on average, balanced. Thus, neither subgroup seems to dominate the market. Nevertheless, the high negative correlation of both variables in *table 6.3* reveals, both groups tend to invest in different campaigns.

The lower third of *table 6.2* provides initial insights on ICO attributes. *35%* of ICOs have conducted a *Pre ICO*, meaning they sold a restricted number of tokens before the actual token sale. Such preliminary token sales have become

FIGURE 6.1: BHAR and ROI distribution in our data sample



Note: Above figure shows distribution of BHAR and ROI in our sample set. ROI is the raw market return of each token to an investor who bought the token upon listing and held is for six months. BHAR is the buy-and-hold abnormal return, the ROI adjusted for the overall market return (all tokens included in our dataset) during the same period.

common practice to test demand and fine-tune the actual ICO design. *Table 5.2* also indicates that *Token price* during the fundraising phase is quite variable. While the average price was *1.2* US Dollar per token, most ventures sold their token for much lower prices, with a median of only *14* cents. Without considering the number of tokens issued, we cannot draw any inferences about the underlying company valuation. It does, however, reveal different pricing strategies by ICOs. Only *9%* of campaigns *Accept fiat* currencies, legal tender issued by governments. The share of ventures with a blockchain industry focus is rather low with *16%*. It follows that most ventures attempting to raise capital via the blockchain, at least in our dataset, use it solely for fundraising. Also, the average project's status is surprisingly advanced, with *51%* of ventures claiming to have a functional product in place and 74% to have at least a *Beta version* available. Moreover, the average CEO has significant prior management experience with a mean value of *4.49* years. The average venture in our sample had *80* stars of *GitHub*, which is equivalent to followers, most of whom are programmers themselves. The number of *GitHub stars* may serve as a proxy for code-quality, project excitement, and visibility in the coding network. Last, the average ICO ended after *36* days. The results slightly differ from the ones presented in *chapter 5*, as our analyses are based on *198* instead of *213* ICOs in the former analysis.

Table 6.3 shows the correlation of all independent variables included in the regression. The project status variable pairs exhibit correlations of around *-0.5*, which makes sense as they belong to the same categorical variable (*Project status*). Nevertheless, most variable pairs show low correlations. Based on this, we are confident that our independent variable composition does not cause multicollinearity issues.

TABLE 6.3: Correlation of regression variables

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log novice share (ICO)	1											
Log experienced share (ICO)	(0.82)	1										
Raised amount (ICO)	0.01	0.09	1									
Pre ICO	(0.12)	0.14	0.01	1								
Token price	(0.03)	0.03	(0.05)	0.08	1							
Accept fiat	0.03	0.02	(0.02)	0.17	(0.03)	1						
Blockchain industry	(0.01)	0.04	(0.05)	(0.09)	(0.06)	(0.09)	1					
Project stage: MVP	(0.04)	0.04	(0.05)	0.02	0.06	0.07	0.11	1				
Project stage: Beta version	0.02	(0.09)	(0.08)	(0.06)	(0.10)	(0.01)	(0.07)	(0.26)	1			
Project stage: Working product	0.06	(0.04)	0.04	0.05	0.03	(0.04)	(0.02)	(0.49)	(0.56)	1		
CEO experience	0.06	(0.03)	0.06	(0.05)	(0.05)	0.28	0.00	(0.02)	(0.03)	0.08	1	
GitHub stars	0.10	(0.06)	(0.02)	(0.17)	(0.06)	(0.02)	0.24	0.08	(0.04)	0.01	(0.01)	1

Note: Above table shows the Pearson correlation of all independent regression variables. Brackets indicate negative correlations.

6.5.2 The stability of crowd composition

Table 6.4 shows the findings for our crowd composition analysis, investigating whether crowd composition remains stable over a ventures life cycle, as approximated by primary phase (ICO) and trading on crypto-exchanges. *Column 1* and *2* report our findings for *Log novice share (Exchange)*, the last *two* columns replicate the analyses for *Log experienced share (Exchange)*.

Our findings suggest that a large proportion of novice investors during fundraising (ICO) leads to a higher share of novice investors on exchanges. The same is true for experienced investors. Accordingly, our findings suggest that crowd composition remains stable between the ICO and trading on crypto-exchanges. Our findings also imply that ICOs attract systematically different crowds. Interestingly, this self-selection seems to be independent of the investigated ICO attributes. Indeed, we find no significant effects of ICO attributes on different investor groups' attraction on exchanges. This does not imply that investors do not employ these cues in decision formation, only that these ICO attributes do not lead to a systematically different crowd attraction by ICOs. Importantly, we only investigate the influence of fundraising phase attributes in this analysis while analyzing the investor attraction on crypto-exchanges. Additional information about individual firms may have become available in the meantime, which we have not accounted for in this analysis. Our findings show that significant differences exist between the investor crowd attracted by different ICOs. In consequence, it becomes relevant to ask whether crowd composition predicts ICO post-campaign success?

TABLE 6.4: Regression results: Stability of ICO crowd composition

	Dependent variable:			
	Log novice share (Exchange)		Log experienced share (Exchange)	
	(1)	(2)	(3)	(4)
Log novice share (ICO)	0.214**	0.205**	0.006	0.024
	(0.092)	(0.095)	(0.078)	(0.079)
Log experienced share (ICO)	−0.174	−0.175	0.397***	0.401***
	(0.114)	(0.117)	(0.095)	(0.098)
Raised amount (ICO)	0.000**	0.000**	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Pre ICO		−0.012		0.022
		(0.021)		(0.017)
Token price		0.001		0.002
		(0.002)		(0.002)
Accept fiat		0.009		−0.017
		(0.034)		(0.029)
Blockchain industry		−0.039		0.025
		(0.027)		(0.022)
Project status: MVP		0.029		−0.011
		(0.042)		(0.035)
Project status: Beta version		0.014		−0.021
		(0.041)		(0.034)
Project status: Working product		0.035		−0.034
		(0.038)		(0.032)
CEO experience		0.001		−0.001
		(0.002)		(0.002)
GitHub stars		−0.00002		0.00001
		(0.0001)		(0.00004)
Constant	184.105*	162.246	−146.678*	−120.844
	(103.310)	(106.454)	(86.429)	(88.015)
Control: Duration	Yes	Yes	Yes	Yes
Control: Quarter	Yes	Yes	Yes	Yes
Control: Age	Yes	Yes	Yes	Yes
Observations	190	190	194	194
Adjusted R^2	0.282	0.266	0.326	0.324

Note: Above table investigates the stability of ICO crowd composition during fundraising (ICO) versus trading on crypto-exchanges (Exchanges), controlling for ICO attributes. The full sample size is 198 ICOs, thereof 8 without novice investors and 4 without experienced investors, reducing our sample size to 190 and 194 respectively. MVP stands for minimum viable product.

6.5.3 ICO post-campaign success and crowd composition

Table 6.5 shows the results of our main analysis, investigating ICO sub-crowd decision-quality. *Column 1-3* employ *Survival* as dependent variable, *Column 4-6* replicate the analyses based on *BHAR* as post-campaign success measure.

