TECHNISCHE UNIVERSITÄT DARMSTADT

DISSERTATION

Market Efficiency, Behavior and Information Asymmetry: Empirical Evidence from Cryptocurrency and Stock Markets

Vorgelegt von:
David HÄFNER, M.Sc.
geboren am 07.12.1990
in Schwäbisch Hall,
Deutschland

Erstgutachter:
Prof. Dr. Dirk SCHIERECK
Zweitgutachter:
Prof. Dr. Peter BUXMANN

Vom Fachbereich Rechts- und Wirtschaftswissenschaften der
Technischen Universität Darmstadt

zur Erlangung des akademischen Grades
Doctor rerum politicarum (Dr. rer. pol.)

Darmstadt 2020

Häfner, David: "Market Efficiency, Behavior and Information Asymmetry: Empirical Evidence from Cryptocurrency and Stock Markets"
Darmstadt, Technische Universität Darmstadt
Tag der Einreichung: 02.07.2020
Tag der mündlichen Prüfung: 15.12.2020
URN: urn:nbn:de:tuda-tuprints-174953
Jahr der Veröffentlichung der Dissertation auf TUprints: 2021

"I don't know what I may seem to the world, but, as to myself, I seem to have been only like a boy playing on the sea shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me."

Sir Isaac Newton

TECHNISCHE UNIVERSITÄT DARMSTADT

Abstract

Fachbereich Rechts- und Wirtschaftswissenschaften
Fachgebiet Unternehmensfinanzierung

Dr. rer. pol.

Market Efficiency, Behavior and Information Asymmetry: Empirical Evidence from Cryptocurrency and Stock Markets

by David HÄFNER

This dissertation is dedicated to the analysis of three superordinate economic principles in varying market environments: market efficiency, the behavior of market participants and information asymmetry.
Sustainability and social responsibility have gained importance as investment criteria in recent years. However, responsible investing can lead to conflicting goals with respect to utility-maximizing behavior and portfolio diversification in efficient markets. Conducting a meta-analysis, this thesis presents evidence that positive (non-monetary) side effects of responsible investing can overcome this burden. Next, the impact of the EU-wide regulation of investment research on the interplay between information asymmetry, idiosyncratic risk, liquidity and the role of financial analysts in stock markets is investigated. An empirical analysis of the emerging primary and secondary market for cryptocurrencies yields further insights about the effects of information asymmetry between investors, issuers and traders. The efficient allocation of resources is dependent on the market microstructure, the behavior of market participants, as well as exogenous shocks. Against this background, this thesis is dedicated to the empirical analysis of limit order books, the rationality of traders and the impact of COVID-19. Due to its young history, the market for cryptocurrencies yields a suitable research subject to test classical financial theories. This doctoral thesis reveals parallels between the microstructure of cryptocurrency and stock markets and uncovers some previously unknown statistical properties of the cryptocurrency market microstructure. An initial examination of the impact of COVID-19 further shows that cryptocurrencies with a high market capitalization seem to react to macroeconomic shocks similar to stock markets.
This cumulative dissertation comprises six stand-alone papers, of which three papers have already been published.

Zusammenfassung (Deutsche Version)

Diese Dissertation widmet sich der Analyse von drei übergeordneten wirtschaftswissenschaftlichen Konzepten in verschiedenen Marktumfeldern: Markteffizienz, Verhalten von Marktakteuren und Informationsasymmetrie.
In den letzten Jahren haben Nachhaltigkeit und soziale Verantwortung als entscheidungsrelevante Investitionskriterien ständig an Bedeutung gewonnen. Verantwortungsvolles Verhalten kann auf effizienten Märkten allerdings zu einem Zielkonflikt hinsichtlich der individuellen Nutzenmaximierung und der Portfoliodiversifikation führen. Diese Arbeit liefert anhand einer Metaanalyse Evidenz dafür, dass (nicht-monetäre) Nebeneffekte verantwortungsvollen Investierens diesen Zielkonflikt überwinden können. Anschließend werden die Auswirkungen der EU-weiten Regulierung von Investment Research auf das Zusammenspiel von Informationsasymmetrie, idiosynkratischem Risiko, Liquidität und die Rolle von Finanzanalysten im Aktienmarkt untersucht. Darüber hinaus liefert eine empirische Analyse des aufstrebenden Primär- und Sekundärmarkts für Kryptowährungen neue Erkenntnisse über die Auswirkungen von Informationsasymmetrien zwischen Investoren, Emittenten und Händlern. Die effiziente Allokation von Ressourcen hängt von der Marktmikrostruktur, dem Verhalten von Marktakteuren, sowie von exogenen Schocks ab. Vor diesem Hintergrund widmet sich diese Arbeit der empirischen Analyse von Limit-Orderbüchern, der Rationalität von Händlern und den Auswirkungen von COVID-19. Der Markt für Kryptowährungen bietet aufgrund seiner jungen Historie einen geeigneten Forschungsgegenstand, um klassische Finanzierungstheorien empirisch zu testen. Diese Dissertation zeigt Parallelen zwischen der Mikrostruktur von Kryptowährungs- und Aktienmärkten auf und deckt einige bisher unbekannte statistische Eigenschaften der Marktmikrostruktur von Kryptowährungen auf. Eine erste Untersuchung der Auswirkungen von COVID-19 zeigt zudem, dass insbesondere Kryptowährungen mit einer hohen Marktkapitalisierung ähnlich wie Aktienmärkte auf makroökonomische Schocks zu reagieren scheinen.
Diese kumulative Dissertation umfasst sechs eigenständige Artikel, von denen drei Artikel bereits veröffentlicht wurden.

Acknowledgements

I want to thank Regina for her unconditional support and understanding during the preparation of this dissertation. I want to thank Prof. Dr. Schiereck for the very fruitful discussions, and his valuable input during my time at his chair.

Contents

List of Figures

List of Tables

List of Abbreviations

ADJR	**Adj**ustierte **R**endite
AMF	**A**utorité des **M**archés **F**inanciers
API	**A**pplication **P**rogramming **I**nterface
BaFin	**B**undes**a**nstalt für **Fin**anzdienstleistungsaufsicht
BCH	**B**itcoin **C**ash
BTC	**B**itcoin
CAPM	**C**apital **A**sset **P**ricing **M**odel
CEO	**C**hief **E**xecutive **O**fficer
COVID-19	**Co**rona**vi**rus **D**isease 20**19**
CSR	**C**orporate **S**ocial **R**esponsibility
EBA	**E**uropäische **Ba**nkenaufsicht
ECN	**E**lectronic **C**ommunications **N**etwork
EMH	**E**fficient **M**arket **H**ypothesis
EPS	**E**arnings **p**er **S**hare
ERC	**E**thereum **R**equest for **C**omments
ESG	**E**nvironmental, **S**ocial and **G**overnance
ESMA	**E**uropean **S**ecurities and **M**arkets **A**uthority
ETF	**E**xchange-**T**raded **F**und
ETH	**E**thereum
EU	**E**uropean **U**nion
EUR	**E**uro
HML	**H**igh **M**inus **L**ow
HFT	**H**igh **F**requency **T**rading
I/B/E/S	**I**nstitutional **B**rokers' **E**stimate **S**ystem
ICO	**I**nitial **C**oin **O**ffering
IPO	**I**nitial **P**ublic **O**ffering
ISIN	**I**nternational **S**ecurities **I**dentification **N**umber
KMU	**K**leine und **m**ittlere **U**nternehmen
LOB	**L**imit **O**rder **B**ook
LTC	**L**itecoin
M&A	**M**ergers and **A**cquisitions
MiFID	**M**arkets **i**n **F**inancial **I**nstruments **D**irective
MOM	**M**omentum
NYSE	**N**ew **Y**ork **S**tock **E**xchange

OLS	**O**rdinary **L**east **S**quares
R&D	**R**esearch and **D**evelopment
RER	**R**ohe **E**missions**r**endite
SEC	United States **S**ecurities and **E**xchange **C**ommission
SIC	**S**tandard **I**ndustrial **C**lassification Code
SMB	**S**mall **M**inus **B**ig
SRI	**S**ocially **R**esponsible **I**nvestment
US	**U**nited **S**tates of America
USD	**US D**ollar
VC	**V**enture **C**apital

Dedicated to my parents.

Chapter 1

Synopsis

1.1 Motivation

Financial markets contain an incredible amount of information about the human behavior in a competitive setting. Market participants express their believes about the value of an asset by indicating the willingness of selling or buying the asset at a specific price. Loosely speaking, market participants "put money where their mouth is" when trading an asset, and are monetary incentivized to act rational. This design of financial markets yields a fertile environment to empirically test economic, financial and behavioral hypotheses.

Financial markets shape the world we live in and play a crucial role in distributing resources efficiently and by uncovering hidden market characteristics and flaws in market design, science can add to a broader understanding of financial markets and thereby benefit the society.

One of the most important discoveries in the finance literature is the concept of market efficiency. Market efficiency guarantees that the best available projects receive funding. Fama (1970) hypothesizes that markets are only efficient, if all available information is reflected in asset prices and no market participant can make long term profits. However, Fama (1970) differentiates between different levels of market efficiency and it remains uncertain to this day, to what extent financial markets are efficient. Furthermore – accompanied by the technological advancement of humankind – markets are in constant change and new markets arise, while others collapse. However, there seem to exist universal characteristics and repetitive patterns hidden in the nature of financial markets and the competitive behavior of market participants, observable across different assets, eras and market environments.

Since establishing the efficient market hypothesis, critics on the validity emerged and empirical data suggests that anomalies exist in financial markets and imperfections are often better explained by cognitive biases and irrational behavioral of market participants. Moreover, social responsibility and regulatory constraints shape the market environment and impact the behavior of market participants.

While traditional financial theories focus on the wealth of one individual and view the maximization of the present value or a typically unobservable individual utility

function as the sole *dictum*, more recent models also account for investors, who may exhibit other strategic patterns depending on their individual behavior.
Moreover, conflicts of aims between stakeholders and the general welfare require concepts like social and environmental responsibility to be taken into account during investment decisions. Despite its importance, the compatibility of social responsibility and rational investment decisions is academically not conclusively clarified and understood making further research necessary.
This dissertation aims to explore the behavior and interaction of agents in financial markets in different market conditions, study the impact of real world features on capital allocation, and is driven by examining the facets of three superordinate concepts, relevant for financial markets:

- **Market Efficiency**
- **Behavior of Market Participants**
- **Information Asymmetry**

This thesis aims to provide new findings that contribute to society and especially academia and is geared towards broadening the understanding of interactions in financial markets, the impact of social and regulatory constraints, the behavior of market participants and the effect of information disparity across agents. Derived from the concepts of market efficiency, the behavior of market participants and information asymmetry, we formulate several research questions that act as a guideline throughout this dissertation and motivate the following chapters:

- **Research Question 1**: Is the concept of market efficiency compatible with social responsibility?
- **Research Question 2**: How do participants in highly efficient stock markets react to an exogenous shock in market design (regulatory change)?
- **Research Question 3**: How do investors cope with information asymmetry?
- **Research Question 4**: How does the market microstructure shape market dynamics?
- **Research Question 5**: How do traders behave when placing limit buy and sell orders in a competitive market setting?
- **Research Question 6**: How does the cryptocurrency market react to an exogenous shock (market crisis)?

1.2 Thesis Structure

This thesis is structured into six stand-alone research papers. The thesis comprises three published papers (Chapter 2, Chapter 4, and Chapter 7) and three unpublished research papers (Chapter 3, Chapter 5, and Chapter 6).
In Chapter 2: "*What do we know about socially responsible investments?*", we conduct a literature review on socially responsible investments (SRIs) from three different perspectives: The perspective of an investor engaging in SRIs, the company and management that acts socially responsible and the performance of institutional funds engaging in SRIs. We discuss corporate social responsibility (CSR) and the three central factors environmental, social and governance (ESG) as major decision drivers in a firm's responsible management. We further analyze investors, who demand social responsibility from firms and institutional funds with an SRI focus, which are often restricted to only invest in companies that fulfill various SRI screening criteria. Incorporating certain social constraints in the investment strategy reduces the size of the possible investment universe which should *ceteris paribus* not lead to a superior financial performance. However, our meta study shows that financial performance is not necessarily the only objective of investors and company management. Investing and acting socially and environmentally responsible can have non-monetary benefits. Thus, while financial gain and social responsibility are oftentimes conflicting goals, the positive side effects of responsible investing and management can overcome this burden. We also highlight the motivation, the behavior and demographics of socially responsible investors and discuss the (non-)financial motivation of managers to act responsible and outline characteristics of socially responsible companies. Finally, we discuss portfolio implications for investment funds under regional considerations, risk and uncertainty and financial performance when engaging in SRIs.

In Chapter 3: "*The Role of Investment Research in view of MiFID II: An Empirical Analysis of Information Asymmetry, Idiosyncratic Risk and Liquidity*", we focus on the impact of regulation on key characteristics of financial markets. Regulation aims to improve market quality by increasing price stability, transparency and thereby securing capital supply for the economy. However, implementing new market regulation is inevitably accompanied by direct and indirect costs that are often borne by market participants and regulators have to carefully weigh up the pros and cons of regulation. Therefore, understanding the consequences of market interventions is important for a broad range of stakeholders, including regulators, listed companies, investors and economists alike. Based on this theoretical background, we study the impact of the revised EU Markets in Financial Instruments Directive (MiFID II) on European financial markets. MiFID II came into force in January 2018, targeting the regulation of investment research services across all EU member states. Investment

research provided by financial analysts plays a crucial role in information dissemination about company specific topics. Meanwhile, financial literature still struggles to explain the exact role of financial analysts. While they should not be able to gather valuable information in efficient markets, they seem to play a crucial role by distributing information between investors and companies via investment research services and reports. We aim to contribute to the literature by analyzing the role of financial analysts during a regulatory change (MiFID II) that heavily affects the way investment research is conducted. Using Data about 1,646 US firms and 1,281 EU firms, we compare the US and EU approach to regulate investment research services. We find a decline in stock liquidity and an increase in idiosyncratic risk for stocks affected by MiFID II. In a second step, we focus on the informational role of research analysts and hypothesize that MiFID II increased the competition between financial analysts, leading to higher quality work and consequently to a higher informativeness of stock prices. While we do not find that analyst coverage decreased in the short term, we find that the amount of coverage in EU markets declined in the medium term. We further find a significant decrease in stock liquidity and an increase in the bid-ask spread, idiosyncratic risk and the level of asymmetric information that can be associated with MiFID II. Finally, we find empirical evidence suggesting that MiFID II affects the informational role research coverage has on the bid-ask spread and a stock's idiosyncratic risk.

In Chapter 4: "*Innovative Finanzierung über Initial Coin Offerings: Struktur und bisherige Performance*", we study how transparency issues and information asymmetry between different parties affects the capital allocation. Initial Coin Offerings (ICOs) are a relatively new way of raising capital and due to their simplicity, low regulatory burdens and marginal direct costs, they are typically carried out by entrepreneurial and smaller companies. We show that ICOs share many similarities with traditional Initial Public Offerings (IPOs). A much researched phenomenon observed during the IPO process is the underpricing effect, which is defined as a positive stock return at the first trading day of a stock. Academia provides different theories that try to explain this persistent phenomenon (e.g Ritter, 1987). While behavioral approaches exist as well, popular theories focus on information asymmetries between the issuing company, the underwriter and investors. In these models, underpricing compensates uninformed investor or signals company well-being. We discuss the different theories in detail and empirically show that underpricing can be observed during ICOs as well and is even more pronounced. During ICOs companies typically do not provide much information about the aspired use of the capital needed, creating an environment of high information asymmetry between investors and the issuing company. Hence, the empirically observed level of underpricing could be a result of investors demand for a payoff compensating uncertainty. We find that companies generally "leave money on the table" during an ICO and could increase their capital raised by optimizing their communication of information towards investors during an ICO.

In Chapter 5: "*Statistical Properties of Cryptocurrency Order Books*", we focus on the market microstructure of a new market segment. Motivated by the results of the previous chapter, which targets information asymmetries in the primary market for cryptocurrencies, we now focus our analysis on the secondary market by analyzing the market microstructure in cryptocurrency markets. In a first step, we analyze traders activities expressed in the limit order book (LOB) and compute the aggregated LOB volume, i.e. the slope of the order book, dependent on the price level for three major cryptocurrencies: Bitcoin, Bitcoin Cash and Ethereum. We find volume peaks at certain relative price levels distant to the best price. We show that these volume peaks are statistically significant and can be found across all observed LOBs. Next, we test, if there is information hidden in the LOB by analyzing the relation of the daily average order book slope and daily price changes, trading volume, and the volume-volatility relation found in stock markets. We find significant links between the slope of the LOB and returns, indicating that investors consider the whole LOB when buying or selling cryptocurrencies. We further find links between trading activity and the slope of the LOB but the relation switches sign when considering the full depth of the LOB. While its existence is quite puzzling, this market anomaly has been documented in stock markets as well by Næs and Skjeltorp (2006). We further find that trading volume and volatility are positively correlated, indicating that the volume-volatility relation, which has been documented in the literature across different stock markets, exists in cryptocurrency markets as well and therefore seems to be a market independent characteristic.

In Chapter 6: "*Heuristics in Cryptocurrency Limit Order Placement*", we focus on the behavior of cryptocurrency investors as these market participants directly impact the market movements with their buying and selling behavior. Using data about incoming limit order prices, we empirically show that cryptocurrency traders apply heuristics during capital allocation. We develop a theoretical model to take this behavior into account and empirically show that this extended model can better predict the limit order placement behavior than previous models suggested in the literature.

In Chapter 7: "*Reaktionen der Kryptowährungsmärkte auf die COVID-19-Pandemie*", we empirically test, how the outbreak of the COVID-19 pandemic affects cryptocurrencies. By constructing two cryptocurrency portfolios based on market capitalization, we show that smaller cryptocurrencies react differently to this global crisis and perform superior compared to larger ones since the outbreak. Further, we employ a fixed-effects regression model and find empirical evidence for a relationship between prior trading volume and returns, and autocorrelation of returns in cryptocurrency markets. Moreover, the intertemporal relation between past trading volume and current returns seems to be enhanced for smaller cryptocurrencies since COVID-19 is shaping the macroeconomic development.