We make *four* interesting observations. First, while most ICO attributes show no significant effects on ICO success, as indicated by the insignificant coefficients and low model fit, we find that more experienced CEOs (*CEO experience*) tend to be associated with higher survival rate. Nevertheless, they are not associated with higher abnormal returns to investors (*BHAR*). Similarly, *Accept fiat* lowers the chances of ICO survival. This finding is interesting. It reaffirms anecdotal evidence suggesting that ICOs accepting fiat currency during their ICO are associated with scams, as fiat currency makes it easier for scammers to cash-out.

Third, the total amount raised during the ICO *Raised amount (ICO)* (funding success) has no predictive power over *Survival* or *BHAR*. The majority of early ICO research has approximated ICO success by fundraising success in the absence of conventional performance metrics such as profit. Our findings suggest that fundraising success is a poor proxy for venture success.

Last, we find weak statistical evidence that experienced investor crowds *Log experienced share (ICO)* collectively select better performing ICOs in terms of *BHAR*. We want to test whether individual experienced investors re-confirm this finding. We replicate the analysis, employing the total number of experienced investors instead of the share. The results are shown in *table 6.6*. The coefficient turns insignificant in this analysis. Our findings imply that while crowds of experienced investors during fundraising lead to higher abnormal returns, the number of experienced investors per se does not.

TABLE 6.5: Regression results: Sub-crowd decision-quality in ICOs

	Dependent variable:					
		Survival			Log BHAR	
		normal			*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)
Log novice share (ICO)		−0.017	−0.099		0.149	0.171
		(0.357)	(0.361)		(0.156)	(0.163)
Log experienced share (ICO)		0.092	−0.010		0.307	0.345*
		(0.440)	(0.446)		(0.192)	(0.200)
Raised amount (ICO)	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pre ICO	0.075		0.071	0.001		−0.0002
	(0.079)		(0.080)	(0.037)		(0.036)
Token price	0.005		0.005	−0.001		−0.001
	(0.007)		(0.007)	(0.003)		(0.003)
Accept fiat	−0.223*		−0.221	0.003		−0.002
	(0.133)		(0.134)	(0.057)		(0.057)
Blockchain industry	0.017		0.017	0.011		0.010
	(0.103)		(0.103)	(0.045)		(0.045)
Project status: MVP	−0.111		−0.108	0.037		0.051
	(0.161)		(0.162)	(0.069)		(0.069)
Project status: Beta version	−0.104		−0.099	0.048		0.071
	(0.156)		(0.159)	(0.065)		(0.066)
Project status: Working product	0.013		0.018	−0.025		−0.010
	(0.147)		(0.148)	(0.061)		(0.062)
CEO experience	0.017**		0.017**	0.001		0.0003
	(0.007)		(0.007)	(0.003)		(0.003)
GitHub stars	0.0002		0.0002	0.00003		0.00001
	(0.0002)		(0.0002)	(0.0001)		(0.0001)
Constant	401.921	352.880	400.139	−31.797	60.593	50.021
	(388.862)	(401.434)	(405.041)	(175.823)	(175.832)	(181.251)
Control: Duration	Yes	Yes	Yes	Yes	Yes	Yes
Control: Quarter	Yes	Yes	Yes	Yes	Yes	Yes
Control: Age	Yes	Yes	Yes	Yes	Yes	Yes
Observations	198	198	198	149	149	149
Adjusted R^2				−0.042	−0.007	−0.032
Akaike Inf. Crit.	299.558	298.376	303.336			

Note: Above table shows the regression results investigating the decision-quality of ICO sub-crowds. Survival is measured on the basis of Twitter activity, BHAR is defined as the market adjusted return to investors. MVP stands for minimum viable product.

TABLE 6.6: Sensitivity analysis: Number of investors and decision-quality in ICOs

	Dependent variable:	
	Survival	Log BHAR
	normal	*OLS*
	(1)	(2)
Log novice investors (ICO)	−0.001	−0.015
	(0.025)	(0.012)
Log experienced investors (ICO)	0.025	0.011
	(0.025)	(0.012)
Raised amount (ICO)	−0.00000	0.00000
	(0.00000)	(0.00000)
Pre ICO	0.046	−0.001
	(0.080)	(0.037)
Token price	0.007	−0.001
	(0.007)	(0.003)
Accept fiat	−0.212	0.011
	(0.133)	(0.057)
Blockchain industry	0.029	0.012
	(0.104)	(0.046)
Project status: MVP	−0.092	0.048
	(0.162)	(0.070)
Project status: Beta version	−0.069	0.062
	(0.158)	(0.067)
Project status: Working product	0.043	−0.014
	(0.149)	(0.062)
CEO experience	0.016**	0.001
	(0.007)	(0.003)
GitHub stars	0.0002	0.00002
	(0.0002)	(0.0001)
Constant	418.218	−10.529
	(392.278)	(176.947)
Control: Duration	Yes	Yes
Control: Quarter	Yes	Yes
Control: Age	Yes	Yes
Observations	198	149
Adjusted R^2		−0.045
Akaike Inf. Crit.	303.287	

Note: Above table shows the regression results investigating the decision-quality of ICO sub-crowds. Survival is measured on the basis of Twitter activity, BHAR is defined as the market adjusted return to investors. MVP stands for minimum viable product. In contrast to the main analysis, we approximate crowd composition by the number of investors rather than share of the overall crowd.

6.6 Discussion and limitations

6.6.1 Discussion

Our findings suggest that *45%* of ICO-backed ventures have failed at the end of our observation period, meaning their first *three* or fewer years of business. While not statistically tested, qualitative comparison of evidence from comparative contexts suggest that ICO-backed ventures have a higher than average failure rate. For example, Kotashev (2019) finds that *30%* of start-ups have failed after their second year, independent of their funding method. Prosser (2016) reports the findings of the crowdfunding platform *Seedr* and shows that *40%* of ventures have failed within a *five* year of less time-horizon. Last, only *25%* of VC backed ventures fail in the long-run (Kotashev, 2019). Taken together with the high negative returns of the average ICO in our data sample, our findings point to the development of a market for lemons in the context of ICOs, the danger of which has been discussed by several ICO studies (e.g., Chod and Lyandres, 2018; Felix, 2018; Lee et al., 2019).
We show that ICOs attract systematically different investor crowds. Our findings are in line with Fisch (2019), and Boreiko and Risteski (2020), who demonstrate the existence of heterogeneous motives and behavior of ICO investors, as well as Wick and Ihl (2019), and Wallmeroth (2019), highlighting differences in decision-making and motives in conventional crowdfunding. Interestingly, our observation is not explained by the ICO attributes employed in this analysis. Therefore, our findings point to other forces' existence, not considered in this analysis, to drive investor decisions. For example, our analysis in *section 5.4.2.2* points to the consideration of social cues. This finding is also supported by other research in the context of ICOs (e.g., Ajaz and Kumar, 2018; Kallinterakis and Wang, 2019). It follows that further research is needed to explore the underlying causes of differential investor attraction of ICO campaigns.
Nevertheless, our findings allow for important inferences about the overall in-

vestor dynamics in the ICO market. Griffin et al. (2011) studied investor behavior during the dot-com bubble, an environment similar to ICOs today. They show that, in the presence of high investor heterogeneity and outcome uncertainty, informed investors deliberately exploited uninformed market participants. They fueled the upwards price spirals and coordinated to sell to uninformed investors at the price peak. The consequential loss of wealth was thus absorbed primarily by uninformed market participants. Our findings suggest that, in the context of ICOs, crowd composition remains stable on exchanges, with experienced and inexperienced investors investing in different ICO campaigns. This implies that informed market participants do not collectively purchase ICO tokens early (during fundraising) to resell at a profit to uninformed investors on crypto-exchanges on the overall market level.

We also approached the question of crowd decision-making quality in ICOs. Our findings indicate that the general crowd, as approximated by overall capital raised by an ICO during fundraising, does not succeed in channeling capital to prospering ventures. Our findings indicate that the *scouting* ability of experienced sub-crowds is weak and not generalizable to individual experts within the crowd, as approximated by experience. In this line of argumentation, our findings are consistent with Herzenstein et al. (2011)'s findings in the context of microloan-lending. We showed in *chapter 5* that ICO investors herd, yet, as our analyses here shows, the crowd does not collectively succeed to select higher-quality ventures. It follows that even experienced ICO investors are not able to distinguish high-quality ICOs better than tossing a coin would. Consequently, in this train of thought, our findings support the opinion of Isenberg (2012), who concludes that crowdfunding is a flawed concept, profiting ventures at the expense of investors. Nevertheless, we find weak evidence that experienced investors, collectively, make superior investment decisions to the general crowd and novices. Our findings complement the findings by Mollick and Nanda (2016), who shows that individual expert investors do not make superior decisions to individuals within the crowd. Our findings also extend the findings by Fisch and Momtaz (2020), who show that institutional investor participation leads to higher *BHAR*. They control for selection effects (*scouting*), arguing that the marginal performance improvement can thus be attributed to professional investor *coaching*. Our analysis is focused on selection and shows that experts, even within the crowd, possess slightly better decision-making capabilities than the general crowd. In this train of thought, the weak evidence could be interpreted as evidence that expert investors, without *coaching*, are worth little, as they alone do not significantly alter start-up success chances. Our finding point to the value of professional *coaching* by VCs. Further research is needed to shed light on this discussion.

6.6.2 Limitations

Our work is not without limitations. First, our base dataset has some missing values due to technical imperfections. This forced us to exclude observations from about *500* campaigns and *6* million investors, many of which inexperienced investors. While we do not expect missing observations to differ from other observations systematically, it is possible and may bias our results. Second, we approximate ICO post-campaign success by *Survival* and *BHAR*. While we feel comfortable that both measures provide valuable insights into the post-campaign success of ICOs, they are proxies with shortcomings. Especially *BHAR* has drawbacks, as it represents the collective investor option of a venture's value. This value may not accurately reflect the intrinsic value of a venture, as it is based on beliefs and expectations rather than facts. Third, our observation period is rather short. The most mature ICO in our sample has operated for *three* years at the end of our observation period and most ventures, while most ventures have operated for less than *two* years. Further research is needed to evaluate the long-run performance of ICOs.

Last, our investor classification is based on inferences drawn from each investor's transaction behavior, as no direct information on investors is available on any blockchain. We likely have incomplete information about an individual investor's behavior. Our data only includes transactions on *Ethereum*-blockchain, investors may use multiple IDs, and our sample does not include all ICOs conducted during the observation period. It follows that some investors may be missing or misclassified. Still, we feel comfortable that our measure of expertise, on average, is a decent investor classification proxy.

6.7 Conclusion

In this work, we have investigated the decision-making quality of ICO crowds. We have also analyzed whether sub-crowds, particularly experienced investors, form superior decisions to the general crowd and inexperienced market participants. Our empirical study is based on a set of *198* ICOs between 2017-18. We extract ownership information based on blockchain transactions and measure ICO post-campaign success by *two* proxies: *Survival*, based on a firms *Twitter* activity and *BHAR*, the buy-and-hold-abnormal return to investors.

Our findings indicate that ICO crowds do not channel capital to prospering, i.e., higher-performing ventures. In contrast, we find weak evidence that experienced investor crowds possess superior decision-making capabilities by channeling capital to firms with higher subsequent *BHAR*. Nevertheless, as our analyses show, these results are not generalizable to individual investors. Our results are in line with relevant literature from ICOs (e.g., Fisch and Momtaz, 2020; Benedetti and Kostovetsky, 2018; Lyandres et al., 2019) and comparable contexts (e.g., Herzenstein et al., 2011; Mollick and Nanda, 2016). They show that crowds in high uncertainty contexts cannot perform the *scouting* role of selecting high-prospect start-ups from the crowd of capital seeking ones, traditionally performed by professional investors such as venture capital firms. We contribute a perspective on crowd decision-making in ICOs.

We show that ICOs attract systematically different investor crowds, not explained by ICO attributes. Further research is required to decompose the crowd selection-process, which has important practical implications for ventures pursuing an ICO. We also show that *45%* of ICOs in our sample have failed within our observation period, and most ICOs generate high negative returns. It follows that the ICO market may have developed into a market for lemons, with important implications for investors and regulators alike. While our findings should heighten investors' cautiousness, we hope they also urge regulators to set boundaries; otherwise, the phenomenon of ICOs will vanish.

Chapter 7

Conclusion

7.1 Synthesis of empirical findings

This dissertation has investigated investors in Initial Coin Offerings, thereby being guided by my central research question: *Who invests in ICOs and how do market participants form decisions in this novel environment, characterized by buzz and high uncertainty?* Motivated by the restricted knowledge that exists on ICO investors, I have attempted to investigate the phenomenon holistically across *three* research studies.

I initiated with an explorative analysis aiming to understand the crowd composition and the degree of investor heterogeneity in the ICO market. The results indicate pronounced heterogeneity: while the majority of the crowd consists of novice investors making cautious, sporadic contributions, *75%* of the capital is contributed by only *3%* of the investors. Employing cluster analysis, I identify *three* distinct behavioral investor clusters. *Crypto Enthusiasts*, who scan the ICO market for opportunities and actively manage their portfolios by diversifying and trading on exchanges. *Risk-avoidant Rationalists*, who concentrate their investment in a few campaigns that have gained market traction to minimize risk and realize short-term profits. Last, *Passive Investors* following market trends and tend to

invest for the longer-term.
I also analyzed how individuals time their investments concerning social media sentiment. While the average market participant is more likely to invest in positive market sentiment times, significant differences exist between investor groups. Experienced investors show higher sensitivity and quicker reactions to market sentiment, while inexperienced investors tend to follow.

Next, I set out to investigate how investors interact with each other, hence, understanding ICO crowd dynamics. I analyzed to what extent herding effects are present in ICOs and whether market participants herd irrationally or strategically reinterpret cues in the light of new information (rational herding).
The findings point to pronounced herding in ICOs. I also show that herding is primarily irrational, as investors blindly mimic others' behavior, whom they believe to possess private information or superior investment analysis capabilities. Nevertheless, these findings also pinpoint the value of disentangling the investor crowd by showing that ICO crowds' herding rationality depends upon investor experience and the observed fundraising phase. Specifically, I show that irrational herding is moderated by investor experience. Novice investors herd after a general market opinion. This tendency disappears as investors gain experience, and they learn to observe others to aid their decision-making strategically. My findings also indicate that herding is more pronounced on crypto-exchanges. I find evidence for an increased tendency to herd rationally across investor groups, indicating the value of secondary markets as an opportunity to resolve information asymmetry. However, I also show that inexperienced investors become more susceptible to sentiment-induced, irrational herding.