The studies conducted in this dissertation are performed in different market environments (see Table 1.1), allowing us to get comprehensive insights into the behavior and mechanisms of competitive markets. We study the impact of different market frictions to better understand how capital can be efficiently allocated and this dissertation aims to contribute to the academic literature by providing new empirical evidence.

While all studies conducted in this dissertation are independent of each other, they have an important common denominator, as they are connected via the concepts of market efficiency, economic behavior and information asymmetry and each chapter highlights a different aspect of these crucial concepts. Table 1.2 provides an overview of the framework in which the three aspects and their features are addressed in each of the following chapters. Table 1.1 summarizes the market conditions and data usage in each chapter. From Chapter 2 to Chapter 6, the observation period steadily decreases while the data granularity increases. While we first analyze established stock markets (Chapter 2 and Chapter 3), we later turn to the global emerging market for cryptocurrencies (Chapter 4 to Chapter 7) and draw parallels between these two markets.

TABLE 1.1: Research Setting and Data

Notes: This table shows the research setting of each chapter of this dissertation and meta information about the empirical data analyzed.

Chapter	**Region**	**Market Environment**	**Observation Period**	**Data Granularity**
2	Global	Established	Decades	Annual/Monthly
3	EU&US	Established/In change	Years	Monthly/Daily
4	Global	Young	Months	Months/Daily
5	Global	Young	Months	Months/Daily
6	Global	Young	Days	Daily/Seconds
7	Global	Young/In change	Months	Daily

TABLE 1.2: Research Focus across Chapters

Notes: This table shows the research topics of each chapter of this dissertation, based on three different aspects: market efficiency, investor behavior and information asymmetry.

Chapter	**Agents**	**Market Efficiency**	**Investor Behavior**	**Information Asymmetry**
2	• Investor • Company	• Compatibility with social responsibility	• Social responsibility	• Between management interests and shareholders
3	• Investor • Company • Analyst	• Efficiency under regulatory constraints	• Reaction to new regulation • Competition between analysts	• Between investors and analysts • Across investors
4	• Investor • Company	• Transparency • Underpricing	• ICO pricing • Information policy	• Between issuing company and investors • Across investors
5	• Investor	• Market impact of LOB characteristics	• Implied demand and supply	• Across investors
6	• Investor	• Rationality of investors	• Order placement behavior	• Across investors
7	• Investor • Company	• Predictability • Efficiency during crisis	• Flight to "save havens"	-

Chapter 2

What do we know about socially responsible investments?

Chapter 2 has been published as a journal article:
Häfner, David, Florian Kiesel, and Lucas Wirthmann (2017). "What do we know about socially responsible investments?" In: *Zeitschrift für Umweltpolitik und Umweltrecht* (4), pp. 299–331, ISSN 0931-0983.

Chapter 3

The Role of Investment Research in view of MiFID II: An Empirical Analysis of Information Asymmetry, Idiosyncratic Risk and Liquidity

The basic idea of regulation in financial markets is to improve the market quality. In this paper, we employ a difference-in-differences approach by using daily data of 1,646 US firms and 1,281 EU firms to investigate the influence of the EU-wide implementation of the EU Markets in Financial Instruments Directive (MiFID II) in January 2018. MiFID II enforces that investment research provided by financial analysts explicitly has to be paid for by investors. Market participants expected a reduction in demand for research services and as a result more information asymmetries in European stock markets. Consistent with these expectations we find that the implementation of MiFID II has led to a significant decline in stock liquidity and an increase in stocks' idiosyncratic risk.
In a second step, we examine the effect of a change in investment research coverage before and after the implementation of MiFID II. We find that an increase in the number of analysts covering a stock significantly affects information asymmetry and idiosyncratic risk, supporting the idea that analysts act as a source of information for investors. We show that both effects are amplified by the introduction of MiFID II.

3.1 Introduction

Financial market regulation aims to direct the market development towards an economically preferable condition where capital supply and price stability is guaranteed. However, the approach and the extent to which regulation is implemented

bears many risks, as complying with more directives comes along with an additional financial burden for affected companies. In a globalized world, competing jurisdictions need to be taken into account as well as capital is not bound to national borders and investors may leave if the legal environment is too restrictive. Hence, regulation needs to be reasonable and only be implemented if severe market inefficiencies predominate.

Focusing on the impact of regulation, Thomsen and Vinten (2014) distinguish between two different hypotheses concerning the costs and benefits of investor protection regulation. The efficiency hypothesis states that benefits of investor protection regulation outweighs the costs leading to the implementation of new regulation to improve the stock markets. For this reason, regulation would likely have a positive effect on company performance and stock prices. On the other hand, the over-regulation hypothesis postulates that regulation originates from rent seeking from powerful economic players which follow their own financial interests leading to costs which exceed the benefits of regulation (Thomsen and Vinten, 2014, p. 800). While the authors focus on the influence of regulation on company delistings, their cost-benefit hypothesis can be transferred to any influence of regulating capital markets.

Information asymmetry between investors and capital seekers describes one form of market failure and can have a serious impact on trading costs. Therefore, market transparency is an important objective for policy makers.

Based on this theoretical background, we study the impact of the revised EU Markets in Financial Instruments Directive (MiFID II) on financial markets. As of January 2018, MiFID II forces, *inter alia*, that investment research has to be billed explicitly. In this paper, we study the effect of MiFID II on European financial markets by employing a difference-in-differences approach around the implementation day of the new regulation, studying the impact of MiFID II on investment research coverage, liquidity and idiosyncratic risk. In a second step, we study the relationship between investment research coverage and information asymmetry and the impact of MiFID II on this relationship. We hypothesize that MiFID II increased the competition between financial analysts. Financial analysts now have to proof their "worth", which should increase the quality of their work and lead to a higher informativeness of stock prices. We find that analyst coverage did not decrease significantly in the short term. However, the amount of coverage overall declined in the medium term. MiFID II further led to a significant decrease in stock liquidity and an increase in the bid-ask spread, idiosyncratic risk and the level of asymmetric information. We also find that MiFID II affects the informational role research coverage has on the bid-ask spread and idiosyncratic risk by employing a fixed effects regression model.

The remainder of this paper is structured as follows: Section 3.2 provides an overview of investment research regulation and compares the treatment of investment research in the European jurisdiction to the regulatory framework in the United States.

Section 3.3 gives a literature review on the theoretical framework regarding the role of financial analysts in capital markets. Section 3.4 defines the terms liquidity, information asymmetry and idiosyncratic risk and introduces empirical measures for these theoretical concepts. Section 3.5 describes the data which is used in our empirical analysis conducted in Section 3.6. Section 3.7 concludes.

3.2 Regulation of Investment Research

In this section we focus on the regulatory treatment of investment research. We highlight the historic way that led to the introduction of MiFID II and compare EU and US approaches to regulate investment research. We find the regulatory framework of the United States to be comparable to the pre-MiFID II European framework making US stocks suitable for the control group of our subsequent empirical analysis.

3.2.1 Investment Research in the European Union

The first Markets in Financial Instruments Directive (MiFID I) was introduced across all member states of the European Economic Area on 31st of November 2007. The aim of MiFID I was to facilitate cross-border trading for private and institutional investors. Further, the European Commission tried to achieve a harmonization across European trading venues. As initially discussed, the superior regulatory aim of MiFID I was to benefit the economy by lowering the overall cost of capital. In order to be reasonable, this effect must exceed the costs of complying with new regulatory requirements. MiFID I led to a significant regulatory overhaul in Europe and is widely regarded as one of the most significant regulatory changes for financial markets (Casey and Lannoo, 2009 and Ferrarini and Wymeersch, 2006).
Since 3rd of January 2018, the revised EU-wide Markets in Financial Instruments Directive (MiFID II) is in force, aiming to increase the transparency and the investor protection in the European financial markets. MiFID II replaces MiFID I and involves regulatory changes targeting investment research. Notably, investment research provided by brokers in the form of stock analysis, research reports and access to the management of covered firms, now has to be paid for by potential investors explicitly – a change which has drawn substantial media attention. For the first time, the exact monetary value of investment research has to be determined, disrupting the market for investment research services. Prior to MiFID II, investment research was commonly provided by brokers free of charge with the intention to encourage investors to engage in trading. Providing investment research has therefore been part of the overall brokerage service, providing investors with information on the one hand and serving as a marketing tool on the other hand. The actual costs for creating investment research have been usually cross-subsidized via the broker margin of subsequent trades. As a consequence, it was almost impossible for investors to

keep track of how much they actually spend for investment research. This procedure has been deemed to be too intransparent by EU regulators and is prohibited since MiFID II is in effect. Consequently, investment research has to be priced explicitly since 3rd of January 2018, raising questions about the monetary value of investment research for brokers and investors.

Uncertainty remains on how the new regulation affects the availability and quality of research. One of the main concerns with the implementation of MiFID II is its effect on the overall research coverage of companies. Especially small companies may suffer from an increase in illiquidity, caused by a decrease in the number of research analysts. The investors' demand for research considering those firms might be too small to be profitable for research providers. The general belief is that the overall supply of investment research decreases, leading to a lower level of analyst coverage and hence to a decrease in liquidity (Deutsche Börse, 2017). Smaller listed companies have been anticipated to be the most affected by MiFID II as they already suffer from a low analyst coverage which potentially drops to zero due to a lack of investor interest (Deutscher Investor Relations Verband, 2017). Providing some early empirical evidence, Fang et al. (2019) find a decline in analyst coverage and show that recommendations of remaining analysts receive more market attention and have greater information content. The authors further find analysts' participation in earnings-conference calls and the number of questions asked to have increased since MiFID II. A survey conducted by the CFA institute also finds that the market place for research has become more competitive and 44% of sell-side respondents believe that the quality of investment research has declined while buy-side respondents do not see a change in research quality (Preece, 2019). Further, 47% of buy-side and 53% of sell-side participants respond with a decrease in coverage for small- and mid-cap stocks since MiFID II is in place.

From a theoretical point of view, market participants need to know whether their source of investment research possesses private information on which trading profits can be generated, as only then it would be worth to pay for it.[1] With the new regulatory framework in place, investors are incentivized to learn about the skill of financial analysts as they face obvious monetary losses, if they base their investments on paid research that does not contain private information. We hypothesize that this relationship leads to an extinction of poorly performing financial analysts as investors will pick their source of information more carefully. We argue that this change should be measurable through a change in the impact of financial analysts on stock specific characteristics. Consequently, we would not expect any measurable effect of analyst coverage at all, if financial analysts are not able to create or unveil private information with their work.

[1]Note that the amount, which is paid for investment research can be interpreted as the cost of information in the model suggested by Grossman and Stiglitz (1980).

Generally, we suppose that an increase in analyst coverage decreases the information asymmetry of a stock as the competition between analysts increases. This is lowering the price for which analysts are willing to disclose their information ultimately increasing transparency. We expect that MiFID II amplifies this effect as financial analysts have to demonstrate their value to investors. We would also expect the overall amount of analyst coverage to drop due to MiFID II, especially for smaller companies as it is likely less profitable to cover them.

3.2.2 Investment Research in the United States

The treatment of investment research in the US is regulated in the Investment Advisers Act of 1940. Section 202(a)(11) defines an investment adviser as "any person, who, for compensation, engages in the business of advising others, either directly or through publications or writings, as to the value of securities or as to the advisability of investing in, purchasing, or selling securities, or who, for compensation and as part of a regular business, issues or promulgates analyses or reports concerning securities". However, Section 202(a)(11)(C) of the Investment Advisers Act of 1940 specifically excludes "any broker-dealer whose performance of such services is solely incidental to the conduct of his business as a broker or dealer and who receives no special compensation therefor". The US Securities and Exchange Commission (SEC) generally treats providing research as the provision of investment advice. Though, Section 202(a)(11)(C) exempts broker-dealers from treatment as investment advisers “since broker-dealers that provide research generally include the cost of that research in the commissions they charge for execution of securities transactions" (Johnsen and Grady, 2017).[2] In the jurisdiction of the US this leads to the distribution of investment research to investors seemingly free of charge with compensation in the form of soft dollars, which is similar to the handling in the EU prior to MiFID II.

MiFID II runs counter to SEC regulations as it requires investment research to be paid for in "hard dollars". Under current SEC regulations, MiFID II would force all providers of cross-border research to register as an investment adviser in the US. For this reason, the SEC Division of Investment Management issued a "No-Action Letter" in October 2017, stating that it will not recommend enforcement action if a US broker-dealer provides research to an EU investment manager that is required to pay for research services (SEC, 2017). The relief is granted for a temporary period of 30 months from MiFID II’s implementation date.

[2]Registering as investment adviser in the US imposes multiple burdens for research providers complicating the dissemination of research, e.g. registered investment advisers need to take on extra fiduciary responsibilities impacting large parts of how the research business is currently done.

3.3 The Role of Investment Research in Capital Markets

In an attempt to explain the impossibility of informational efficiency in capital markets, Grossman and Stiglitz (1980) propose a noisy rational expectations equilibrium model. In their model, prices only partly reflect the information of informed traders, implying that those who expend resources to obtain information will also be able to generate profit off of this information. In their model, prices act as a delivery system of information from informed to uninformed traders. As investors can only profit from their superior information advantage by trading, they inevitably reveal their information to uninformed traders through the resulting price movement of the respective asset. The authors hypothesize that information can be noisy, i.e. markets are not fully efficient. Their model leads to an information asymmetry between two distinct groups of market participants: informed and uninformed (noise) traders. Grossman and Stiglitz (1980) show that the market price will reveal most of the information of informed traders when information is cheap or precise. Prices, however, cannot completely reflect the information which is available to the informed traders because in that case, those who paid for the information or created it cannot benefit from the information as it would already be fully incorporated in the market price. Transferred to our problem setting, we hypothesize that analysts might be able to offer private information by creating investment research and market participants can decide to buy this information.

Nevertheless, disagreement persists in the literature regarding the actual role of investment research produced by financial analysts. While one strain of literature provides evidence that financial analysts actually produce and provide valuable information (e.g. Womack, 1996, Barber et al., 2001, Gleason and Lee (2003), Kelly and Ljungqvist, 2012, Piotroski and Roulstone, 2004) other studies present findings indicating that investment research analysts do not seem to be able to create private information (e.g. Easley, O'Hara, and Paperman, 1998).

Above all this stands a well known information paradox regarding the existence of informational market efficiency: If all information is already incorporated in the market price, information itself has no financial value. Consequently, market participants would not gather new information which would lead to the case of new information not being reflected in market prices. In this stage, markets can not be information-efficient as it would be beneficial to gather private information again, leading to the paradox.

Focused on the role of research analysts, French and Roll (1986) hypothesize that more volatile prices during exchange trading hours could be caused by the frequency of information arrival during business days. The authors further state that public information affects prices before anyone can trade on it, e.g. weather, while private information is produced by investors and security analysts. The authors argue that more private information might occur while security analysts actively scan

company documents and thereby reveal previously unknown information. This could also explain the observation of higher volatility during business days. To further analyze the behavior of stock return variances, French and Roll (1986) split daily stock returns into a rational information component, a mispricing component and a bid-ask error. Although they identify mispricing errors, they state that the magnitude of these errors is too small to explain the difference in variances during trading and non-trading days. The authors attribute the observed effect to differences in the flow of information during trading and non-trading hours. Groundbreaking work in the analysis of the intrinsic value of stock research reports has also been conducted by De Bondt and Thaler (1990). Studying security analyst behavior, the authors use analysts' earnings forecasts to determine, whether forecasted changes are too extreme and whether the prediction bias grows with uncertainty, i.e. when the predicted changes range into the more distant future. De Bondt and Thaler (1990) argue that the question of distortedness in professional recommendations is especially important, as most investors neither have the time nor the necessary skill to produce independent predictions and therefore depend on buying earnings forecasts. The authors try to predict the actual change in earnings per share using the forecasted change as an explanatory variable, which – imposing efficient markets – should not be possible. They find evidence for excessive optimism and conclude that the findings can be attributed to an agency problem arising from analysts who work for broker houses and make money by encouraging trading. However, they question if this agency problem can be considered the sole reason for the proven overreaction, as similar patterns can be observed in scenarios where no agency conflict is present. Womack (1996) analyzes the influence of buy and sell recommendations issued by investment research analysts on stock prices. As those recommendations originate from predictions of stock values which should incorporate all industry and firm-specific information, Womack (1996) states that they allow to directly test, whether informed investors can outperform the stock market. He observes permanent changes following a recommendation, which indicates that recommendations include private information. He also observes that buy recommendations occur more frequently than sell recommendations. Together, these findings would legitimate the compensation of brokerage firms.[3] The author further analyzes stock prices and company-specific recommendations during the 1989-1991 time period and categorizes all recommendations into: "buy", "hold" and "sell" and finds cumulative average abnormal returns of 4.00% for a stock added to the "buy" list in a three-day event window. In contrast, the cumulative average return of a stock added to the "sell" list is -4.32% in the same time frame. This finding indicates that stock prices are influenced by analyst recommendations. Womack (1996) also finds evidence, that this influence persists over the long term and that the market reaction

[3]The compensation for providing investment research services is traditionally earned by soft dollar commissions.