In my third research study, I asked whether ICOs benefit society as a whole by channeling capital to prosperous ventures. I investigated ICO post-campaign success and the crowd's ability to identify prospering ventures. I find that *45%* of ICO-backed ventures have failed within the observation period. This failure

rate is significantly higher than the overall start-up space, hence pointing to the development of a market for lemons (Chernev, 2020). I also show that crowds do not succeed in channeling capital to prospering (higher performing) ventures.
My results show that crowds in high uncertainty contexts cannot perform the *scouting* role of selecting high-prospect start-ups from the crowd of capital seeking ventures, traditionally performed by professional investors such as venture capital firms. Therefore, they highlight the value of professional investors and the complementary (not substituting) role of crowdfunding.

My findings paint a dramatic picture of the ICO market. The crowd in ICOs is composed mainly of inexperienced investors, who make small contributions each and usually do not return for a second investment within my observation period. A small number of experienced investors provides the majority of capital. As such, fundraising success crucially depends on a small number of individuals. One could argue, consequently, that ICOs do not democratize start-up financing. Instead, retail investor contributions and the campaign as such are a vehicle for ventures to raise visibility. At the same time, their capital-backing remains in the hands of a small, privileged club of investors. It follows that ICOs shift, yet to not spread, decision-making power.
Furthermore, past episodes of speculative bubbles were characterized by a shift in crowd composition to inexperienced investors. In the past, such compositions have triggered a dynamic in which informed market participants (professional investors) exploited uninformed investors (e.g., Greenwood and Nagel, 2009; Griffin et al., 2011; Binglin et al., 2017). My findings point to a similar development in the ICO market: novices dominate the market and, in the presence of high decision-making uncertainty, turn to social cues to aid their decision formation, making them vulnerable to manipulation. Indeed, 2019 developments point to the burst of a speculative bubble. As my analyses showed, almost *70%* of value was lost by the median ICO, and 2019 saw a drastic decline in market activity of close

to *100%* (Statista, 2020).

My research is inconclusive about the beneficiaries of the ICO-boom. Throughout this work, the analyses support the argumentation of Isenberg (2012), who argues that ventures are the net winner of the crowdfunding phenomenon. Evidence shows that ventures raise more capital in ICOs than they could with alternative funding methods (Fundera.com, 2020). Yet, ventures do not have to trade equity for capital in ICOs. My dataset also indicates that the average venture-quality is lower in ICOs, as compared to the overall start-up space[1].

My research suggests that investors, on average, lose wealth by investing in ICOs. As the analyses in *chapter 6* showed, not even experienced investors within the ICO crowd can consistently distinguish high-performing ICOs. Consequently, while individual investors may have benefited from ICO investing, overall wealth has been destroyed by ICOs thus far.

The implications of my findings could follow either of *two* alternative lines of argumentation. First, ICOs should be prohibited altogether, as they do not benefit society as a whole but rather shift unequally distributed power from financial intermediaries to ventures and potentially privileged investors. (Low quality) ventures obtain *risk-free* capital, and a few informed market participants profit from temporary arbitrage opportunities.

Alternatively, one could argue that ICOs may benefit society by democratizing venture-financing if regulation achieves what the free market forces have not yet accomplished: control ICOing ventures' quality and protect the retail investor wealth. Further research is needed to understand which of the *two* alternative interpretations is the predominant force and conclude the debate on Initial Coin Offerings' sustainability.

[1] As suggested by qualitatively comparing the failure rate of ICOs in the data sample to overall start-up failure rates, irrespective of the financing method (Chernev, 2020).

7.2 Future research directions

With the rapid pace at which ICO research is picking up, it would undoubtedly help to comprehend the state of research on ICOs and systematically point to future research directions. The research recommendations pointed out here are based on the discoveries during my research, thus not systematic. Moreover, the recommendations in this section complement the research directions provided in each chapter.

While this work, to my knowledge, is among the first empirical investigations of ICO post-campaign success, we are, due to the recency of the phenomenon, analyzing a relatively short observation period of around *two* years. Further research is needed to evaluate the ICO-backed ventures' long-term success. A longer-termed investigation would allow for alternative success measures looking beyond capital market valuation to more conventional success measures (e.g., profit, time-to-market). Further research is also needed to understand how ventures self-select to the ICO market, thus shed light on the causes for the lower performance observed for ICO-backed ventures.

Furthermore, although qualitatively debated in this work, further research is needed to understand secondary-market trading implications on ventures. To assess ICO-financing sustainability, we need to understand whether the price fluctuations on exchanges negatively affect a venture's development. For example, it may reduce innovativeness, lead to myopic behavior, or reduce managerial willingness to take risks - all vital factors to business success.

Last, the choice of financing may also negatively affect longer-run venture success. Further research is needed to understand how ICOs affect, e.g., the ability of ventures to secure additional funding or how the general reputation and network ties of ventures are affected by the financing method's choice. Research on both is crucial to understand the sustainability of ICOs as means for early-venture financing.

Bibliography

Abdi, H. and Williams, L. J. 2010. Principal component analysis. ***Wiley Interdiscip. Rev. Comp. Stat.***, 2: 433–459.

Abrams, E. 2017. Securities Crowdfunding: More than Family, Friends, and Fools? Available at SSRN: https://ssrn.com/abstract=2902217.

Adhami, S., Giudici, G., and Martinazzi, S. 2018. Why do businesses go crypto? An empirical analysis of initial coin offerings. ***Journal of Economics and Business***, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3046209.

Ajaz, T. and Kumar, A. S. 2018. Herding in Crypto-Currency Markets. ***Annals of Financial Economics***, 13(02): 1–15.

Amsden, R. and Schweizer, D. 2018. Are Blockchain Crowdsales the New 'Gold Rush'? Success Determinants of Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3163849.

An, J., Hou, W., and Liu, X. 2018. Initial Coin Offerings: Investor Protection and Disclosure. ***Proceedings of Financial Management Association Annual Meeting, San Diego***. Available at http://fmaconferences.org/SanDiego/Papers/Initial_Coin_Offerings,_Investor_Potection_and_Disclosure.pdf.

Anson, M. 2018. Initial Coin Offerings: Economic Reality or Virtual Economics? ***Journal of Private Equity***, 21(4): 41–52.

Arora, P. and Varshney, S. 2016. Analysis of K-Means and K-Medoids Algorithm For Big Data. ***Procedia - Procedia Computer Science***, 78(December 2015): 507–512.

Avery, C. and Zemsky, P. 1998. Multidimensional Uncertainty and Herd Behavior in Financial Markets. ***The American Economic Review***, 88(4): 724–748.

Baum, J. A. and Silverman, B. S. 2004. Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. ***Journal of Business Venturing***, 19(3): 411–436.

Benedetti, H. and Kostovetsky, L. 2018. Digital Tulips? Returns to Investors in Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3182169.

Berkhin, P. 2002. ***A Survey of Clustering Data Mining. In: Kogan J., Nicholas C., Teboulle M. (eds) Grouping Multidimensional Data.*** Springer, Berlin, Heidelberg.

Bikhchandani, S., Hirshleifer, D., and Welch, I. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. ***Journal of Political Economy***, 100(5): 992–1026.

Binglin, G., Deng, P., and Donghui, S. 2017. New investors and bubbles: An analysis of the baosteel call warrant bubble. ***Management Science***, 63: 2493–2508.

Black, A. and Fraser, P. 2002. Stock market short-termism - an international perspective. ***Journal of Multinational Financial Management***, 12: 135–158.

Blaseg, D. 2018. Dynamics of voluntary disclosure in the unregulated market for Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3207641.

Bolton, P. 2006. Executive compensation and short-termist behaviour in speculative markets. ***Review of Economic Studies***, 73(3): 577–610.

Boreiko, D. and Risteski, D. 2020. Serial and large investors in Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3318781.

Bourveau, T., De George, E. T., Ellahie, A., and Macciocchi, D. 2018. Initial Coin Offerings: Early Evidence on the Role of Disclosure in the Unregulated Crypto Market. ***Working paper Columbia Business School***. Available at SSRN:https://ssrn.com/abstract=3193392.

Brennan, M. J. 2004. How Did It Happen? ***Economic Notes***, 33(1): 3–22.

Burago, D., Burago, Y. D., and Ivanov, S. 2001. ***A course in metric geometry, Graduate Studies in Mathematics***, volume 33. American Mathematical Society.

Carnahan, S., Agarwal, R., and Campbell, B. 2010. The Effect of Firm Compensation Structures on the Mobility and Entrepreneurship of Extreme Performers. ***Business***, 801(January): 1–43.

Catalini, C. and Gans, J. S. 2018. Initial Coin Offerings and the Value of Crypto Tokens. ***MIT Sloan Research Paper No. 5347-18, Rotman School of Management Working Paper No. 3137213***. Available at SSRN: https://ssrn.com/abstract=3137213.

Cerchiello, Paola Toma, A. M. 2018. DEM Working Paper Series ICOs success drivers: a textual and statistical analysis ICOs success drivers. ***DEM Working Papers Series 164, University of Pavia, Department of Economics and Management***. Available at https://ideas.repec.org/s/pav/demwpp.html.

Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. 2014. NbClust: An R package for determining the relevant number of clusters in a data set. ***Journal of Statistical Software***, 61(6): 1–36.

Chen, D. and Lin, Z. 2014. Rational or irrational herding in online microloan markets: Evidence from china. ***SSRN Electronic Journal***. Doi: 10.2139/SSRN.2425047.

Chernev, B. 2020, What percentage of startups fail? [67+ stats for 2020], https://review42.com/what-percentage-of-startups-fail/, accessed on 2020-10-06.

Chiang, Y.-m., Hirshleifer, D., and Sherman, A. E. 2011. Do Investors Learn from Experience? Evidence from Frequent IPO Investors. ***Review of Financial Studies***, 24(5): 1560–1589.

Chiu, I. H.-Y. 2018. Decoupling tokens from trading: reaching beyond investment regulation for regulatory policy in initial coin offerings. ***International Business Law Journal/ Revue de Droit des Affaires Internationales***, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3156277.

Chod, J. and Lyandres, E. 2018. A Theory of ICOs: Diversification, Agency, and Information Asymmetry. ***Management Science***, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3159528.

Clarida, R. H., Kraay, A., and Ventura, J. 2013. The Dot-Com Bubble, the Bush Deficits, and the U.S. Current Account. ***NBER working papers***, 11,543: 457–496.

Cohen, A. 2020, Python package fuzzywuzzy 0.18.0, https://pypi.org/project/fuzzywuzzy/, accessed on 2019-12-20.

CoinCheckup 2020, Coincheckup - the ico research platform, https://ico.coincheckup.com/, accessed on 2020-07-22.

coinmarketcap 2021, Today's cryptocurrency prices by market cap, https://coinmarketcap.com/, accessed on 2021-01-08.

Cordova, A., Dolci, J., and Gianfrate, G. 2015. The Determinants of Crowdfunding Success: Evidence from Technology Projects. ***Procedia - Social and Behavioral Sciences***, 181: 115–124.

Crowdfundinsider 2019, Initial coin offerings are dead for now, https://www.crowdfundinsider.com/2019/04/145920-report-initial-coin-offerings-are-dead-for-now/, accessed on 2020-01-13.

Cumming, D. and Dai, N. 2010. Local bias in venture capital investments. ***Journal of Empirical Finance***, 17(3): 362–380.

Cumming, D. and Johan, S. 2013. Demand-driven securities regulation: evidence from crowdfunding. ***Venture Capital***, 15(4): 361–379.

Datanovia 2019, Determining the optimal number of clusters, https://www.datanovia.com/en/lessons/determining-the-optimal-number-of-clusters-3-must-know-methods/, accessed on 2020-07-01.

Dell'Erba, M. 2018. The implementation of the "do no harm" approach: From inactivity to full enforcement. The case of ICOs. ***Michigan Telecommunications and Technology Law Review***, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3194863.

Devenow, A. and Welch, I. 1996. Rational Herding in Financial Economics. ***European Economic Review***, 40(3-5): 603–615.

Dubuisson, M.-P. and Jain, A. K. 1994. A modified hausdorff distance for object matching. ***Proceedings of 12th International Conference on Pattern Recognition, Jerusalem, Israel***, 1: 566–568. doi: 10.1109/ICPR.1994.576361.

Edelmann, L. 2002. German venture-capital firms and portfolio-company performance: what types of management support make a difference? ***Academy of Management Executive***, 1(16): 156–157.

European Parliament 2020, Listing of tax havens by the eu, https://www.europarl.europa.eu/cmsdata/147404/7%20-%2001%20EPRS-Briefing-621872-Listing-tax-havens-by-the-EU-FINAL.PDF, accessed on 2020-07-24.

Felix, T. 2018. Underpricing in the Cryptocurrency World: Evidence from Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3202320.

Fisch, C. 2019. Initial coin offerings (ICOs) to finance new ventures. ***Journal of Business Venturing***, 34(1): 1–22.

Fisch, C. and Momtaz, P. P. 2020. Institutional investors and post-ICO performance: an empirical analysis of investor returns in initial coin offerings (ICOs). ***Journal of Corporate Finance***, 64(June): 101679.

Franke, N., Gruber, M., Harhoff, D., and Henkel, J. 2006. What you are is what you like-similarity biases in venture capitalists' evaluations of start-up teams. ***Journal of Business Venturing***, 21(6): 802–826.

Fried, V. H. and Hisrich, R. D. 1994. Toward a Model of Venture Capital Investment Decision Making. ***Financial Management***, 23(3): 28–37.

Fromberger, M. and Haffke, L. 2020. ICO Market Report 2018/2019 – Performance Analysis of 2018's Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3512125.

Fundera.com 2020, Crowdfunding statistics (2020): Market size and growth, https://www.fundera.com/resources/crowdfunding-statistics, accessed on 2020-10-22.