is significantly larger for firms with a smaller market capitalization. Easley, O'Hara, and Paperman (1998) further investigate the role of analysts in financial markets, as disagreement between previous studies persists. French and Roll (1986) as well as Easley, O'Hara, and Paperman (1998) suppose that private information exists in the market and financial analysts might uncover private information with their work. However, the worth of a financial analyst's research output can range from a stock selling marketing tool up to an exploitable information edge. The latter case implies that financial analysts possess valuable private information. Easley, O'Hara, and Paperman (1998) try to investigate the informational role of financial analysts by estimating the probability of information-based trading using trade data of NYSE stocks. They further investigate, whether the amount of analyst coverage on a firm level increases the likelihood of a private information disclosure of a company. By computing the probability of informed trading, their technique allows to investigate, whether analysts create private information. Surprisingly, they find that a company that is highly covered by analysts has a lower risk of information-based trading. However, the amount of information-based trades is higher. This phenomenon can be explained by a high number of noise traders trading these stocks, diluting the overall risk of facing an information-based trade. Easley, O'Hara, and Paperman (1998) observe that the probability of private information events does not depend on the number of analysts covering a stock. Calculating the probability of information-based trading for each stock in the sample and regressing stock spreads on the number of analysts covering the stock, they show that analysts do not appear to create private information. These results contradict the findings of Womack (1996).
More recently, Wallmeier (2005) examines analysts' earnings forecasts for German DAX100 firms during the stock market boom of the 1990s using five alternative forecasting models. Based on I/B/E/S data from 1991 to 2000, Wallmeier (2005) tests whether the aggregated recommendations of analysts are too optimistic. The author defines the forecasting error of a firm as the difference between the predicted and the actual book equity rates of return. By removing the forecasting error, i.e. the empirically observed tendency of too optimistic stock recommendations, Wallmeier (2005) is able to outperform five alternative benchmark models. This result indicates that earnings forecasts contain information, although the market level of earnings is valued too generous by analysts. Wallmeier (2005) also investigates whether the optimistic bias diminishes over the long term and finds evidence for a decline of the bias over time. Bessler and Stanzel (2007) analyze the quality and efficiency of earnings forecasts of analysts in the German stock market. They attribute a central role in the information efficiency of capital markets to the research carried out by financial analysts. However, to guarantee an efficient allocation of resources, they argue that financial research needs to be free of conflicts of interests. Further, it is crucial that regulation supports rather than restricts financial research, e.g. by reducing potential conflicts of interest. Bessler and Stanzel (2007) analyze the German stock market

from 1995 to 2004 and find positive and biased analyst forecasts. They argue that the persistence of this bias could be explained by differences in the publicity and accounting regulation of different firms which affects the number of analysts following a company and the overall quality of available information. Additionally, behavior-oriented factors could influence the precision and overall quality of predictions, e.g. business relations with the management of the analyzed company. Benchmarking against a naïve forecast, Bessler and Stanzel (2007) use a straightforward approach to access the forecasting error of financial analysts. The authors compute the average relative forecasting error for analyst forecasts and compare the results with the error a naïve prediction would produce on a fiscal year basis. Kelly and Ljungqvist (2012) state that analysts are among the most influential information producers in financial markets. Using the closure of research departments of forty-three brokerage firms as an exogenous source of variation in the extent of analyst coverage, the authors employ a difference-in-differences approach to assess the influence of a decrease in analyst coverage on multiple liquidity proxies. Most notably, the authors take care to only consider exogenous changes to analyst coverage in order to prevent endogeneity concerns. They further investigate whether a loss of analyst coverage affects the price of the respective stock. Providing evidence that coverage terminations increase information asymmetry, they also find that information asymmetry has a substantial effect on asset prices and identify liquidity as the primary link between asset prices and information asymmetry.

In summary, the link between investment research conducted by financial analysts has been broadly studied using a multitude of different approaches. The general objective of these studies is to understand which role investment research analysts play in financial markets and how analyst coverage affects the interplay between liquidity, returns, information asymmetry and market efficiency.

We contribute to the literature in the following sections by empirically examining the influence of analyst coverage on information asymmetry and the effect of putting a concrete price on investment research as is stipulated by MiFID II. We show that analyst coverage affects information asymmetry, leading to the conclusion that the level of information asymmetry can be linked to the number of analysts covering a stock and that research analysts possess and distribute private information. We also show that the role and impact of financial analysts in European stock markets changed since the implementation of MiFID II.

3.4 Variable Definition

Efficient markets are characterized by certain desirable features: high liquidity, low information asymmetry and low transaction costs. An efficient market guarantees the efficient allocation of capital and supports the general economy. A high level of liquidity allows all market participants to trade anytime at low trading costs while

only marginally affecting asset prices. Market efficiency also requires that all available information is reflected in market prices, ruling out information asymmetry. Therefore, measuring the existence and dependencies of information asymmetry and changes in liquidity is important to understand the prevailing level of market efficiency and the impact of MiFID II.

3.4.1 Illiquidity

In the literature, there are many liquidity measures proposed, which can be divided into two major groups: price-based liquidity measures and volume-based liquidity measures. We choose to focus on the volume-based illiquidity measure proposed by Amihud (2002) as it is widely used in the literature and is straight forward to compute.

As stated above, a key characteristic of liquid markets is the possibility to trade an asset without heavily impacting the asset price, which implies that asset prices are not easily manipulated by large trades. This necessity of a liquid market makes it important to analyze, by how much the number or size of transactions affects the asset price. By setting the daily trading volume into ratio with daily returns, the illiquidity measure proposed by Amihud (2002) allows to directly assess this key market characteristic. Amihud (2002) shows that expected stock returns are an increasing function of expected stock illiquidity, supporting the illiquidity premium hypothesis. The author also shows, that the effect of expected illiquidity on expected stock returns is higher for smaller stocks.

Based on Amihud (2002) we compute the illiquidity measure as a proxy for stock illiquidity according to the following formula:

$$Illiquidity_{i,t} = \frac{|R_{i,t}|}{VOL_{i,t}} \tag{3.1}$$

$Illiquidity_{i,t}$ describes the Amihud (2002) illiquidity measure, defined as the ratio of the absolute stock return ($|R_{i,t}|$) to the trading volume in US Dollar ($VOL_{i,t}$) for stock i on trading day t. $Illiquidity_{i,t}$ directly incorporates the trading volume of a stock and can be interpreted as the change in stock price per unit of trading volume. Amihud (2002) also provides an alternative interpretation of his measure: If investors agree about the implication of new information, the stock price changes without trading, while disagreement leads to an increase in trading volume. Thus, $Illiquidity_{i,t}$ can also be interpreted as a measure of consensus belief among investors about new information (Amihud, 2002, p. 43). While not mechanically induced, the illiquidity measure should be positively correlated with the bid-ask spread from an economic perspective, as illiquidity is associated with a high bid-ask spread and a low trading volume. The return should be higher for illiquid stocks too, as investors demand a liquidity premium for holding stocks that are harder to convert into cash.

3.4.2 The Bid-Ask Spread

We compute the bid-ask spread as the difference between the best ask price $ask_{i,t}$ and the best bid price $bid_{i,t}$ for stock i at closing of trading day t:

$$Absolute\ Spread_{i,t} = ask_{i,t} - bid_{i,t} \tag{3.2}$$

We also compute the relative bid-ask spread as follows:

$$Relative\ Spread_{i,t} = \frac{ask_{i,t} - bid_{i,t}}{\frac{1}{2} * (ask_{i,t} + bid_{i,t})} \tag{3.3}$$

The denominator can be interpreted as the mid price of stock i at closing of trading day t. The relative bid-ask spread normalizes the size of the bid-ask spread by the respective share price.

The bid-ask spread resembles a widely used proxy for information asymmetry. Copeland and Galai (1983) state that a dealer or market maker sets the size of the bid-ask spread in order to maximize his own profits. By increasing the bid-ask spread, the market maker faces a trade-off between losing expected revenue from trades with liquidity (noise) traders and protecting himself from losses caused by trades against informed investors. Chung et al. (1995) propose two contrary hypotheses regarding the relationship between financial analysts and the bid-ask spread: On the one hand it might be beneficial for analysts to cover stocks with a high level of information asymmetry. Market makers would then observe the number of analysts covering a stock and increase the bid-ask spread for highly covered stocks to protect themselves against trading with informed investors. On the other hand the likelihood of insider trading will be reduced if the firm is followed by many financial analysts because analysts reveal most of the firm-specific information to investors and market makers. In this case, analyst coverage should decrease the bid-ask spread. The authors argue that determining which theory is more realistic needs to be tested empirically. Using a simultaneous equation regression analysis and I/B/E/S data, Chung et al. (1995) find empirical evidence for a bidirectional relationship between the bid-ask spread and the number of analysts covering a stock and support for the first hypothesis.

Glosten and Milgrom (1985) state that the problem of matching buyers and sellers is particularly eminent in trading shares of small companies. The authors propose a theoretical model based on the idea that even under the assumption that a market maker's cost of trading as well as his expected returns are zero, a bid-ask spread can emerge purely due to informational reasons. In the proposed model, informed and uninformed investors as well as the market maker are risk-neutral. Glosten and Milgrom (1985) describe an adverse selection problem which is based on the information disparity between traders and market makers. In their model, informed traders possess some kind of private information or superior analysis skills they trade on. The authors show that adverse selection by itself can be responsible for the existence

of a bid-ask spread. The size of the bid-ask spread depends on the proportion of informed and uninformed investors as well as the quality of private information.
Glosten (1987) more distinctively separates the bid-ask spread into two components: One part caused by exogenous costs for the market maker and the other part caused by informational asymmetry. If market makers have to compete for the best bid-ask spread, the magnitude of the bid-ask spread is mainly determined by the level of information asymmetry in the market.[4] Kim and Verrecchia (1994) show that market makers lose by trading against informed traders (Copeland and Galai, 1983 and Glosten and Milgrom, 1985). A higher level of information asymmetry causes market makers to increase the bid-ask spread in order to recover from these losses. These findings make the bid-ask spread a suitable proxy for information asymmetry. However, by decomposing the bid-ask spread into three cost components, Stoll (1989) suggests that the bid-ask spread is also determined by inventory holding costs and order processing costs[5]:

$$\text{Spread} = \text{Adverse Information Costs} + \text{Inventory Holding Costs} + \text{Order Processing Costs}$$

Based on this theoretical model, the bid-ask spread would also change when order processing costs or inventory holding costs change.
While we can not observe inventory holding costs directly, we can observe stock illiquidity by computing the illiquidity measure proposed by Amihud (2002) and use it as a proxy for inventory holding costs. Illiquidity and inventory holding costs should be highly correlated, as illiquid stocks have high inventory holding costs and vice versa.
It is well documented that less frequently traded stocks typically possess a higher bid-ask spread. Easley, Kiefer, et al. (1996) offer multiple explanations for large spreads. First, an investor may not be able to liquidate an infrequently traded stock, thus demanding an illiquidity premium for compensation. Moreover, the risk of facing an information-based trade might be higher in illiquid stocks. In this case, a larger bid-ask spread would compensate for the risk of facing an informed counterparty.
While order processing costs theoretically co-determine the size of the bid-ask spread, we argue that the impact is neglectable in modern financial markets. Since the formalization of the first theoretical models explaining the bid-ask spread (Stoll, 1989), exchanges have become more cost-efficient due to automatization of the order execution process. Hence, we choose to exclude order processing costs from our further examination. In our proposed empirical fixed effects model, it is even sufficient to

[4] In the extreme case of perfect competition, market makers would be forced to aim for zero exogenous costs to stay competitive.

[5] Stoll (1989) develops and empirically tests a model to derive the relative composition of the quoted spread. He shows that adverse information costs account for 43%, inventory holding costs for 10% and order processing costs for 47% of the total quoted spread.

assume that order processing costs remain stable during the observation period. Using fixed effects, our results are unbiased even if order processing costs are not zero, as long as they do not significantly vary during the observation period. In the case of time-constant order processing costs, we get rid of the influence of order processing costs on the bid-ask spread, as it is part of the time-constant error term.[6]

3.4.3 Stock Price Informativeness and Idiosyncratic Risk

Recent findings in the literature suggest that idiosyncratic risk is related to stock price informativeness (e.g. Morck, Yeung, and Yu, 2000 and Chen, Goldstein, and Jiang, 2006), however it is not *a priori* clear whether higher idiosyncratic risk implies more or less informed stock pricing (Durnev et al., 2003).
Roll (1988) points out the problem, suspecting a link between a low value of R^2 in asset pricing models and the existence of private information. Not finding improvements in R^2 by controlling for public news events, the author concludes that the firm-specific return variation might be caused by traders acting on private information. Roll (1988) declares that there are two possible contradictory explanations for his findings. The first explanation proposes that firm-specific price variation reflects the incorporation of private information into prices. The second explanation states that firm-specific return variation might actually reflect noise trading. Durnev et al. (2003) conclude that this problem is an empirical question and several subsequent studies emerged supporting either the first or the latter view.
Piotroski and Roulstone (2004) link idiosyncratic risk to the role of financial analysts by investigating the influence of trading and trade-generating activities of informed market participants. They measure the firm-specific, industry-level, and market-level information impounded into stock prices, as measured by stock return synchronicity:

$$SYNCH_i = ln\left(\frac{R_i^2}{1-R_i^2}\right) \tag{3.4}$$

R^2 is the part of the volatility of stock i at time t which can be explained by market volatility, i.e. the systematic risk component. Consequently, $1 - R^2$ resembles the unsystematic (idiosyncratic) risk. R^2 can be estimated via OLS by computing the goodness-of-fit of the baseline market model for each stock individually:

$$R_t = \alpha + \beta R_{m,t} + \epsilon_t \tag{3.5}$$

R_t is the daily stock return of a stock on day t and $R_{m,t}$ is the daily market return. Hence, stock return synchronicity is a measure of the extent to which market returns are able to explain firm specific returns. Piotroski and Roulstone (2004) hypothesize that informed parties (notably analysts, institutional investors and insiders)

[6] In addition, we would not expect to introduce an omitted variable bias as to our knowledge, there is no direct link between analyst coverage and order processing costs.

"contribute or disseminate information into the price formation process in different ways, which should lead to a systematic variation in stock return synchronicity with the presence or absence of these party's activities" (Piotroski and Roulstone, 2004, p. 1120). The authors state that analysts convey their information through earnings forecasts and investment recommendations, concluding that analysts can improve the price efficiency by dissemination information into the price formation process of covered firms. As there is empirical evidence that analyst activities trigger trades (e.g. Givoly and Lakonishok, 1979), Piotroski and Roulstone (2004) argue that analyst activity should cause prices to reflect this additional information, resulting in greater stock return synchronicity.
More recently, a slightly modified measure has been established in the literature. Zhu, Jog, and Otchere (2014) use the following measure for idiosyncratic risk, which is also based on the market model:

$$\Psi_i = ln\left(\frac{1 - R_i^2}{R_i^2}\right) \tag{3.6}$$

Note that $\Psi_{i,t}$ and $SYNCH_{i,t}$ are negatively correlated by definition. Both measures are log-transformed to create an unbound continuous dependent variable with a more normal distribution as both R^2 and $1 - R^2$ are bound between zero and one.
Studying the influence of analyst coverage and MiFID II on idiosyncratic risk bears the advantage of obtaining a measure that is linked to information asymmetry and which is derived without a direct link to the bid-ask spread. This allows us to add to the literature regarding the role of idiosyncratic risk as a proxy for informativeness of stock prices and also serves as a robustness check for our empirical results shown in Section 3.6. Following the argumentation of Piotroski and Roulstone (2004), we would expect that an increase in analyst coverage leads to a decrease in the idiosyncratic risk measure Ψ. We hypothesize that MiFID II increased competition between analysts to provide economically valuable information and therefore amplifies this effect. Considering the role of financial analysts disseminating private information it is worth to note that information may not be strictly private or public. Initially private information becomes "more public" as more uninformed investors become informed by either learning about the information from trading or by directly buying the information when the costs to access the information decrease over time.[7]

3.5 Data

We use Thomson Reuters Datastream to obtain a data set consisting of daily data of publicly traded stocks covered by the Thomson Reuters Global Equity Index from

[7] A similar reasoning is given by Chung et al. (1995), who argue that an analyst will first inform his favored clients about a new piece of information before putting it in a newsletter and finally disclosing it publicly (Chung et al., 1995, p. 1028).

each trading day in the months October 2017, November 2017, February 2018 and March 2018. We examine the constituents of the Thomson Reuters Global Equity Index in order to obtain a representative sample of the stock market universe. The Thomson Reuters Global Equity Index consists of more than 8,000 publicly traded companies across 51 countries and is designed to serve as a broad market benchmark and to track the performance of liquid equities worldwide. For our analysis we use the companies listed in this index. A major advantage of this index is that it is designed to incorporate more than 99.50% of the market capitalization of all liquid stocks.[8]

We use the International Securities Identification Number (ISIN) to determine in which country a company is listed and drop all companies from our sample, which are not listed in EU28 member states or the United States. We further exclude all financial services companies identified by the respective 2-digit Standard Industrial Classification Code (SIC) from our sample. Consequently, all remaining companies are either listed in EU28 member states or the US, the first defining our treatment group and the latter defining our control group.

Similar to Chung et al. (1995) and Easley, O'Hara, and Paperman (1998), we use the number of earnings-per-share (EPS) estimates reported by the Institutional Brokers' Estimate System (I/B/E/S)[9] in order to assess the investment research coverage of a stock. We find that if a stock is covered by an analyst, EPS is one of the most reported variables and the number of EPS estimates should therefore provide a reasonable proxy for the amount of investment research coverage of a company. Following the literature, we use the number of EPS estimates reported as a proxy for investment research coverage (e.g. Piotroski and Roulstone, 2004). We further gather information about the market capitalization of a company, daily stock trading volume and the market closing bid and ask prices using Thomson Reuters Datastream.

Our final sample consists of four months of daily data covering 2,927 companies, thereof 1,281 European companies and 1,646 companies listed in the United States. The five largest economies (United Kingdom, Sweden, France, Germany and Italy) account for more than half of all European observations and the United States account for more than half of the total observations in the sample. A detailed breakdown of our sample, divided by country, can be found in Table 3.1.

[8]Equity closed funds, equity derivatives, exchange traded funds, some units, investment trusts and Limited Partnerships/Master Limited Partnerships are excluded. Further, in order to be eligible, constituents must have a market capitalization of at least USD 150mn. This restriction should not raise concerns, as the median company listed in the German small cap index SDAX already has a market value which is four times higher.

[9]Thomson Reuters I/B/E/S Estimates provides sell side analyst estimates for a company's future quarterly and annual financial performance including real-time and historical estimates from 900 contributors across 100 developed and emerging markets, totaling over 13,000 individual analysts. Thomson Reuters maintains a close relationship with sell-side contributors who must pass a rigorous screening process before their research gets accepted. Coverage includes 99% of MSCI Asia, 98% of MSCI World and 100% of S&P500, resulting in a coverage total of over 22,000 active companies (+ 20,000 inactive) across over 87 countries (https://developers.thomsonreuters.com/content/ibes-estimates).

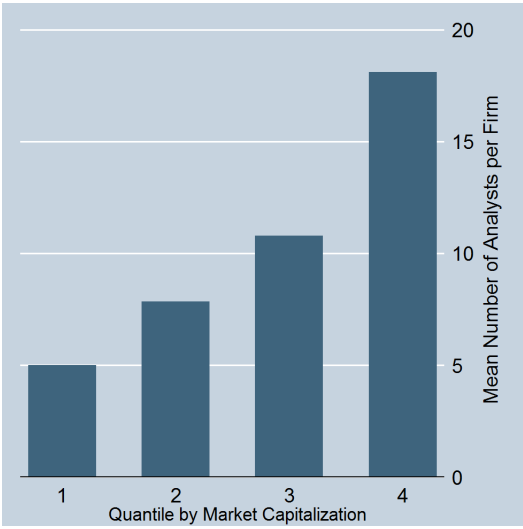
As we are mainly interested in persistent effects of MiFID II, we focus on comparing daily trading data from before and after the implementation of MiFID II.[10] Doing so, we face a trade-off between minimizing the potential threat of introducing measurement errors by measuring effects not caused by MiFID II and making sure that EU markets had enough time to adjust to the new market conditions. We increase the robustness of our analysis by estimating results for two different time frames. Table 3.2 provides summary statistics for the final sample used for our subsequent empirical analysis in Section 3.6. It can be seen that the average characteristics of the treatment and control group barely differ.

3.5.1 Descriptive Statistics

In this section, we highlight several properties to get an insight into statistical dependencies in our sample. First, we examine the dependency between the market size of a company and the number of analysts following the company's stock by computing the average number of analysts covering a firm dependent on the market capitalization (in USD). We find research coverage to be highly correlated with firm size (Figure 3.1). An evident explanation would be that larger companies generally gain more market exposure and publicity, therefore demand for investment research and consequently investment research coverage is larger for those firms. Moreover, certain companies might be too small to attract large investors as there are just not enough shares available to purchase making these stocks uninteresting. Figure 3.1 depicts the mean number of analysts per stock for quartiles of market capitalization with Q_1 being composed of companies with the smallest market capitalization. Interestingly, investment research coverage increases almost steadily between the three smallest quantiles of market capitalization (from Q_1 to $Q2$: + 2.96 analysts, from Q_2 to $Q3$: + 2.92 analysts), while a company in the highest quantile is covered by 7.09 more analysts than a company in the second largest quantile on average. A possible explanation for this finding might be that there exist some "must have" companies which most of the analysts cover, independent of their own research focus.

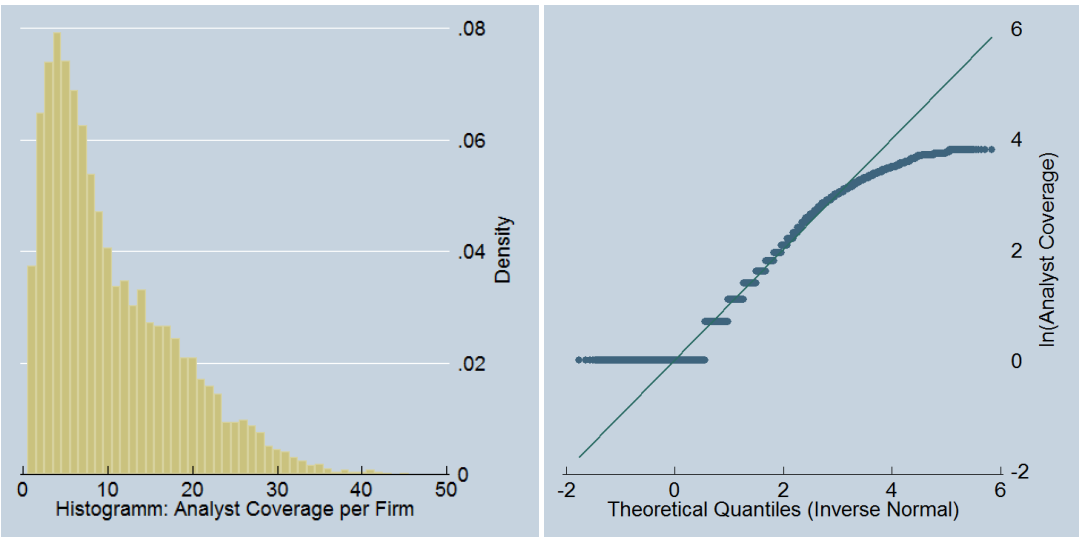
We also examine the empirical distribution of research coverage and find that on average, a firm is covered by 10.41 analysts across our data. We further observe, that the distribution of analyst coverage reaches a maximum at four analysts. Based on the QQ-plot shown in Figure 3.2, analyst coverage seems to be reasonably well fitted by a log-normal distribution. It appears that there are only few companies which

[10] Analyzing short term effects of MiFID II could be achieved by conducting an event study around the implementation day. However, this approach would be prone to confounding events around the fiscal year change. However, we do not find confounding international events during the estimation windows, which arguably affect the US and the EU in a different way.



Notes: This figure shows the average number of analyst coverage for four quantiles of market capitalization (Q_1: 0% – 25%, Q_2: 25% – 50%, Q_3: 50% – 75%, Q_4: 75% – 100%). Market capitalization is measured in US Dollar for all 2,927 observed firms.

FIGURE 3.1: Research Coverage dependent on Market Capitalization

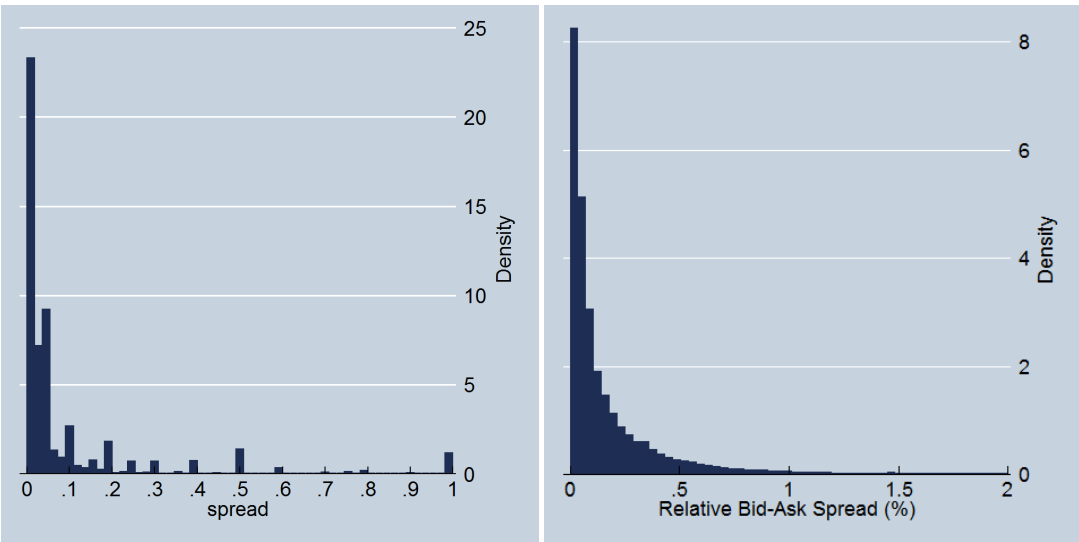


Notes: This figure shows the empirical distribution of analyst coverage per firm (left) and the QQ-plot of the logarithmized analyst coverage (right).

FIGURE 3.2: Empirical Distribution of Research Coverage per Firm

are covered by a high amount of analysts.[11] These findings also provide some support for the hypothesis that there might be some saturation regarding the amount of available research coverage, i.e. an analyst might face a trade-off between the difficulty to unveil new information about a stock which has already been analyzed by competing analysts on the one hand and the investor demand for his research on the other hand. In this case, the analyst has to weigh the positive effect of covering a highly sought after stock against the increased difficulty to unveil private information about this stock.

Figure 3.3 depicts the empirical distribution of the bid-ask spread and the relative bid-ask spread in our sample, which we compute based on Formula 3.2 and Formula 3.3. While the majority of observed spreads is rather small, we find peaks at round figures (0.1, 0.2, ..., 1.0). The peaks are not easily explained from an economic perspective assuming rational agents. We hypothesize that this pattern stems from a human tendency for round figures serving as an anchor point on a discrete price grid.[12] While a deeper analysis of this phenomenon and its compatibility with the efficient market hypothesis is certainly an interesting research topic on its own, it is not the focus of this study and shall not be further discussed. We are interested in the



Notes: This figure shows the empirical distribution of the bid-ask spread (left) and the relative bid-ask spread (right).

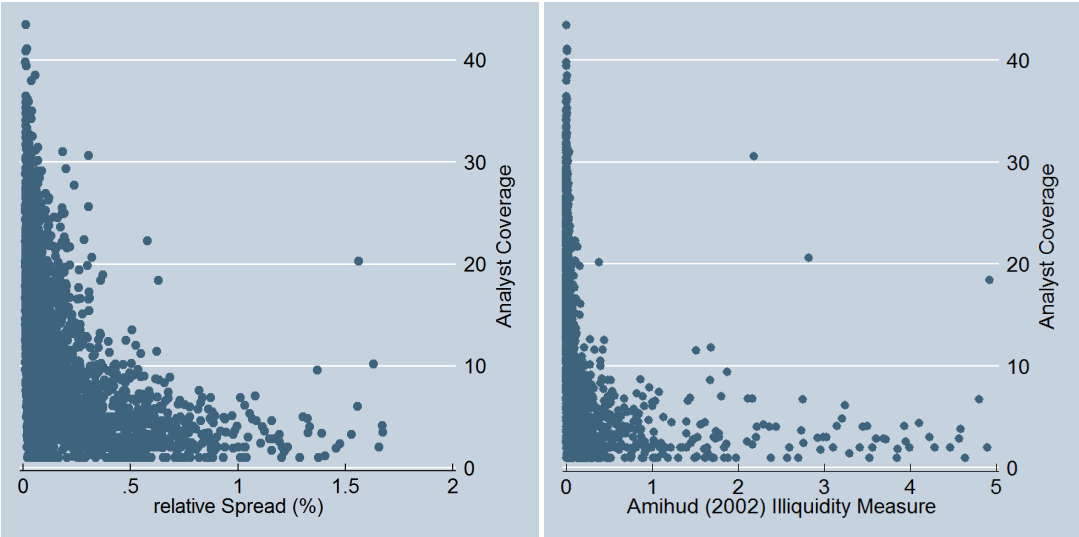
FIGURE 3.3: Sample Distribution of Stock Spreads

general interplay between analyst coverage, the bid-ask spread and stock liquidity.

[11] We find that the American cloud-based software company SALESFORCE.COM to have the highest analyst coverage overall (45 analysts). ADIDAS, the German manufacturer of sports goods, has the highest coverage across all EU countries (39 analysts).

[12] The occurrence of peaks could also be amplified by the definition of a minimum tick size for certain illiquid stocks equal to the level of the peak. In the recent past there has been a race to the bottom regarding the price granularity however, indicating that this explanation can not be the sole reason for our observation.

We proxy the illiquidity of a stock by computing the illiquidity measure proposed by Amihud (2002) based on Formula 3.1. The average relative bid-ask spread per stock in our sample is plotted against the average number of analysts in Figure 3.4 revealing a negative link. A similar pattern arises between the illiquidity measure and analyst coverage, indicating that highly covered stocks have a lower bid-ask spread and are more liquid than stocks with low analyst coverage.[13] In summary,



Notes: This figure shows a plot of the relative bid-ask spread against analyst coverage (left) and a plot of the Amihud (2002) illiquidity measure against analyst coverage (right).

FIGURE 3.4: Analyst Coverage and the Relative Bid-Ask spread/the Amihud (2002) Illiquidity Measure

our descriptive analysis shows that there seems to be a link between analyst coverage and market size, illiquidity and the bid-ask spread. Especially market capitalization seems to be highly correlated with analyst coverage, which makes it mandatory to control for this factor in our empirical analysis.

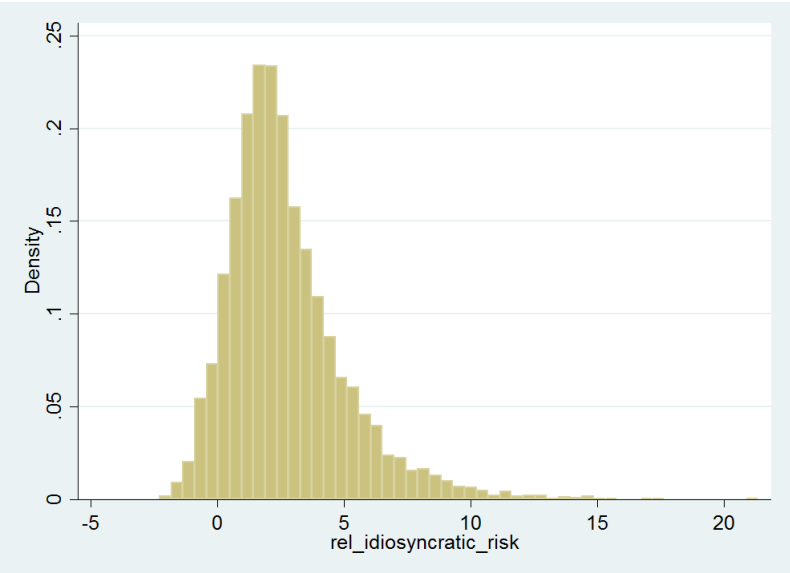
Finally, we compute the idiosyncratic risk measure Ψ_i for each stock in our sample in a three-step procedure: First, we compute the daily market return as the market value-weighted return of all stocks included in the Thomson Reuters Global Equity Index. Second, we compute the R_i^2 for each stock i in our sample by regressing the market model for each stock in our sample based on Equation 3.5. As R_i^2 gives a measure for the variation in stock return explained by variation in market returns, $1 - R_i^2$ describes the idiosyncratic risk. Following the literature, we compute Ψ_i by

[13]Note that we can not derive any causal relation from these findings. Further, we do not control for market size, which we have shown to be correlated with analyst coverage as well. We find a positive correlation of 0.35 between the relative bid-ask spread and illiquidity in our data.

transforming the idiosyncratic risk:

$$\hat{\Psi}_i = ln\left(\frac{1-\hat{R}_i^2}{\hat{R}_i^2}\right) \tag{3.7}$$

Figure 3.5 shows the sample distribution of $\hat{\Psi}_i$. As we are interested in changes to Ψ_i associated with the introduction of MiFID II, we compute Ψ_i for different sub periods before and after MiFID II. Note that a value of $\hat{\Psi}_i = 0$ indicates the idiosyncratic risk of a stock to be responsible for 50% of the total stock return variation.



Notes: This figure shows the empirical distribution of the idiosyncratic risk measure Ψ, where $\hat{\Psi}_i = ln((1-\hat{R}_i^2)/\hat{R}_i^2)$. $\hat{R}_i^2$ denotes the determination coefficient from regressing the stock return of stock i on the market return, i.e. $\hat{R}_i^2$ gives the return variation explained by the market model, commonly referred to as the systematic risk component of stock i. Note that a hypothetical value of $\hat{R}_j^2 = 50\%$ implies $\hat{\Psi}_j = 0$.

FIGURE 3.5: Empirical Distribution of the Idiosyncratic Risk Measure Ψ

3.6 Empirical Analysis

Our empirical analysis is structured as follows: First, we compute the difference-in-differences estimator for five market shaping factors (analyst coverage, illiquidity, trading volume, bid-ask spread and idiosyncratic risk) to estimate the effect of MiFID II on these factors. Second, we analyze the impact of analyst coverage on the bid-ask spread and idiosyncratic risk by employing a fixed effects regression model. We include interaction terms in our regression to test, whether MiFID II has an observable impact on the relationship between analyst coverage, information asymmetry and idiosyncratic risk.

3.6.1 Difference-in-Differences Approach

Using the following regression model, we estimate the difference-in-differences estimator for our variables of interest:

$$y_i = \alpha + \beta treat_i + \gamma after_i + \tau(treat_i \times after_i) + u_i \tag{3.8}$$

where y_i is the measurement value of the dependent variable for stock i. $treat_i$ indicates, whether company i is based in the EU or the US. We identify the jurisdiction in which stock i is listed based on the ISIN. $after_i$ indicates, whether the observation stems from before or after the implementation of MiFID II. Consequently, τ resembles the difference-in-differences estimator. We use daily data from October 2017, November 2017, February 2018 and March 2018 surrounding the implementation of MiFID II on 3rd of January 2018. We choose the respective time frames, considering that by increasing the estimation window, we face a trade-off between measuring a persisting long-term effect and the risk of measuring the effect of confounding events. Our sample consists of 1,281 listed companies that are affected by MiFID II (treatment group) and 1,641 US-based companies (control group) during the same time window (see Table 3.2). Table 3.3 shows the estimation results for the mean treatment effect τ of the implementation of MiFID II on European stocks. We further compute the difference-in-differences estimator for three different subgroups based on quantiles for analyst coverage in order to study whether MiFID II impacts stocks with low or high coverage differently. We find our results not to be largely driven by a specific subgroup. Contrary to our expectations, the number of analysts covering a stock did not significantly decrease due to the implementation of MiFID II.[14] We find the largest decrease of analyst coverage by -0.41 analysts per stock across stocks with a high coverage but this effect is not statistically significant. We derive that the mere consumption of research did not change significantly with the introduction of MiFID II in the short run.