Gage, D. 2012, The venture capital secret: 3 out of 4 start-ups fail, https://www.wsj.com/articles/SB10000872396390443720204578004980476429190, accessed on 2020-09-29.

Github 2020, Where the world builds software, https://github.com, accessed on 2021-01-05.

Giudici, G. and Adhami, S. 2019. The impact of governance signals on ICO fundraising success. ***Journal of Industrial and Business Economics***, 46(2): 283–312.

Goldberg, A., Hannan, M. T., and Kovács, B. 2015. What does it mean to span cultural boundaries? Variety and atypicality in cultural consumption. ***American Sociological Review***, 81(2): 215–241.

Greenstein, S. and Zhu, F. 2014. Do experts or collective intelligence write with more bias? Evidence from Encyclopædia Britannica and Wikipedia. ***Havard Business School Working Paper 15-023***. Available at https://www.hbs.edu/faculty/Publication%20Files/15-023_e044cf50-f621-4759-a827-e9a3bf8920c0.pdf.

Greenwood, R. and Nagel, S. 2009. Inexperienced investors and bubbles. ***Journal of Financial Economics***, 93(2): 239–258.

Griffin, J. M., Harris, J. H., Shu, T. A. O., and Topaloglu 2011. Who drove and burst the tech bubble? ***The Journal of Finance***, 66(4): 1251–1290.

Griffin, J. M. and Shams, A. 2018. Is Bitcoin Really Un-Tethered? Available at SSRN: https://ssrn.com/abstract=3195066.

Gurdgiev, C. and O'Loughlin, D. 2020. Herding and anchoring in cryptocurrency markets: Investor reaction to fear and uncertainty. ***Journal of Behavioral and Experimental Finance***, 25(100271).

Hartigan, J. A. and Wong, M. A. 2012. Algorithm AS 136 A K-Means Clustering Algorithm. ***Journal of the Royal Statistical Society. Series C (Applied Statistics)***, 28(1): 100–108.

Hervé, F. 2016. Investor Motivations in Investment-Based Crowdfunding Investor Motivations in Investment-Based Crowdfunding. ***European Conference on Information Systems (ECIS) (accepted for publication), Tel Aviv, Israel***. Available at http://pubs.wi-kassel.de/wp-content/uploads/2015/04/JML_5321.pdf.

Herzenstein, M., Dholakia, U. M., and Andrews, R. L. 2011. Strategic Herding Behavior in Peer-to-Peer Loan Auctions. ***Journal of Interactive Marketing***, 25(1): 27–36.

Hornuf, L. and Schwienbacher, A. 2017. Should securities regulation promote equity crowdfunding? ***Small Business Economics***, 49(3): 579–593.

Hornuf, L. and Schwienbacher, A. 2018. Market mechanisms and funding dynamics in equity crowdfunding. ***Journal of Corporate Finance***, 50: 556 – 574.

Hui, J. S., Greenberg, M. D., and Gerber, E. M. 2014. Understanding the role of community in crowdfunding work. ***Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14***, 62–74.

icorating.com 2017, Ripio, https://icorating.com/ico/ripio-/, accessed on 2020-11-09.

Isenberg, D. 2012, The Road to Crowdfunding Hell - Daniel Isenberg, https://hbr.org/2012/04/the-road-to-crowdfunding-hell, accessed on 2020-09-22.

Jain, A. K. 2010. Data clustering : 50 years beyond K-means. ***Pattern Recognition Letters***, 31(8): 651–666.

Janis, I. 1991. Groupthink in A first look at communication theory, 235–246. New York: McGrawHill.

Jolliffe, I. T., Cadima, J., and Cadima, J. 2016. Principal component analysis: a review and recent developments Subject Areas. ***Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences***, 374(2065).

Kallinterakis, V. and Wang, Y. 2019. Do investors herd in cryptocurrencies – and why? ***Research in International Business and Finance***, 50(January): 240–245.

Kher, R., Terjesen, S., and Liu, C. 2020. Blockchain, Bitcoin, and ICOs: a review and research agenda. ***Small Business Economics***. Available at https://doi.org/10.1007/s11187-019-00286-y.

Kim, K. and Viswanathan, S. 2018. The Signals in the Noise: The Role of Reputable Investors in a Crowdfunding Market. ***SSRN Electronic Journal***. doi:10.2139/ssrn.2258243 .

Kotashev, K. 2019, Startup failure rate: Ultimate report + infographic [2020], https://www.failory.com/blog/startup-failure-rate, accessed on 2020-09-29.

Kuppuswamy, V. and Bayus, B. L. 2015. Crowdfunding creative ideas: the dynamics of projects backers in Kickstarter. Forthcoming. Available at SSRN: https://ssrn.com/abstract=2234765.

Kuppuswamy, V. and Bayus, B. L. 2018. Crowdfunding Creative Ideas: The dynamics of project backers. The Economics of Crowdfunding, 151–182.

Lee, J., Li, T., and Shin, D. 2019. The Wisdom of Crowds and Information Cascades in FinTech: Evidence from Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3195877.

Lee, L. 2016. New Kids on the Blockchain: How Bitcoin's Technology Could Reinvent the Stock Market. ***Hastings Business Law Journal***, 12(2).

Leone, A. J., Rock, S., and Willenborg, M. 2007. Disclosure of intended use of proceeds and underpricing in initial public offerings. ***Journal of Accounting Research***, 45(1): 111–153.

Li, X., Tang, Y., Yang, N., Ren, R., Zheng, H., and Zhou, H. 2016. The value of information disclosure and lead investor in equity-based crowdfunding: An exploratory empirical study. ***Nankai Business Review International Article information***, 7(3).

Lin, M., Prabhala, N. R., and 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.

Lin, M., Sias, R., and Wei, Z. 2015. “Smart Money”: Institutional Investors in Online Crowdfunding. ***Proceedings of 2015 INFORMS Conference on Information Systems and Technology, University of California, Berkeley***. Available at http://www.fmaconferences.org/Boston/II_LSW.pdf.

Lin, Y., Boh, W., and Goh, K. 2014. Examining Archetypes of Crowdfunders and their choice of projects. ***Academy of Management Proceedings***, 2014(1).

Liu, H. and Wang, Y. 2018. The Value of Crowdfunding: An Explanation Based on Demand Uncertainty and Comparison with Venture Capital. ***Emerging Markets Finance and Trade***, 54(4): 783–791.

Lyandres, E., Palazzo, B., and Rabetti, D. 2019. Are Tokens Securities? An Anatomy of Initial Coin Offerings. ***SSRN Electronic Journal***, (April).

Mackay, C. 2008. Memoirs of Extraordinary Popular Delusions and the Madness of Crowds. Richard Bentley.

Malmström, M., Voitkane, A., Johansson, J., and Wincent, J. 2020. What do they think and what do they say? Gender bias, entrepreneurial attitude in writing and venture capitalists’ funding decisions. ***Journal of Business Venturing Insights***, 13(December 2019).

Markowitz, H. 1939. Portfolio selection. ***The Journal of Finance***, 7(1): 77–91.

Megginson, W., L. and Weiss, K. 1991. Venture Capitalist Certification in Initial Public Offerings. ***The Journal of Finance***, 46(3): 879–903.