We observe that MiFID II led to a small decrease in overall stock liquidity. As expected and motivated in Section 3.2, this effect is most eminent for stocks with low coverage. While observable in both Panels (Table 3.3), the effect seems to decline over time. We further find a decrease in trading volume which is significant at the 1% level for stocks with low coverage across both time windows.

Observing the bid-ask spread we find a significant effect across all subgroups and time frames except for highly covered stocks in the six months time frame. This effect does not seem to be caused by a change in illiquidity, as we find an effect of MiFID II on the bid-ask spread of almost the same magnitude when we specifically control for changes in illiquidity. Investigating the surprising unambiguousness of

[14] The difference-in-difference approach is especially useful to determine short-term effects. In Table 3.6 we show that the research coverage decreased substantially in EU markets since the implementation of MiFID II.

this finding, we hypothesize that changes to the minimum tick size, which were also introduced with MiFID II might drive this effect. Article 49 of MiFID II introduces a new tick size regime, which defines the minimum tick sizes for stocks based on the average daily number of transactions and the stock price with the aim of not constraining spreads by a too large tick size. It is important to note that this adjustment only affects the granularity of the pricing grid on which limit and market orders can be placed on. To our knowledge, there is no reason to believe that this change leads to a change of the average bid-ask spread. The French financial supervision Autorité des marchés financiers (AMF) issued a report shortly after the implementation of MiFID II stating that an increase in stock spreads can be observed which they state can be traced back to changes in the minimum tick size of affected stocks (AMF, 2018).[15] While our results confirm this observation it also highlights the importance of controlling for MiFID II in our subsequent multivariate regression model. Another cause of an increase in bid-ask spreads which should not be overlooked is a potential change in the behavior of investors: They might be more reserved in sharing information gathered from research reports that they consciously paid for.
Computing the difference-in-differences estimator for the measure of the stock specific idiosyncratic risk (Ψ_i) we find that the idiosyncratic risk increased significantly since the implementation of MiFID II.[16] The effect is highly significant across all subsamples with an increase in the average Ψ_i ranging from 1.43 to 3.22. This change is quite substantial, considering that the average value for Ψ_i for EU stocks before MiFID II is 2.19 in our sample.

3.6.2 Multiple Regression on the Bid-Ask Spread

Next, we want to study the relationship between analyst coverage and information asymmetry. We do so by estimating the following baseline regression model:

$$\begin{aligned} log(Spread)_{i,t} = \beta_0 + \beta_1 log(Analyst\ Coverage)_{i,t} + \beta_2 Illiquidity_{i,t} \\ + \beta_3 log(Market\ Capitalization)_{i,t} \\ + \beta_4 MiFID\ II \times log(Analyst\ Coverage)_{i,t} + c_i + u_{i,t} \end{aligned} \quad (3.9)$$

$log(Spread)_{i,t}$ describes the logarithmic form of the bid-ask spread of stock i at the end of trading day t and serves as proxy for the level of information asymmetry. Our main interest lies in the coefficient of $log(Analyst\ Coverage)$, which yields insight into the influence of analyst coverage on the asymmetric information component of

[15] Consulting the European Securities and Markets Authority (ESMA), Deutsche Börse Group also states that they find an increase in spreads for equities traded at Frankfurter Wertpapierbörse after MiFID II (ESMA, 2018).

[16] We use an estimation period of one month of daily trading data to compute Ψ_i before and after MiFID II.

the bid-ask spread.[17] Note that we include both the bid-ask spread and the analyst coverage in their logarithmic form[18]. This procedure allows us to estimate the elasticity of the bid-ask spread with respect to analyst coverage. We include the illiquidity measure proposed by Amihud (2002) to capture changes in the bid-ask spread caused by an increase in inventory holding costs. An increase in the illiquidity measure should *ceteris paribus* lead to an increase in the bid-ask spread as it increases the inventory holding costs of a stock. As shown in Section 3.5.1, analyst coverage is highly dependent on the market capitalization of a company. To account for this size effect, we include the market value of a company as an explanatory variable in our regression model. Note that the error term c_i captures time-constant order processing costs. Using fixed effects allows us to completely get rid of c_i. $MiFID\ II$ represents a dummy variable indicating whether the observation is affected by MiFID II, i.e. $MiFID\ II$ equals 1, if the observation stems from an EU stock in the year 2018 and 0 otherwise. We include $MiFID\ II$ for two distinct reasons: First, we only get an unbiased estimator for the coefficient of our interaction term between $MiFID\ II$ and $log(Analyst\ Coverage)$ by including both variables in their primal form, and second, we are thereby able to capture the effect of MiFID II on the bid-ask spread. Additionally, we control for year and country fixed effects. *Year* 2018 indicates, whether the observation stems from the calendar year 2017 or 2018. We include this variable in order to control for seasonal changes in the level of information asymmetry, which else would be falsely attributed towards the implementation of MiFID II. Our regression results are shown in Table 3.4.

We find that an increase in analyst coverage leads to a statistically significant decrease in the bid-ask spread. The effect is also significant at the 1% level in all subgroups. Across our sample, an increase in the analyst coverage by 10% decreases the average bid-ask spread by 2.6%, indicating that the observed effect is also of economic relevance. This finding supports our hypothesis of an increase in analysts reducing the costs of information leading to a more efficient distribution of private information. The interaction term between $MiFIDII$ and $log(Analyst\ Coverage)$ is also highly significant. This finding shows that the influence of analyst coverage on the information asymmetry component of the bid-ask spread significantly increased with the implementation of MiFID II. As opposed to medium and highly covered firms, we find that the introduction of MiFID II reduces the impact of an increase in analyst coverage on the bid-ask spread for companies with low coverage, indicating that MiFID II impacts companies with low coverage in a different way. Our result could be explained by an increased competition between investment research analysts having to proof the economic value of their work or analysts focusing their

[17] Variation in $log(Spread)_{i,t}$ effectively proxies variation in the level of information asymmetry. We control for illiquidity effects regarding the bid-ask spread by including $Illiquidity_{i,t}$ in our regression model.

[18] We follow Easley, O'Hara, and Paperman (1998) by using the logarithmized value of analyst coverage.

resources on producing high-quality research about companies with high coverage. As expected, we find an increase in illiquidity to be associated with a significant increase in the bid-ask spread. We find this effect across all subgroups to be of equal magnitude. Examining the economic significance, an increase in the illiquidity measure by 0.1 (which equals approximately one standard deviation) increases the average bid-ask spread by a mere 0.17%.
With an R^2 ranging from 28.3% to 31.0%, the explanatory power of our regression results is rather high. Further, many estimated coefficients show the expected sign boosting the confidence regarding the validity of our proposed model specification.

3.6.3 Multiple Regression on Idiosyncratic Risk

We estimate $\Psi_{i,t}$ on a monthly basis for each stock in our regression model by first regressing the return of stock i at trading day t on the market return on trading day t for the months surrounding the implementation of MiFID II: October 2017, November 2017, February 2018 and March 2018. We compute the market return as the market value weighted average return of stocks included in the Thomson Reuters Global Equity Index as described in Section 3.5.1. We use the resulting $R^2_{i,month}$ to compute $\Psi_{i,t}$ for each stock in our sample. In a second step, we study the impact of analyst coverage on $\Psi_{i,t}$:

$$\begin{aligned}\Psi_{i,t} = \beta_0 &+ \beta_1 log(Analyst\,Coverage)_{i,t} + \beta_2 Illiquidity_{i,t} \\ &+ \beta_3 log(Market\,Capitalization)_{i,t} \\ &+ \beta_4 MiFID\,II_{i,t} \times log(Analyst\,Coverage)_{i,t} + c_i + u_{i,t}\end{aligned} \tag{3.10}$$

While the coefficient of $log(Analyst\,Coverage)_{i,t}$ does not have a straightforward economic interpretation, we can derive a statement by observing its sign: An increase in analyst coverage leads to a significant decrease in the amount of idiosyncratic risk of a stock, indicating that more analysts covering a stock lead to a higher level of synchronicity between stock market returns and individual stock returns. This finding is consistent with the results of Piotroski and Roulstone (2004) who find a positive link between analyst activities and stock price synchronicity as well. Piotroski and Roulstone (2004) attribute their findings to analysts increasing the informativeness of prices through intra-industry information transfers.
We further find higher illiquidity to be associated with a higher level of idiosyncratic risk. A stock that is traded less frequently might be more separated from the general market development. Moreover, the information transfer from informed traders via trading and the ability of uninformed traders to learn about the intrinsic value of a stock from price movements might be limited, if less trading takes place. Interestingly, this effect can not be observed for highly covered stocks suggesting a link between analyst coverage and illiquidity. This effect would be consistent with the

consideration of analysts functioning as a company promoter, making a stock more popular across investors by issuing company specific research.
Regarding the coefficient of the interaction term $MiFID\ II \times log(AnalystCoverage)$ we find that the magnitude of the impact of analyst coverage on idiosyncratic risk increases with the implementation of MiFID II. This effect seems to be most apparent for highly covered stocks and is not observable for stocks exhibiting low to medium levels of analyst coverage. We also find that market capitalization significantly decreases idiosyncratic risk. This effect could be explained by small companies being more specialized in the products or services they offer, while large conglomerates tend to possess a more diversified portfolio and operate in many different industries at the same time. Subsequently, general market variation might be better in explaining returns of larger companies than smaller companies. The coefficient of $MiFID\ II$ provides further evidence supporting our findings from our difference-in-differences analysis: We observe a significant increase in the overall idiosyncratic risk caused by the implementation of MiFID II. As discussed in Section 3.4.3, an increase in the idiosyncratic risk can either be interpreted as an increase or a decrease in information asymmetry with two different strains in the existing literature supporting either the first or the latter view. For this reason, we analyze the behavior of the bid-ask spread based asymmetric information measure and idiosyncratic risk separately. While we do not directly try to assess the link between idiosyncratic risk and information asymmetry, our results suggest that a lower level of idiosyncratic risk can be associated with a lower level of information asymmetry. We base our view on the regression results shown in Table 3.4 and Table 3.5. An increase in analyst coverage seems to decrease both information asymmetry as well as idiosyncratic risk.
Our results also provide empirical evidence for a more fundamental question: What is the role of an analyst in financial markets? We show that financial analysts play an important role in disseminating information. We further find evidence which supports the theoretical model proposed by Grossman and Stiglitz (1980) in which private information can only be observed by investors, if a specific price is paid. We argue that an increase in the number of analysts increases the competition between analysts which in turn leads to a decrease in the price for private information making it easier for investors to access the information. This reduces the discrepancy of the information level between market participants. We observe the resulting decrease in asymmetric information costs by a reduction in the bid-ask spread.

3.6.4 Summary Statistics on Medium-Term Effects on Research Coverage

First explorative studies begin to emerge considering the medium-term effects of MiFID II on research coverage, e.g. the 11th Annual IR Survey conducted by the company CITIGATE DEWE ROGERSON finds that sell-side analyst research declined

for 52% of UK based companies and 39% for EU companies excluding UK (Citigate Dewe Rogerson, 2019).
While our study focuses on short-term effects surrounding the implementation of MiFID II – not least to avoid attributing effects of confounding events to the effect of MiFID II – we provide descriptive statistics on the general development of the market for investment research following the implementation of MiFID II (Table 3.6). On a year-on-year basis, we find that the overall coverage decreased from a total of 12,353 available earnings forecasts for 1,271 companies at the beginning of 2018 to 11,387 available earnings forecasts for the same companies in 2019, which amounts to an overall decrease in coverage by 7.82% since implementation of MiFID II. Moreover, 3.23% of all companies and 5.91% of companies with already low coverage (25%-Quantile) in 2018 completely lost coverage. However, it is tough to link this decline in coverage solely to the impact on MiFID II without further analysis. In addition, variation of analyst coverage seems to be low for companies with low coverage, as the coverage did not change at all for 41.90% of companies, which already had low coverage in 2018.

3.7 Conclusion

How analysts affect stock markets is of widespread interest and has been widely discussed in the literature. We add to the literature by studying the impact of a regulatory change on analyst behavior. With the introduction of MiFID II, research services conducted by financial analysts have to be paid for explicitly for the first time, creating a unique opportunity to study the role of financial analysts. We hypothesize that analysts face enhanced competition, which should have a measurable effect on the bid-ask spread of a covered stock. Using a difference-in-differences approach, we first show that MiFID II had a substantial impact on European financial markets. We show that MiFID II led to an increase in the overall bid-ask spread as well as the stock specific idiosyncratic risk.
We further show that an increase in analyst coverage reduces the bid-ask spread of a stock. Our results are robust when we control for factors potentially influencing the bid-ask spread. By controlling for liquidity, we capture the variation of the bid-ask spread caused by a change in the implicit costs of asymmetric information and find analysts to reduce asymmetric information. We find that this effect is amplified by MiFID II.
With a total of 2,927 observed firms, we are confident that our results can be generalized to a certain level. Nevertheless, more studies have to be conducted focusing on long-term and industry specific effects to understand the full impact of MiFID II on European financial markets. Another promising chance to study the effect of a regulatory change affecting analyst behavior in the future will certainly arise as

SEC has to decide on how to proceed after their 30 month non-action relief issued in October 2017 concerning US research providers.

TABLE 3.1: Summary Statistics by Country

Notes: This table shows summary statistics for four months of daily data surrounding the implementation of MiFID II (October 2017, November 2017, February 2018 and March 2018). Analyst Coverage is proxied by the number of available earnings-per-share estimates per firm in the I/B/E/S database.

Country	**Firm-Day Observations**			**Analyst Coverage**		
	Freq.	Percent	Cum.	Min	Max	Mean
European Countries						
AT	2,175	0.87	0.87	2	19	7.19
BE	3,480	1.40	2.27	1	31	6.58
CZ	174	0.07	2.34	2	11	6.09
DE	13,875	5.57	7.92	1	40	12.49
DK	3,247	1.30	9.22	1	32	10.60
ES	4,894	1.97	11.19	1	34	14.15
FI	5,216	2.10	13.28	1	28	8.04
FR	14,470	5.81	19.09	1	32	10.10
GB	22,322	8.97	28.06	1	33	12.14
GR	1,698	0.68	28.74	1	17	6.01
HU	348	0.14	28.88	3	9	6.33
IE	2,926	1.18	30.06	3	31	13.64
IT	7,596	3.05	33.11	1	29	7.55
LU	1,467	0.59	33.70	2	25	11.04
MT	261	0.10	33.80	1	5	2.67
NL	5,112	2.05	35.86	1	31	11.49
PL	5,099	2.05	37.91	1	17	5.22
PT	1,131	0.45	38.36	2	24	9.84
SE	14,525	5.83	44.19	1	30	5.56
United States of America						
US	138,920	55.81	100.00	1	45	10.90
Total	248,936	100.00				

TABLE 3.2: Summary Statistics for Analyst Coverage

Notes: This table shows summary statistics about analyst coverage derived from four months of daily data surrounding the implementation of MiFID II (October 2017, November 2017, February 2018 and March 2018) for 1,646 US and 1,281 EU firms.

Jurisdiction	**EU**		**US**	
	Mean	Std. Dev.	Mean	Std. Dev.
Analyst Coverage (#)	9.81	7.38	10.90	7.79
Quantile 1 – LOW	3.17	1.44	3.61	1.47
Quantile 2 – MEDIUM	8.38	2.12	8.51	2.22
Quantile 3 – HIGH	18.99	4.91	19.76	6.30
Firms (#)	1,281		1,646	
Firm-Day Observations (#)	110,016		138,920	

TABLE 3.3: The Effect of MiFID II on European Financial Markets

Notes: This table shows the difference-in-differences estimator for multiple variables of interest surrounding the implementation of MiFID II. The control group consists of 1,646 US firms (not affected by MiFID II) and the treatment group consists of 1,281 EU firms (affected by MiFID II). Illiquidity is measured by the illiquidity measure proposed by Amihud (2002). The estimator is multiplied by 10^6. The estimator for the relative bid-ask spread is multiplied by 100 for readability.

Analyst Coverage	**(ALL)**		**(LOW)**		**(MEDIUM)**		**(HIGH)**	
	Mean of Diff-in-Diffs (Treatments vs Controls)	t-statistic	Mean of Diff-in-Diffs (Treatments vs Controls)	t-statistic	Mean of Diff-in-Diffs (Treatments vs Controls)	t-statistic	Mean of Diff-in-Diffs (Treatments vs Controls)	t-statistic
Panel A: Four Months (M−2 vs. M+2)								
Analyst Coverage	−0.03	−0.07	0.16	1.20	0.03	0.15	−0.09	−0.17
Illiquidity	0.01***	4.24	0.03***	3.39	0.01*	1.87	0.00	0.09
log(Trading Volume)	−0.29***	−3.05	−0.38***	−2.58	−0.20	−1.50	−0.09	−0.71
rel. Bid-Ask Spread	0.07***	5.29	0.09***	3.16	0.08***	4.01	0.03***	3.45
rel. Bid-Ask Spread$^{(1)}$	0.06***	5.21	0.09***	3.19	0.07***	3.65	0.03***	3.65
Idiosyncratic Risk (Ψ)	1.70***	16.82	1.65***	9.24	1.45***	8.40	1.93***	11.42
Panel B: Six Months (M−3 vs. M+3)								
Analyst Coverage	−0.34	−0.84	−0.10	−0.71	−0.18	−0.85	−0.41	−0.77
Illiquidity	0.01**	2.41	0.01*	1.70	0.00	0.90	0.00	1.47
log(Trading Volume)	−0.35***	−3.68	−0.59***	−3.90	−0.15	−1.13	−0.21*	−1.74
rel. Bid-Ask Spread	0.03**	2.50	0.06**	2.06	0.04**	2.06	0.00	−0.01
rel. Bid-Ask Spread$^{(1)}$	0.03**	2.27	0.06**	2.08	−0.04*	1.92	0.00	−0.43
Idiosyncratic Risk (Ψ)	1.53***	14.09	3.22***	32.98	1.43***	7.66	1.85***	10.36

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, $^{(1)}$Including a control variable for liquidity.

TABLE 3.4: The Impact of Analyst Coverage on Information Asymmetry

Notes: This table shows fixed-effects regression results on $log(Spread)$ for the whole sample (ALL) and three sub-samples based on quantiles of analyst coverage (LOW, MEDIUM, HIGH). We measure *Illiquidity* by computing the illiquidity measure proposed by Amihud (2002). *Market Capitalization* is measured in US Dollar.

	log(Spread)			
Analyst Coverage	**(ALL)**	**(LOW)**	**(MEDIUM)**	**(HIGH)**
Variables of Interest				
log(Analyst Coverage)	-0.26^{***}	-0.12^{***}	-0.25^{***}	-0.42^{***}
	(-65.10)	(-10.56)	(-12.48)	(-22.62)
Illiquidity	1.70^{***}	1.92^{***}	1.63^{***}	2.22^{***}
	(37.91)	(36.12)	(14.02)	(8.05)
Market Capitalization	0.00^{***}	-0.00	-0.00^{***}	0.00^{***}
	(10.26)	(-1.03)	(-23.88)	(21.27)
MiFID II×log(Analyst Coverage)	-0.03^{***}	0.08^{***}	-0.14^{***}	-1.16^{***}
	(-3.37)	(3.83)	(-3.02)	(-26.34)
Control Variables				
MiFID II	0.02	-0.07^{**}	0.32^{***}	3.22^{***}
	(1.28)	(-2.32)	(3.29)	(24.98)
Year [2017=0;2018=1]	0.06^{***}	0.05^{***}	0.05^{***}	0.08^{***}
	(8.23)	(3.81)	(4.37)	(6.55)
Country [US=0;EU=1]	1.65^{***}	1.56^{***}	1.68^{***}	1.71^{***}
	(213.19)	(110.15)	(128.55)	(134.15)
Constant	-3.50^{***}	-3.64^{***}	-3.39^{***}	-3.11^{***}
	(-365.44)	(-219.35)	(-81.16)	(-57.22)
Observations	241,598	80,558	80,277	80,763
R^2	31.0%	28.3%	30.9%	29.8%

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 3.5: The Impact of Analyst Coverage on Idiosyncratic Risk

Notes: This table shows fixed-effects regression results on the idiosyncratic risk measure Ψ for the whole sample (ALL) and three sub-samples based on quantiles of analyst coverage (LOW, MEDIUM, HIGH). We measure *Illiquidity* by computing the illiquidity measure proposed by Amihud (2002). *Market Capitalization* is measured in US Dollar.

	Relative Idiosyncratic Risk ($\Psi_{i,month}$)			
Analyst Coverage	**(ALL)**	**(LOW)**	**(MEDIUM)**	**(HIGH)**
Variables of Interest				
log(Analyst Coverage)	-0.19^{***}	-0.05^{***}	-0.24^{***}	-0.29^{***}
	(-31.43)	(-2.91)	(-7.85)	(-10.31)
Illiquidity	1.12^{***}	0.92^{***}	0.97^{***}	-0.59
	(16.71)	(11.69)	(5.53)	(-1.43)
Market Capitalization	-0.00^{***}	-0.00^{***}	-0.00^{***}	-0.00^{***}
	(-40.01)	(-29.91)	(-28.23)	(-33.39)
MiFID II$\times$*log(Analyst Coverage)*	-0.15^{***}	0.03	0.09	-0.53^{***}
	(-12.89)	(0.99)	(1.35)	(-8.20)
Control Variables				
MiFID II	1.90^{***}	1.39^{***}	1.24^{***}	3.43^{***}
	(67.99)	(31.27)	(8.52)	(17.92)
Year [2017=0;2018=1]	-2.25^{***}	-1.93^{***}	-2.14^{***}	-2.64^{***}
	(-208.20)	(-92.94)	(-121.62)	(-148.79)
Country [US=0;EU=1]	0.10^{***}	0.28^{***}	0.22^{***}	-0.18^{***}
	(8.64)	(13.33)	(11.32)	(9.69)
Constant	3.58^{***}	3.39^{***}	3.77^{***}	4.07^{***}
	(250.37)	(137.39)	(59.84)	(50.36)
Observations	242,170	80,742	80,578	80,850
R^2	21.5%	16.2%	20.6%	26.4%

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 3.6: Medium-Term Development of Analyst Coverage since the Introduction of MiFID II (2018 – 2019)

Notes: This table shows the change in analyst coverage between March 2018 to March 2019 for EU stocks affected by MiFID II. We measure the change in analyst coverage be computing the change in the amount of available earnings-per-share estimates available in the I/B/E/S. If the drop in coverage is based on the delisting of the respective stock, we drop the observation from our sample.

	All Firms		**Low Coverage**	
	$N = 1,271$		$N = 389$ ($Q_{0.25}$)	
Δ**Coverage** (2018 $\rightarrow$ 2019)	Total	%	Total	%
$\geq +3$	72	5.66	12	3.08
$+2$	86	6.77	26	6.68
$+1$	155	12.20	74	19.02
0	311	24.47	163	41.90
-1	275	21.63	77	19.79
-2	139	10.93	13	3.34
≤ -3	192	15.11	1	0.26
No more Coverage	41	3.23	23	5.91
Total	1,271	100.00	389	100.00

Chapter 4

Innovative Finanzierungen über Initial Coin Offerings: Struktur und bisherige Performance

Chapter 4 has been published as a journal article:
Häfner, David and Dirk Schiereck (2020). "Innovative Finanzierungen über Initial Coin Offerings: Struktur und bisherige Performance", in: *Zeitschrift für KMU und Entrepreneurship (ZfKE)*, 68.2, pp. 73-98, ISSN 1860-4633.
DOI: https://doi.org/10.3790/zfke.68.2.73

Chapter 5

Statistical Properties of Cryptocurrency Order Books

5.1 Introduction

Securities exchanges or multilateral trading platforms allow us to analyze the behavior of interacting market participants. In these financial markets all market participants trade the same asset and pursue the same goal (profit) which should incentivize each market participant to act rational. This setting allows to study not only general economic and financial theories, but also theories of individual human behavior. By analyzing market microstructure data, apparently *universal* laws have been discovered in the literature, that seem to hold true, independent of the traded asset or the exchange (e.g. Bouchaud, Mézard, and Potters, 2002 and Næs and Skjeltorp, 2006). These statistical laws help to better describe the competitive behavior of market participants and some of these laws have revealed noticeable similarities across different assets, time periods and regions (Potters and Bouchaud, 2003, Cont, 2001 and Mantegna and Stanley, 1999).

In this study we investigate, whether properties in established securities markets can also be found in cryptocurrency markets. Cryptocurrencies did not yet exist when some of the statistical laws were first discovered and verifying these properties in cryptocurrency markets would strongly boost their validity and the robustness of those characteristics adding important evidence to the existing microstructure literature.

Groundbreaking work has been done by O'Hara (1997), who analzyes the development of microstructure theory and the evolution of the literature to that point and more recently the impact of high-frequency trading on market microstructure (O'Hara, 2015).

Studying market microstructure adds to a deeper understanding of financial markets in general and the real economy benefits from efficient markets in many ways, e.g. risk sharing and an efficient capital allocation. As market microstructure and the market design directly affect market efficiency, inefficient markets on the micro level lead to higher transaction costs and imply that someone can earn money on

someone elses expense. If a market is inefficient on a microscopic level, direct and indirect transaction costs will inevitably increase. Further, an ideal price discovery may be disturbed, which can lead to gaps between the price and the perceived value of an asset. This again affects the real economy, as an investor is only willing to hold an asset if he is confident about its value. This is crucial for companies to be able to use financial markets as a reliable source of capital. These relations between the market microstructure and the real economy are especially relevant for emerging cryptocurrency markets. Recent developments in the acceptance of cryptocurrencies as a source of capital for companies – e.g. via Initial Coin Offerings (ICOs) – or for portfolio diversification makes analyzing cryptocurrencies on a micro level more important than ever before.

The remainder of this study is structured as follows: In Section 5.2 and Section 5.3, we first describe the operating principle of a Limit Order Book (LOB) and give a brief overview of the cryptocurrencies examined in this study. Further, a detailed description of the data collection process is provided and the data set used for the subsequent empirical analysis is described and descriptive statistics are presented. In Section 5.4, we reconstruct the limit order books of three different cryptocurrencies from our data. We compute the aggregated limit order book volume, also referred to as the slope of the order book. As shown in our descriptive analysis, we find empirical evidence for volume peaks in the LOB at specific price levels relative to the best price. We hypothesize that these peaks do not appear at random but are rather caused by "lazy" investors disregarding the granularity of the price grid. We test the empirical significance of our observation and find these peaks to be statistically highly significant and to appear across all examined LOBs.

We further find that the slope of the LOB varies over time, i.e. the aggregated LOB changes its shape. This finding raises the question, whether information is hidden in the slope of the order book. We examine the explanatory power of the slope in Section 5.5. Similar to Næs and Skjeltorp (2006) who study stock LOBs, we investigate three groups of models, using the slope of the order book to explain changes in prices, trading volume, and the correlation between price changes and trading volume. Our results suggest that the slope of the LOB can explain changes in the dependent variables. We further find evidence, suggesting that limit orders placed far away from the best price still carry price-relevant information. Section 5.6 concludes and discusses implications of our findings.

5.2 Structure of a Limit Order Book

An asset can only be traded, when a buyer and a seller find to each other. While the buyer wants to pay as little as possible for the respective asset, the sellers wants sell for the highest possible price. Exchanges serve as a platform to enable trading

by helping potential buyers and sellers to find a trading partner. Since the development of stock exchanges, two alternative systems have emerged differing in the way of how trading is organized: In quote-driven markets, market makers will post buy and sell quotes for which they are willing to trade an asset.[1] A market maker is responsible to provide the market with liquidity by matching incoming orders against other orders or by buying or selling from his own inventory. While market makers provide liquidity – which is beneficial to all market participants – quote-driven markets lack transparency, as market participants do not know the identity of their counterpart.[2] The second way to organize trading on an exchange is via a limit order book. Exchanges operating via a LOB are commonly referred to as order-driven markets. In an order-driven market, all valid buy and sell orders are listed in the (electronic) LOB, which each market participant has access to, providing full transparency to all market participants. Order driven markets became increasingly popular in the past and many of the largest stock exchanges currently operate with LOBs (e.g. NYSE, Euronext, Deutsche Börse, Nasdaq, the Tokyo Stock Exchange and the London Stock Exchange). The matchmaking of a LOB is based on a set of predefined mechanical rules and strictly follows a first-come, first serve principle and meets a price-time priority. Further, the order processing mechanics of LOBs is automated and follows transparent rules. On the contrary, the flow of incoming orders is solely based on traders internal decision making processes.

From a scientific point of view, the arrival of limit orders in an order-driven market is a complex field of study as the order placement is not bound to any rules and purely stems from the decision of an individual to indicate the willingness to buy or sell a specific quantity of the respective asset during a specific point in time for a self-set price. Thus, studying limit order books yields the most detailed insight into dynamic behavior in financial markets. As every decision by a market participant is directly linked to a potential financial loss or gain, market participants are strongly incentivized to act rational, providing an ideal experimental setting to study economic behavior.

Figure 5.1 depicts a stylized version of a LOB. Each block on the demand side resembles one unit of the traded asset. For simplicity, we assume that each limit order refers to exactly one unit of the traded asset. We define the best bid $b(t)$ as the highest price level, at which at least one market participant is willing to buy one unit of the asset. $a(t)$ describes the lowest price for which at least one market participant is willing to sell. In this steady state of the order book, no trade would occur, as

[1] While cryptocurrency markets are not order-driven and there are no market makers in Bitcoin markets (Marshall, Nguyen, and Visaltanachoti, 2019), Holste and Gallus (2019) find empirical evidence for "market maker quotes" at cryptocurrency exchanges. "Market maker" type of traders issue limit orders with attractive prices, however the volume of their offers is rather small and they should therefore not be regarded as liquidity providers.

[2] Anonymity is no distinct feature of quote-driven markets. When trading is organized by operating a limit order book, the identity of the counterpart if often unknown to traders and only referred to by a unique number. Solely the exchange operator knows about the true identity of the buyer and seller.

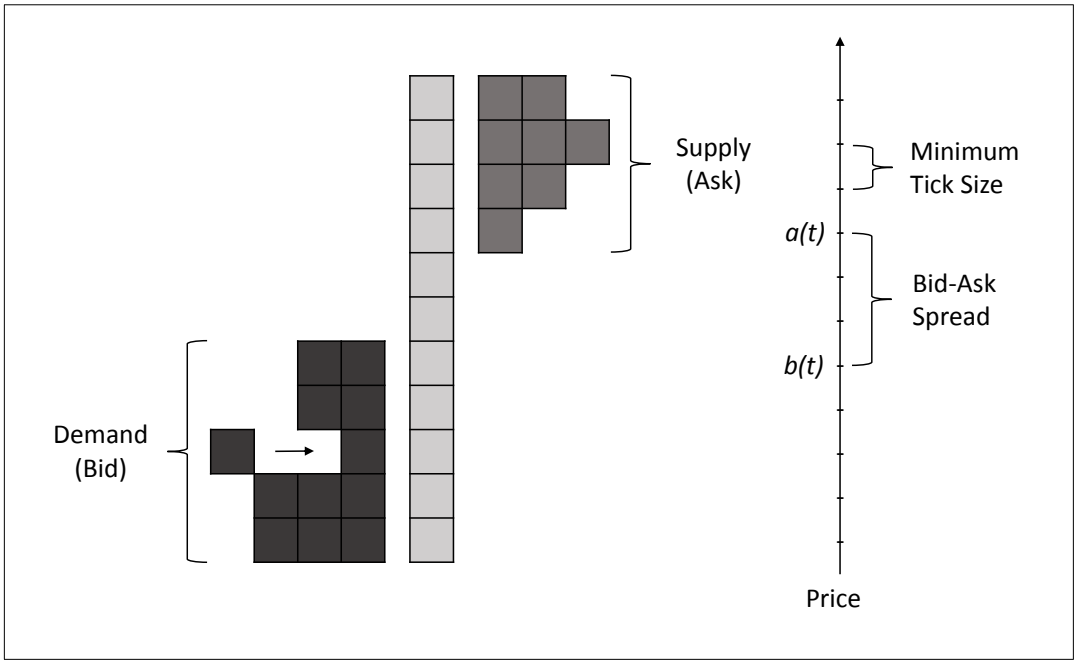


FIGURE 5.1: Structure of a Limit Order Book. Own Representation based on Preis et al. (2006)

the lowest sell price is higher than the highest ask price. Consequently, $a(t) - b(t)$ describes the implicit bid-ask spread in this model.

Observing the state of the LOB shown in Figure 5.1, imagine the case that a new market buy order arrives comprised of two units at a price level of $a(t) + 1$. This market order would be matched against the best limit sell order available in the order book. In the current state of the order book, there exists exactly one unit at a price level of $a(t)$ and two units at a price level of $a(t) + 1$. Thus, the unit at price level $a(t)$ will be matched against the incoming order. The next best price is at $a(t) + 1$. However, as there are two units available, it is not clear which one should be matched against the remaining unit from the market order. In this case a time priority is met, i.e. the limit sell order at price $a(t) + 1$ which has been submitted first will be matched against the market order. *Ceteris paribus*, this would lead to an increase in the bid-ask spread by one tick as there is no volume left at $a(t)$ after the market order has been matched. One would describe the arrival of a market sell order in an analogous manner.

If a limit order arrives at a price level, which can not directly be matched, it will be added to the LOB. This scenario is shown for a limit buy order with a size of one unit at a price level of $b(t) - 2$ in Figure 5.1. As there exist higher bids in the order book, this order will not be executed immediately, but rather added to the order book (indicated by the arrow) and remain in the order book until either all higher bid orders are already matched or canceled and a sell order at price level $b(t) - 2$ arrives, or the order is canceled by the trader who submitted the order.

It is important to note that besides the arrival of the market buy order or the arrival

of the limit buy order, all subsequent processes are based on predefined mechanical rules. The order book follows two simple principles to decide how an order is matched: price priority and time priority, with the first dominating the latter. Due to the seemingly nondeterministic arrival of orders, the volume at each price level in the order book is nondeterministic as well and provides an interesting research field on its own. Recent studies also discuss the endogeneity of the size of the bid-ask spread. As shown above, the bid-ask spread changes based on the arrival of new orders and is thus based on the behavior of the market participants. However, when the size of the bid-ask spread almost always equals the minimum tick size – e.g. when the traded asset is very liquid – the question arises, whether the bid-ask spread is in fact perceived as exogenous rather than endogenous by market participants (Biais, Hillion, and Spatt, 1995). In this case, the granularity of the price grid – which is set by the minimum tick size – may prevent the observability of the endogeneity of the size of the bid-ask spread.
While determining the price level of a limit order, market participants have to take two antagonizing criteria into account: Opting to place a limit order close to or at the bid-ask spread leads first to, a higher probability that the order will be fulfilled; and secondly, a decrease in the mean time passing until the order will be matched. On the other hand, placing a limit order close to the bid ask-spread yields the risk of having the limit order executed at an unfavorable price, abandoning the chance of taking advantage of price movements towards the desired direction (e.g. the possibility to sell at a higher price or buy at a lower price). Thus, every market participant has to select a price level for his limit order which takes into account his individual price and time preferences.

5.3 Data Collection and Description

In order to analyze order book characteristics in cryptocurrency markets, we collect trading data from one of the largest US-based cryptocurrency exchanges. The exchange generates a turnover of 1 bn. USD and offers one of the largest online trading platforms for cryptocurrencies to a peak of more than 10 mn. worldwide users. At the time of data collection, the cryptocrurrencies Bitcoin (BTC), Bitcoin Cash (BCH), Ethereum (ETH) and Litecoin (LTC) could be traded against each other and against US Dollar (USD) and Euro (EUR).
Bitcoin is a peer-to-peer electronic cash system first proposed by Nakamoto (2008). The main advantage of Bitcoin is the redundancy of intermediaries as transactions can be made peer-to-peer and are tamper-proof as all transactions are recorded and stored in a blockchain. Bitcoins can either be created via a process known as "mining" or purchased on an exchange. While a traditional money transfer typically involves a third party (e.g. a bank), no such entity is necessary to transfer Bitcoin.