Mendelson, M. 2019. From Initial Coin Offerings to Security Tokens: A U.S. Federal Securities Law Analysis. ***HeinOnline***, 52(1): 52–94.

Merre, R. 2020, ICO 101 — History of Initial Coin Offerings (ICOs), https://medium.com/hackernoon/ico-101-history-of-initial-coin-offerings-icos-part-1-from-mastercoin-to-ethereum-4689b7c2326b, accessed on 2020-10-29.

Mochkabadi, K. and Volkmann, C. K. 2018. Equity crowdfunding: a systematic review of the literature. ***Small Business Economics***, 2018 edition: 9–78.

Mollick, E. 2014. The dynamics of crowdfunding: An exploratory study. ***Journal of Business Venturing***, 29(1): 1–16.

Mollick, E. and Nanda, R. 2016. Wisdom or madness? Comparing crowds with expert evaluation in funding the arts. ***Management Science***, 62(6): 1533–1553.

Mollick, Ethan, R. 2013. Swept away by the crowd? Crowdfunding, venture capital and the selection of entrepreneurs. Available at SSRN: https://ssrn.com/abstract=2239204.

Momtaz, P. P. 2018a. Initial Coin Offerings. ***PLOS ONE***. Available at SSRN: https://ssrn.com/abstract=3166709.

Momtaz, P. P. 2018b. Putting Numbers on the Coins: The Pricing and Performance of Initial Coin Offerings. Available at SSRN: https://ssrn.com/abstract=3680858.

Moritz, A. and Block, J. H. 2016. Crowdfunding : A literature review and research directions. In ***Brüntje D., Gajda O. (eds) Crowdfunding in Europe. FGF Studies in Small Business and Entrepreneurship***, 25–53. Springer, Cham. Available at https://doi.org/10.1007/978-3-319-18017-5_3.

Morris, S. 1996. Speculative Investor Behavior and Learning. ***The Quarterly Journal of Economics***, 111(4): 1111–1133.

Nakamoto, S. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. ***Whitepaper***. Available at https://bitcoin.org/bitcoin.pdf.

Neupane, S., Paudyal, K., and Thapa, C. 2014. Firm quality or market sentiment: What matters more for IPO investors? ***Journal of Banking and Finance***, 44(1): 207–218.

OECE 2020, Venture capital investments, https://stats.oecd.org/Index.aspx?DataSetCode=VC_INVEST, accessed on 2020-10-26.

Parrino, R. J. and Greenslade, K. K. 2014. Tweeting corporate disclosures: SEC staff issues guidance permitting hyperlinks to cautionary statements and legends in social media communications. ***Journal of Investment Compliance***, 15(3): 28–31.

Penke, M. 2018, Rekord-ico sammelt 4 milliarden dollar ein – für ein besseres android, https://www.gruenderszene.de/fintech/blockone-eos-ico-rekord-4-milliarden-dollar-ethereum, accessed on 2020-10-22.

pnswire 2019, The Global Crowdfunding Market was Valued at 10.2 Billion US$ in 2018 and is Expected to Reach 28.8 Billion US$ with a CAGR of 16% by 2025 - Valuates Reports, https://reports.valuates.com/sreport/QYRE-Auto-1598/Global$_$Crowdfunding$_$Market, accessed on 2020-10-26.

Prosser, D. 2016, Crowdfunding delivers a 40% return - and a 40% failure rate, https://www.forbes.com/sites/davidprosser/2016/09/13/crowdfunding-delivers-a-40-return-and-a-40-failure-rate/#3462afc63473, accessed on 2020-09-29.

Rhue, L. 2018. Trust is All You Need: An Empirical Exploration of Initial Coin Offerings (ICOs) and ICO Reputation Scores. Available at SSRN: https://ssrn.com/abstract=3179723.

Ritter, J. R. 1991. The Long Run Performance of initial Public Offerings. ***The Journal of Finance***, 46(1): 3–27.

Rousseeuw, P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. ***Journal of Computational and Applied Mathematics***, 20: 53–65.

Schefczyk, M. and Gerpott, T. J. 2001. Management Support for Portfolio Companies of Venture Capital Firms: An Empirical Study of German Venture Capital Investments. ***British Journal of Management***, 12(3): 201–216.

Scheinkman, J. A. and Xiong, W. 2003. Overconfidence and speculative bubbles. ***Journal of Political Economy***, 111(6): 1183–1219.

SEC 2015, Regulation crowdfunding: A small entity compliance guide for issuers, https://www.sec.gov/info/smallbus/secg/rccomplianceguide-051316.html, accessed on 2020-10-26.

Sherman, N. 2018. A behavioral economics approach to regulating Initial Coin Offerings. ***Georgetown Law Journal***, 107(1990): 17–35.

Simonsohn, U. and Ariely, D. 2008. When rational sellers face nonrational buyers: Evidence from herding on eBay. ***Management Science***, 54(9): 1624–1637.

Statista 2020, Worldwide amount crytocurrency ico projects, https://www.statista.com/statistics/804748/worldwide-amount-crytocurrency-ico-projects/, accessed on 2020-10-29.

Streletzki, J. G. and Schulte, R. 2013. Start-up teams and venture capital exit performance in Germany: venture capital firms are not selecting on the right criteria. ***Journal of Small Business and Entrepreneurship***, 26(6): 601–622.

Suberg, W. 2018, Eos about to secure a record $4 bln in year-long ico, https://cointelegraph.com/news/eos-about-to-secure-a-record-4-bln-in-year-long-ico, accessed on 2020-10-14.

Techcrunch.com 2018, Exit scammers run off with $660, https://techcrunch.com/2018/04/13/exit-scammers-run-off-with-660-million-in-ico-earnings/, accessed on 2020-09-22.

Turan, S. S. 2015. Stakeholders in Equity-Based Crowdfunding: Respective Risks Over the Equity Crowdfunding Lifecycle. ***Journal of Financial Innovation***, 1(2): 141–151.

Velmurugan, T. and Santhanam, T. 2010. Computational Complexity between K-Means and K-Medoids Clustering Algorithms for Normal and Uniform Distributions of Data Points. ***Journal of Computer Science***, 6(3): 363–368.

Vismara, S. 2018. Information cascades among investors in equity crowdfunding. ***Entrepreneurship: Theory and Practice***, 42(3): 467–497.

Wallmeroth, J. 2019. Investor behavior in equity crowdfunding. ***Venture Capital***, 21: 273–300.

Welch, I. 1992. Sequential Sales, Learning, and Cascades. ***The Journal of Finance***, 47(2): 695–732.

Wick, J.-N. and Ihl, C. 2019. Herding or Avoiding? Disentangling Social Influence Among Heterogeneous Investors. ***Conference Proceedings Druid18***. Available at https://conference.druid.dk/acc_papers/bto5poplqa0deatam7lcvtzc0td83b.pdf.

Wilson, T. D. and Schooler, J. W. 1991. Journal of Personality and Social Psychology Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions. ***Journal of Personality and Social Psychology***, 60(602): 181–192.