In order to send Bitcoin from one party to another, only a bitcoin wallet and an internet connection is necessary. In the Bitcoin network, many "miners" define the exact structure of Bitcoin using an algorithm, thus the currency is not controlled by one single authority. This organizational structure leads to continuous majority decisions where computational power is used to vote. Bitcoin is transparent as each transaction is stored and traceable in the blockchain. A new transaction is verified only if the majority of miners in the system confirms the transaction. Although the transaction history is visible, the counterparties of a transaction remain anonymous, as solely the address of the Bitcoin wallet of the sending and receiving entity can be observed. With approximately ten minutes, the average transaction time in the Bitcoin network is many times faster than a traditional bank transaction which still takes some business days to be completed. However, Bitcoin is currently not able to compete with the speed of large credit card operators.
As of September 2019, Bitcoin has a total market capitalization of approximately 189 billion USD which is about 70% of the total cryptocurrency market. The market capitalization constantly changes by either a change in the BTC/USD price or an increase in the available amount. The maximum amount of Bitcoin is mathematically limited to 21 mn. BTC of which approximately 18 mn. BTC have been mined as of 2019. Computing power needed to mine the remaining amount increases dynamically with the total computing power in the network. Bitcoin has many possible applications as it can be used as means of payment, a protection against inflation or as a value storage. With a maximum of 21 mn. BTC, its limited availability makes BTC a scarce resource. Further, Bitcoin is decentralized and available through the internet making it portable and tradeable in small units with low storage costs – one major advantage over storing values physically, e.g. by buying gold.
A big problem of the Bitcoin network is the transaction speed and energy consumption. Currently, Bitcoin can only verify seven transactions per second. With an increase in popularity, more transactions need to be verified per second creating a potential bottleneck for the mainstream adoption and usability in everyday transactions. Moreover, one single Bitcoin transaction consumed at least 300 kWh in 2018 (De Vries, 2018, p. 804), while a bank transfer by credit card only needs 0.001 to 0.002 kWh. In 2017, the cryptocurrency Bitcoin Cash (BCH) was introduced to tackle this scaling challenge. Bitcoin Cash was created by a hard fork of the Bitcoin blockchain making Bitcoin Cash technically almost identical to Bitcoin except for an increased block size of the blockchain, resulting in a higher number of transactions that can be verified per second. While Bitcoin Cash is faster than Bitcoin, larger blocks are harder to process, favoring larger miners which is diametral to the initial Bitcoin concept of decentralization. Since its creation, Bitcoin Cash developed towards an independent cryptocurrency and is the fourth largest cryptocurrency in terms of market capitalization (5.3 bn. USD) as of September 2019.

Similar to Bitcoin and Bitcoin Cash, Ethereum is based on the blockchain technology. However, Ethereum acts as both a cryptocurrency and a decentralized computing platform that allows developers to execute decentralized computer programs. Ethereum was introduced in 2015 by Vitalik Buterin and is used as a means of payment in the Ethereum network. As of September 2019, Ethereum is the second largest cryptocurrency after Bitcoin with a market capitalization of approximately 19 bn. USD. From a consumer perspective, Ethereum can be used to do everything that is possible with fiat money. Ethereum can be spent, invested, saved or it can be transferred to peers. For companies, Ethereum can be used to finance M&A activities and other investment decisions without a financial intermediary. Neither a bank nor a payment processor is needed to transfer Ethereum, i.e. bank fees for granting a financing loan can be avoided completely by using Ethereum. On a broader scale, Ethereum has the potential to make the economy more efficient and increase productivity as the present value of projects increases due to lower capital costs, leading to more profitable projects overall.

In the Ethereum network, the currency ETH is used to pay participating computers for providing computational power, i.e. Ether can also be mined. While the financial industry is assumed to be the primary user of the blockchain concept (Nofer et al., 2017), the Ethereum blockchain also receives increasing attention from more distant industry sectors lately. E.g. a USD 30 mn. real estate property was tokenized with blockchain in Manhattan in 2018 (Wolfson, 2018). The transaction was based on a theoretical "Two Token Waterfall" model proposed by Lippiatt and Oved (2018) utilizing the Ethereum blockchain, demonstrating the wide range of possible applications of Ethereum. The model provides a structural framework to tokenize real assets and is based on two tokens representing debt and equity classes. Consequently, both classes combined represent the total capitalization of a transaction. Lippiatt and Oved (2018) state, that this tokenized structure can increase liquidity of real assets. The waterfall depicted in Figure 5.2 represents the flow of cash in the case of liquidation of the tokenized asset. From the flow of payments it appears that debt token holders enjoy seniority. Compared to traditional debt, interest payments are not paid on a recurring basis in this model but the accrued amount is paid at the time of liquidation. This benefits the equity token holder as fewer cash requirements are necessary and the date of sale does not have to be predetermined, allowing equity holders to sell during favorable market conditions. In exchange for this flexibility, equity token holders must share their excess sales profit with debt token holders, on a prenegotiated split. Lippiatt and Oved (2018) utilize the distributed ledger technology as a decentralized clearing house that stores all financial transactions and where tokens represent the ownership of real assets as smart contracts.[3] The authors further show that the return profiles of the debt and equity token imply an underlying present value of the asset which would create arbitrage opportunities for traders

[3]In their paper, Lippiatt and Oved (2018) select Ethereum smart contracts.

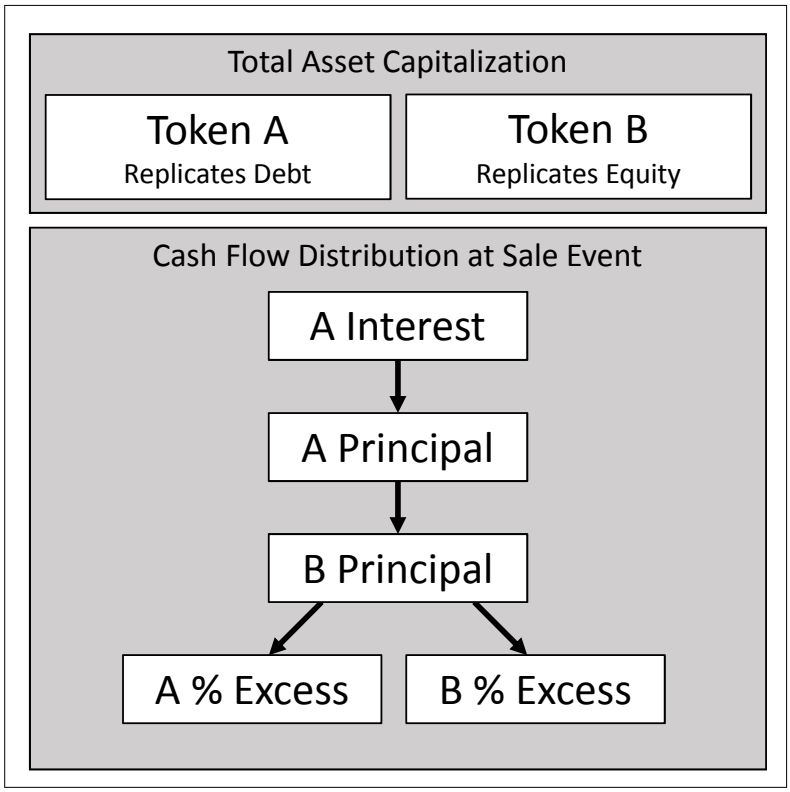


FIGURE 5.2: Waterfall Model for Tokenized Real Assets. Own representation based on Lippiatt and Oved (2018)

if the tokens are not priced correctly. Lippiatt and Oved (2018) claim that the tokenization of real assets can reduce illiquidity. The most obvious increase in liquidity is due to the simplification of trading and pricing of a real asset which creates a secondary market for those alternative investments. In accordance with Glosten and Harris (1988), the authors decompose the bid-ask spread and show that the asymmetric information component is reduced as a consequence of the transparency of smart contracts. Smart contracts allow to see the supply and holdings of all market participants leading to more educated investment decisions and thus reducing the likelihood of asymmetric information (Lippiatt and Oved, 2018). Further, clearing costs equal almost zero through the peer-to-peer transfer on the Ethereum network. Moreover, a potential investor is not tied to either provide equity or debt but can create his individual risk/return profile by blending debt and equity tokens of the same asset making an investment attractive for a broader range of investors. While tokenized real estate transactions are still often conducted in fiat currency in the end – mainly due to a lack of investor acceptance – the real estate transaction discussed above demonstrates how the transaction process for alternative assets is changing and how Ethereum differs from Bitcoin and Bitcoin Cash in it's potential use.
Due to the volatile nature of the cryptocurrency market, it is hard to derive a meaningful statement about the relevance of the above mentioned cryptocurrencies. However, we find that the market share of Bitcoin, Ethereum and Bitcoin Cash combined remains somewhat stable over time and accounts for roughly 80% of the total cryptocurrency market capitalization as of September 2019. Bitcoin alone has a total market share of approximately 71% (Ethereum c. 7%, Bitcoin Cash c. 2%). Including these three major currencies in our analysis, we are convinced to capture a representative picture of the cryptocurrency market.[4]

5.3.1 Data Acquisition Process

We use real-time market data updates for orders and trades provided by a websocket feed, which can be used to reconstruct a real-time order book. While the websocket feed is publicly available, connections to it are rate-limited. By creating a script, we are able to store all updates received by the websocket feed locally. Observing the built-in Sequence Number we can assure that we do not miss any updates and in fact can recreate the full order book at any point in time during our period of observation. Table 5.1 provides an example excerpt of the data that we are able to record. Updates to the order book happen, when the status of an order changes. "Type" defines the event that occurs and is split into the four distinct categories: "received", "open", "match", or "done". Whenever a new order arrives, it is first "received" and then "open" in the LOB. The order, which is identified by its unique "Order ID", will remain in the order book until either the entity who created the order cancels it, or

[4]Based on a market capitalization of BTC: 188.9 bn. USD, ETH: 19.2 bn. USD, BCH: 5.3 bn. USD, as of 04 September 2019.

TABLE 5.1: Example of events occuring in a Cryptocurrency Order Book

Notes: This table shows an excerpt of the development of the Bitcoin/Euro order book on 18th April 2018. Note that all events shown in this table occur in just 0.043 seconds. If a new limit order is "received", it will remain "open" in the order book until it either gets "canceled" by the trader or "matched" against another order. Figures have been slightly simplified for readability, e.g. the true "Order ID" is composed of 32 digits to guarantee uniqueness. "Sequence" increases by 1, whenever a new event occurs and proves that no updates in the order book were missed. We further have information on the best bid and ask during each event.

Type	Order Type	Side	Price	Size	Time	Sequence	Order ID	Remaining Size	Reason
open		sell	6,628.77		21:59:18.320	0		0.001	
received	limit	buy	6,528.61	0.001	21:59:18.323	1			
open		buy	6,528.61		21:59:18.323	2		0.001	
done		sell	6,636.27		21:59:18.336	3		0.001	canceled
done		sell	6,645.27		21:59:18.356	4		0.001	canceled
done		sell	6,637.77		21:59:18.360	5		0.001	canceled
received	limit	buy	6,507.61	0.001	21:59:18.362	6			
open		buy	6,507.61		21:59:18.362	7		0.001	
received	limit	buy	6,512.11	0.001	21:59:18.363	8			
...									

the order is matched against another order. The "Order Type" defines, whether an order has been issued as a market order ("market") or a limit order ("limit"). "Side" indicates from which side an order has been issued, i.e. if someone wants to "sell" (ask side) or "buy" (bid side) some amount of the respective cryptocurrency. "Size" defines the volume of the order.

We track the evolution of the order book for every currency pair combination which can be traded at the time of observation, allowing us to obtain an immense amount of trading data (see Table 5.2). We are able to store and analyze more than 60 gigabyte worth of trading data for eleven different currency pairs.

Further, a Rest API is available which we use to compute periodic order book snapshots. While the Rest API is more bulky and slow, it allows us to gather preprocessed data, which would be unfeasible to reconstruct from the websocket feed in a timely manner. We use the Rest API to capture a snapshot of the order book every ten minutes. We use this data to examine statistical properties of the average shape of the order book.

5.3.2 Descriptive Statistics

Table 5.2 gives an sample overview and some meta data about the gathered trading data. Our observation period spans from April 2018 to August 2019 (including gaps). Due to hardware and software restrictions (e.g. forced reboots, operating system updates, server connection losses and data storage limitations), we are not able to gather 100% of the daily order flow and therefore split the observation into multiple "sessions" spread across the day.

Table 5.3 provides a breakdown of order types and cancellation rates. We find that across all currency pairs and order types, the cancellation rate is remarkably high with the BTC/USD sell side (98.36%) being the lowest overall cancellation rate. Moreover, we find that market orders are rare, i.e. orders are almost exclusively issued as limit orders.

High cancellation rates as depicted in Table 5.3 indicate the existence of high-frequency trading (HFT). While a strict definition of HFT does not exist, we find that the recent definition of "high-frequency algorithmic trading" as a subset of algorithmic trading issued by the European Securities and Markets Authority (ESMA) under the rules of MiFID II is a useful definition. According to ESMA (European Commission, 2014, p.36), HFT is mainly characterized by:

- an infrastructure intended to minimize network latencies (e.g. co-location);
- the controlling of order initiation, generation, routing or execution by machines without human interaction;
- high "message intraday rates" concerning orders, quotes or cancellations.

ESMA further states that HFT is characterized by a high daily portfolio turnover, a high order-to-trade ratio and ending the trading day at or close to a flat position (European Commission, 2014, p.10).

HFT is discussed controversial as the net economic benefit or loss remains unclear. The real economy benefits from HFT in many ways, namely a higher liquidity, an increased trading and order book volume, a reduction of the bid-ask spread and a better price formation and execution coming along with an overall improved price quality which ultimately decreases the cost of capital. However, enhanced HFT bears many risks, e.g. an overload of the trading systems due to high order cancellation rates, an increased price volatility and the potential for market manipulation. In addition, slower traders could stop trading, suspecting that high-frequency traders use their informational advantage against slower traders to earn a profit off of them.

In Germany, HFT is regulated since 2013 by the high-frequency trading bill (Hochfrequenzhandelsgesetz) which includes some major amendments in order to prevent dangers and the misuse of HFT. Notably, in addition to improved system control, risk control and transparency guidelines, an order-to-trade ratio has been stipulated, which aims to limit the amount of updates a trading system is allowed to send towards an exchange. The definition and measurement of the order-to-trade ratio has to be provided in the stock exchange regulations (§26a BörsG).

HFT regulation has also been discussed on an European level, and responsibilities of firms engaging in HFT have also been defined in MiFID II to ensure market quality, notably to store records of their trading systems for a minimum of five years and the implementation of measures to prevent market distortion.

Unfortunately, we can not infer the share of market participants engaging in HFT

and/or algorithmic trading from our data, as this number is even tough to measure at regulated stock exchanges.[5]
Similar to Biais, Hillion, and Spatt (1995), we compute the conditional probabilities of LOB events. Our results are presented in Table 5.4. The variation of conditional probabilities in each column of Table 5.4 indicates that order book events are not statistically independent from previous order book events.

5.4 Limit Order Book Characteristics in Cryptocurrency Markets

5.4.1 The Shape of the Average Order Book

Incoming limit orders are stored in the LOB until they are either executed or canceled. The sum of the demanded or supplied quantity of the traded assets currently available in the LOB at a given price level defines the order book volume at that price level and represents the current queue size. The sum over all price levels is now referred to as the depth of the order book. As incoming orders are determined by market participants, so is the volume of the order book at a given price level.
In Figure 5.2, the sum of the squares at each price level represents the volume of the order book at that price level.
Our analysis of LOBs provides insights into the microstructure of cryptocurrency markets and serves two major purposes: First, studying market behavior at cryptocurrency exchanges is crucial in order to understand, whether cryptocurrencies can extend the bandwidth of options for corporate finance by providing an alternative way of raising capital. Second, identifying market behavior at cryptocurrency exchanges analogous to those of security exchanges would indicate that major market characteristics are determined by intrinsic (economic) human behavior and are less determined by asset specific characteristics. Thus, studying cryptocurrency LOBs is important for the fields of market microstructure but also yields valuable empirical insights for corporate and behavioral finance theory. In this section, we study, whether the behavior of market participants creates similar patterns in cryptocurrency limit order books as have been found in stock markets.
Examinig stock markets, Potters and Bouchaud (2003) argue that the shape of the average order book is not clear *a priori*. While most of the incoming orders arrive in proximity to the current bid or ask price, an order placed close to the current price has a larger probability to be executed and disappearing from the order book. Bouchaud, Mézard, and Potters (2002) find the time-averaged shape of the limit order book to be characterized by a maximum, distant to the current bid-ask spread.

[5]Nonbinding estimations for the amount of HFT as a share of total trading activity mostly lie close to the 50% range for stock exchanges. Ultimately, one may not abandon the fact, that all algorithmic trading strategies act according to rules made by humans pursuing the goal to generate profit, thus they do not act independent and are just an extension of the human capabilities.

TABLE 5.2: Sample Overview and Meta Statistics of Trading Data.

Notes: We recorded data from 18th of April 2018 until 31th of August 2019 of live order book updates. During this time frame, we aggregate 60 GB of raw data, tracking every event in the order book of the respective currency pair. The deviation from "Recorded (Days)" and "Full Data (Days)" stems from exogenous events, e.g. forced operating system updates or a reboot of the server used for data storage. We exclude days, where we are missing data due to those events. Trading is allowed nonstop. One "Day" refers to a full 24h cycle rather than a traditional trading day which is dependent on the opening hours of the respective exchange. BTC refers to Bitcoin. BCH refers to Bitcoin Cash. ETH refers to Ethereum. USD refers to US Dollar. There are no records of currency pairs including BCH in April 2018 as we were not able to receive live order book updates in this time frame. We observe some interrupts tracking the order flow since mid-2019. Upon closer inspection, these interrupts do not follow a pattern and appear to be unsystematic. We tackle this data issue by controlling for monthly fixed effects in our empirical analysis, thus this occurrence does not raise any concerns regarding the unbiasedness of our results.