Wöckner, K., Lösing, C., Diehl, T., and Kutzbach, A. 2018. Regulation of Initial Coin Offerings. ***White & Case White Paper***. Available at https://www.whitecase.com/publications/alert/regulation-initial-coin-offerings.

Wold, S., Esbensen, K. I. M., and Geladi, P. 1987. Principal Component Analysis. ***Chemometrics and Intelligent Laboratory Systems***, 2(1-3): 37–52.

Yujian, L. and Bo, L. 2007. A Normalized Levenshtein Distance Metric. ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, 29(6): 1091–1095.

Yum, H., Lee, B., and Chae, M. 2012. From the wisdom of crowds to my own judgment in microfinance through online peer-to-peer lending platforms. ***Electronic Commerce Research and Applications***, 11(5): 469–483.

Zacharakis, A. L. and Shepherd, D. A. 2001. The nature of information and overconfidence on venture capitalists' decision making. ***Journal of Business Venturing***, 16(4): 311–332.

Zetzsche, D., Buckley, R., and Amer, D. 2017. The ICO Gold Rush: It's a scam, it's a bubble, it's a super challenge for regulators. ***Law Working Paper Series, University of New South Wales***, 011(2017).

Zhang, J. and Liu, P. 2012. Rational Herding in Microloan Markets Rational Herding in Microloan Markets. ***Management Science***, (January 2015).

Zhang, S., Zhang, D., Zheng, J., and Aerts, W. 2020. Does policy uncertainty of the blockchain dampen ICO markets? ***Accounting and Finance***, 2019 China Issue. doi:10.1111/acfi.12639.

Zhang, Y. and Gimeno, J. 2016. Earnings pressure and long-term corporate governance: Can long-term-oriented investors and managers reduce the quarterly earnings obsession? ***Organization Science***, 27(2): 354–372.

Appendix

TABLE 7.1: Full list of ICO attributes

Variable	Description	Measurement	Missing ICOs	Regression set
Campaign attributes				
Blockchain industry	Specifies whether the ventures core business is related to blockchain technology	Dummy	0 (0%)	Yes
Duration	Length of ICO campaign in days.	Days	23 (2%)	Yes
Pre ICO	Specifies whether an ICO has carried out a preliminary token sale before the actual ICO.	Dummy	0 (0%)	Yes
Softcap	Minimum fundraising amount, if the softcap is not reached during the ICO, the venture commits to returning funds to all investors (ventures are not required to set a softcap).	US Dollar	847(66%)	No
Hardcap	Maximum amount to be raised in the ICO (ventures are not required to set a hardcap).	US Dollar	391 (31%)	No
Token supply	Number of tokens issued during ICO fundraising.	Integer	801 (63%)	No
Funding goal	Funding goal set and communicated by the ICOing venture	US Dollar.	1,067 (84%)	No
Minimum investment	Minimum ticket size per investment	US Dollar.	1,144 (90%)	No
Additional emission	Statement by venture on whether they plan to issue additional tokens in future fundraising, thus diluting ownership rights of initial token holders.	Dummy	424 (33%)	No
KYC required	Know your customer process required during ICO, that is investors must register and reveal information about themselves (each ICO may custom-design this process).	Dummy	497 (39%)	No
Retained Equity	Share of equity retained by founders in the ICO.	Percentage	766 (60%)	No
Accept fiat	Tokens can be purchased directly in exchange for fiat currency (legal tender) in the ICO.	Dummy	786 (62%)	No
Token price	Price of ICO tokens during fundraising.	US Dollar	3 (0%)	No
Tax haven	Is the venture listed in a country that is regarded as tax haven by the European Union (European Parliament, 2020)?	Dummy	1,177 (92%)	No
Country	Country of listing	Categorical	1 (0%)	No
Token platform	Technical token setup, list of 15 token types, most commonly ERC-20, a token template provided by Ethereum blockchain	Categorical.	1,010 (79%)	No

Continued on next page

TABLE 7.1 – continued from previous page

Variable	Description	Measurement	Missing ICOs	Regression set
Minable	Additional units of cryptocurrency can be created through mining on the blockchain, that is, in exchange for computing power provided to the blockchain.	Dummy	880 (69%)	No
Open source	Project code is editable by other developers.	Dummy	876 (69%)	No
Venture attributes				
GitHub stars	Measure of popularity of a project on GitHub, a widely known and used open source coding repository (Github, 2020).	Integer	948 (74%)	Yes
Project status	Founding stage of the venture: Just an idea (base level), Minimum viable product (MVP), Beta version or Working product (CoinCheckup, 2020).	Categorical	889 (70%)	Yes
CEO experience	Years of CEO experience, based on LinkedIn profile (CoinCheckup, 2020).	Years	876 (69%)	Yes
CEO business experience	Years of business experience, based on LinkedIn profile (CoinCheckup, 2020).	Years	951 (75%)	No
CEO large corporation	Has the CEO managed multi-million dollar projects or companies before (CoinCheckup, 2020)?	Dummy	884 (69%)	No
CEO honest	Does the CEO mention ICO alliance on her LinkedIn profile (CoinCheckup, 2020)?	Dummy	916 (72%)	No
CTO code experience	Years of experience as a developer (CoinCheckup, 2020).	Years	1,022 (80%)	No
CTO years as CTO	CTO business experience as CTO.	Years	876 (69%)	No
CTO honest	Does the CTO mention ICO alliance on her LinkedIn profile (CoinCheckup, 2020)?	Dummy	988 (78%)	No
CTO GitHub link	Is the CTO linked to this project on GitHub (CoinCheckup, 2020)?	Dummy	1,164 (91%)	No
CTO large corporation	Has the CTO worked for a multi-million dollar projects before (CoinCheckup, 2020)?	Dummy	901 (71%)	No
GitHub community	Number of followers on GitHub (Github, 2020).	Integer	876 (69%)	No
GitHub commits	Saved changes to project code on GitHub, thus a proxy for activity and active development of code (Github, 2020).	Integer	948 (74%)	No
GitHub forks	Number of times the projects GitHub repository has been copied by other developers (Github, 2020).	Integer	948 (74%)	No
Number of developers	Number of developers in the team (CoinCheckup, 2020).	Integer	876 (69%)	No
Team size	Number of team members excluding advisors (CoinCheckup, 2020).	Integer	876 (69%)	No
Team bio available	Biography visible on the website for all team members (CoinCheckup, 2020)?	Dummy	876 (69%)	No
Team social media link	Does the venture provide social media links for most team members (CoinCheckup, 2020)?	Dummy	876 (69%)	No
Advisor bio available	Biography visible on the website for all advisors (CoinCheckup, 2020)?	Dummy	889 (70%)	No
Advisor large corporation	Has at least 1 advisor managed a multi-million dollar project before (CoinCheckup, 2020)?	Dummy	886 (70%)	No
Advisor social media bio	Do most of the advisory board members on the team page have links to social media profiles (Facebook, LinkedIn, GitHub) (CoinCheckup, 2020)?	Dummy	885 (69%)	No
Incorporated company	Does the venture provide social media links for most advisors (CoinCheckup, 2020)?	Dummy	876 (69%)	No
Incorporated official	Is the venture officially incorporated as a company (CoinCheckup, 2020)?	Dummy	926 (73%)	No