Currency Pair	**Recorded (Days)**	**Full Data (Days)**	**Avg. Filesize (MB/Day)**	**Avg. No. of Sessions (per Day)**	**Avg. Length of one Session (min)**	**% of OF Recorded (per Day)**
April 2018						
BTC/USD	12	11	262	24	19	31
ETH/USD	12	11	239	27	18	33
BCH/USD	–	–	–	–	–	–
May 2018						
BTC/USD	31	30	252	23	37	57
ETH/USD	31	30	251	28	14	27
BCH/USD	11	10	233	34	3	8
June 2018						
BTC/USD	22	20	233	24	26	40
ETH/USD	24	23	244	25	19	31
BCH/USD	23	22	237	23	45	41
July 2018						
BTC/USD	–	–	–	–	–	–
ETH/USD	6	5	105	20	7	9
BCH/USD	13	13	80	25	2	3
February 2019						
BTC/USD	17	15	261	25	19	32
ETH/USD	16	13	242	17	57	66
BCH/USD	17	15	207	4	310	83
June 2019						
BTC/USD	7	7	72	13	4	3
ETH/USD	7	6	71	13	10	9
BCH/USD	8	8	71	13	8	7
July 2019						
BTC/USD	11	11	70	13	6	5
ETH/USD	11	11	69	12	12	10
BCH/USD	11	11	68	12	20	17
August 2019						
BTC/USD	17	17	69	12	8	7
ETH/USD	17	17	68	10	19	14
BCH/USD	17	17	68	9	36	25

TABLE 5.3: Breakdown of Order Types and Cancellation Rates

Notes: This table provides conditional probabilities for specific events in the order book. If a new order arrives in the order book, it is labeled as "received" and be either a market or a limit order. Each order remains in the order book until it is "done", which can be either due to a cancellation or because the order was matched ("filled"). The interpretation is shown on the following example: The value 99.58% in the first row and column (BTC/USD, Panel A: Sell Side, Received, Limit) is the probability of an incoming sell order in the BTC/USD order book to be a limit order, i.e. $P(Order\ Type = limit | Currency\ Pair = BTC/USD, Type = received, Side = sell) = 99.58\%$. Data Source: Cryptocurrency limit order books obtained from April to August 2019 (see Table 5.2 for a data overview).

	BTC/USD	**ETH/USD**	**BCH/USD**	**Mean**
Panel A: Sell Side				
Received				
Limit	99.58%	99.76%	99.91%	99.75%
Market	0.42%	0.24%	0.09%	0.25%
Done				
Canceled	98.36%	98.44%	98.81%	98.54%
Filled	1.64%	1.56%	1.19%	1.46%
Panel B: Buy Side				
Received				
Limit	99.67%	99.79%	99.68%	99.71%
Market	0.33%	0.21%	0.32%	0.29%
Done				
Canceled	99.22%	99.36%	98.77%	99.12%
Filled	0.78%	0.64%	1.23%	0.88%

TABLE 5.4: Conditional Probabilities of the Order Type in Time t given the Order Type in $t-1$

Note: The table shows the conditional probabilities of specific order types following each other in the order book. Each column represents the probability of the respective order type given the order type of the previous order equals the order type of the row. The table can be read as follows: The conditional probability that a Small Buy Order is followed by a Large Buy Order is: $Prob.(\text{Large Buy}_t|\text{Small Buy}_{t-1}) = 0.22\%$. The sum of each row equals 100%. Note that the sum of each column does not give the unconditional probability of the respective event. Similar to Biais, Hillion, and Spatt (1995), we define market orders as large if their volume exceeds the existing volume at the best bid or ask price and the price allows the order to climb through the opposite side of the order book. A small market order does not possess at least one of those characteristics.

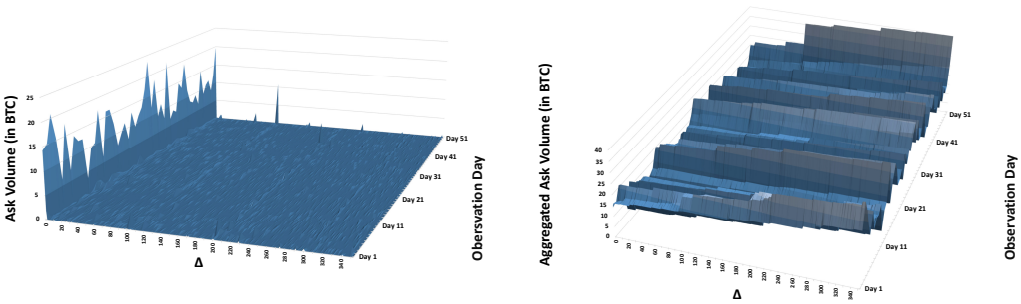
Time t / **Time** $t-1$	Large Buy	Small Buy	Buy Order inside Spread	Buy Order at Spread	Buy Order below Spread	Canceled Buy Order	Large Sell	Small Sell	Sell Order inside Spread	Sell Order at Spread	Sell Order above Spread	Canceled Sell Order
Large Buy	**1.13%**	2.58%	12.86%	2.13%	14.43%	35.45%	0.06%	0.10%	2.13%	0.45%	15.11%	13.58%
Small Buy	0.22%	**2.58%**	4.09%	2.24%	21.76%	33.24%	0.05%	0.29%	0.72%	1.69%	12.60%	20.52%
Buy Order inside Spread	0.05%	0.11%	**6.76%**	5.68%	17.45%	50.58%	0.01%	0.04%	2.35%	0.47%	7.20%	9.29%
Buy Order at Spread	0.06%	0.30%	5.31%	**6.36%**	17.63%	38.75%	0.01%	0.14%	0.94%	1.36%	12.63%	16.51%
Buy Order below Spread	0.03%	0.23%	1.43%	1.06%	**25.61%**	39.75%	0.03%	0.14%	0.89%	0.90%	15.14%	14.80%
Canceled Buy Order	0.03%	0.16%	3.84%	1.86%	38.50%	**27.53%**	0.02%	0.09%	0.95%	0.79%	9.81%	16.41%
Large Sell	0.07%	0.19%	2.59%	0.35%	18.67%	13.54%	**3.54%**	8.35%	12.91%	1.60%	8.67%	29.51%
Small Sell	0.04%	0.48%	0.70%	2.10%	18.28%	24.51%	0.73%	**5.53%**	8.27%	2.20%	11.94%	25.23%
Sell Order inside Spread	0.02%	0.09%	3.92%	0.79%	13.93%	16.29%	0.07%	0.14%	**5.95%**	4.09%	12.70%	42.00%
Sell Order at Spread	0.01%	0.25%	1.10%	1.60%	18.40%	23.37%	0.05%	0.26%	3.03%	**5.99%**	14.66%	31.30%
Sell Order above Spread	0.04%	0.24%	1.12%	1.18%	24.18%	22.04%	0.03%	0.16%	0.93%	1.17%	**24.87%**	24.03%
Canceled Sell Order	0.02%	0.15%	1.27%	1.04%	13.67%	25.68%	0.02%	0.11%	2.98%	1.81%	26.72%	**26.54%**

This finding is surprising as the probability of an incoming order to be placed is highest at the best price. The authors find this shape across different stocks listed at the Paris Bourse. In addition, they find that the shape of the time-average order book is roughly symmetric between the bid and the ask side.[6] Potters and Bouchaud (2003) analyze the average shape of the order book of two exchange traded funds that track the NASDAQ and the S&P500 performance respectively, confirming a maximum of the queue size lying away from the bid-ask spread for one of the ETFs. For the other ETF however, the queue has a maximum at the current bid-ask spread.[7]
In order to buy a cryptocurrency (e.g. Bitcoin, Bitcoin Cash or Ethereum) a buyer needs to find someone to trade the respective currency for another currency, e.g. US Dollar. Bringing together buyers and sellers of cryptocurrencies is the purpose of cryptocurrency exchanges, which operate a LOB following the same set of rules as stock exchanges, in particular the price-time priority and the first-come, first-served principle. In this section, we restrict our work and empirical analysis to the LOB of three major cryptocurrencies: Bitcoin (BTC), Bitcoin Cash (BCH) and Ethereum (ETH) all of which can be bought and sold for US Dollar at a cryptocurrency exchange. We exclusively consider US Dollar order books, to retain a common denominator, allowing to compare our empirical results across cryptocurrencies.

5.4.2 The Bitcoin Order Book

Below, we analyze characteristics of the Bitcoin/US Dollar limit order book. Existing literature mainly focuses on the empirical shape of the order book of stocks. While Biais, Hillion, and Spatt (1995) and Næs and Skjeltorp (2006) examine the slope of the aggregated order book and potential connections to volume and volatility, Bouchaud, Mézard, and Potters (2002) show that a snapshot of the order book can deviate substantially from its average shape.
The slope of the order book is derived by computing the gradient for the additional supplied or demanded volume for deeper levels in the order book. Economically, the slope of the order book is the elasticity $\partial q/\partial p$ describing how quantity (q) provided in the order book changes as a function of the price (p) (Næs and Skjeltorp, 2006, p. 415).
First, we plot the empirical order book volume and the aggregated volume of the BTC/USD LOB as a function of the absolute distance measured in ticks (Δ) towards the best price.[8] The results are shown on a daily-average basis in Figure 5.3 and Figure 5.4 for the ask and bid side of the LOB respectively. It can be seen that the available volume is highest directly at the bid-ask spread. Both sides of the order

[6]These findings have been empirically confirmed by Zovko and Farmer (2002) and Mike and Farmer (2008).
[7]The authors argue that the observed deviation may be due to the order book of this ETF not being the dominant player at NASDAQ as Island ECN is just one of many trading platforms of NASDAQ covering only 20% of the total trading volume of this specific ETF (Potters and Bouchaud, 2003, p.136).
[8]One tick corresponds to one US Dollar cent in the BTC/USD LOB.



Notes: The figure shows both the average volume (left) and the average aggregated volume (right) dependent on the distance Δ to the best ask price for the first 350 ticks on a daily basis across the sample period. We derive the data by computing the average volume at each tick per day based on snapshots of the order book taken every 10 minutes. Accordingly, each single data point in the above chart represents the average value of 144 observations. Most of the volume is available directly at the spread, however, a pattern of slightly higher volume manifests surrounding definite numbers (100, 200 and 300 ticks away from the best ask price) emerge.

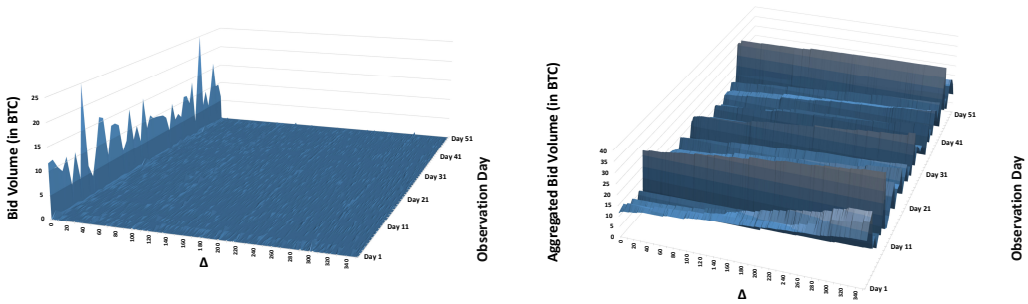
FIGURE 5.3: Ask Volume and Aggregated Ask Volume of the BTC/USD Limit Order Book relative to the Distance (Δ) towards the best Ask Price

book seem to behave almost symmetrical, however, the aggregated volume at the bid side increases more steadily in Δ. Looking at the shape of the aggregated order book in Figure 5.3 and Figure 5.4, we observe substantial variation of the shape across observation days, raising the question if there is information hidden in the slope of the order book.

For both, the bid and ask side, a higher average volume at multiples of 100 ticks away from the best price can be observed. This effect can be found consistently across the observation period and is shown in more detail in Figure 5.5. A high average volume across the whole sample period can be observed at round figures, especially for Δs that are a multiple of 100. Preferences for round figures have been documented in the financial literature before, notably Corwin (2003) finds that the underpricing in seasoned equity offers is significantly related to underwriter pricing conventions such as price rounding and pricing relative to the bid quote. Further links between numeric fluency and human preferences have been documented by Kettle and Häubl (2010) and we are confident that the observed pattern in LOB volumes can be linked to human preferences as well.[9]

Next, we examine the slope of the BTC/USD LOB. The slope of the order book represents the elasticity of the market supply and demand of the respective cryptocurrency. We test how the aggregated order book volume increases in Δ by comparing the empirical fit of three alternative sets of models. The first model assumes a linear relationship between aggregated order book volume and Δ, the second model

[9]We test the statistical significance of this finding in Section 5.4.5.



Notes: The figure shows both the average volume (left) and the average aggregated volume (right) dependent on the distance Δ to the best bid price for the first 350 ticks on a daily basis across the sample period. We derive the data by computing the average volume at each tick per day based on snapshots of the order book taken every 10 minutes. Accordingly, each single data point in the above chart represents the average value of 144 observations. Most of the volume is available directly at the spread, however, a pattern of slightly higher volume surrounding definite numbers (100, 200 and 300 ticks away from the best ask price) emerge.

FIGURE 5.4: Bid Volume and Aggregated Bid Volume of the BTC/USD Limit Order Book relative to the Distance (Δ) towards the best Bid Price

assumes a logarithmic relation, while the third model assumes a square root relation:[10]

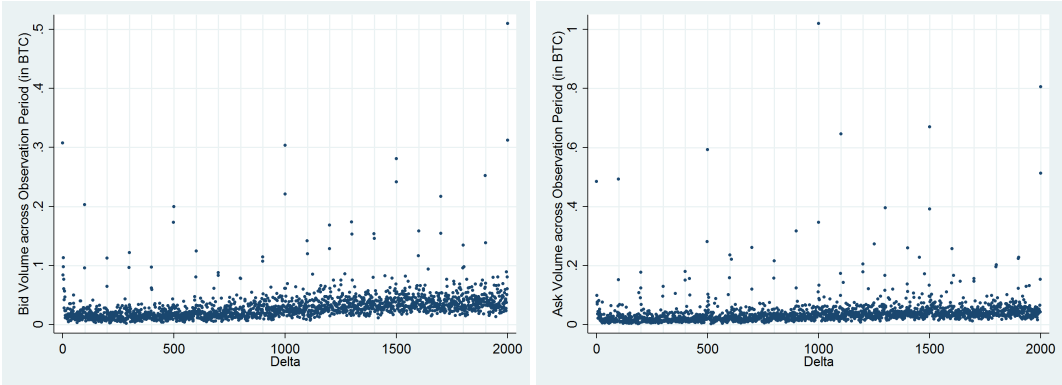
1. $Aggregated\ Average\ LOB\ Volume_{BTC/USD}(\Delta) = \beta_0 + \beta_1 \times \Delta$
2. $Aggregated\ Average\ LOB\ Volume_{BTC/USD,}(\Delta) = \beta_0 + \beta_1 \times ln(\Delta)$
3. $Aggregated\ Average\ LOB\ Volume_{BTC/USD}(\Delta) = \beta_0 + \beta_1 \times \sqrt{\Delta}$

The empirical fit is presented in Table 5.5. We find that the logarithmic model performs worst for both the bid and ask side of the BTC/USD LOB. Surprisingly – with an R^2 of 98.85% and 98.88% – the square root specification beats the linear model for both, the ask and bid side of the order book by 6.23 and 1.14 percentage points, respectively. Considering only the most relevant part of the aggregated order book ($\Delta \leq 100$), we find the square root model to still fit better than the linear model, however the difference in explanatory power between the two models diminishes. Considering only the first 100 ticks away from the current best price, the square root specification yields an R^2 which is 1.45 (ask side) and 0.86 (bid side) percentage points higher than in the linear model.

Our results suggest that the slope of the BTC/USD LOB decreases when the distance towards the best price increases. This finding is valid for both sides of the order book. Close to the best price however, a steady slope appears to be a reasonable approximation for the slope of the order book. Our results also indicate that assuming

[10]Based on our graphical analysis (see Figure 5.3 and Figure 5.4), we do not test for other functional forms. However, we test whether the functional form is concave or convex in Section 5.9. Note that the aggregated average LOB volume increases in Δ by definition.

a logarithmic relationship to describe the link between the aggregated average order book volume and Δ is not feasible.



Notes: The left (right) figure shows the aggregated bid (ask) volume relative to the best bid (ask) price across the observation period for the BTC/USD limit order book for the first 2,000 ticks. Volume directly at the spread is omitted due to scaling. Volume peaks occur at round numbers at both sides of the BTC/USD limit order book.

FIGURE 5.5: Average Ask and Bid Volume of the BTC/USD Limit Order Book relative to the Distance Delta (Δ) towards the Best Bid or Ask Price

TABLE 5.5: Slope of the Aggregated Order Book for BTC/USD

Notes: The table shows the slope of the average order Bitcoin/US Dollar LOB across our sample period. The slope is derived from a linear regression for different model specifications. Model specifications (1.1)-(3.2) consider the slope of the supply side. Respectively Model specifications (1.3)-(3.4) show the empirical results for the demand side. Each observation has a high level of confidence, as it denotes the average aggregated volume of the order book Δ ticks away from the best price. We measure the volume every ten minutes continuously across our sample period.

Ask Side						
Sample	ALL			$\Delta \leq 100$		
Model	(1.1)	(2.1)	(3.1)	(1.2)	(2.2)	(3.2)
Δ	0.02***			0.02***		
Std. Err.	0			0		
$\ln(\Delta)$		260.22***			0.62***	
Std. Err.		0.41			0.02	
$\sqrt{\Delta}$			5.18***			0.26***
Std. Err.			0			0
Constant	yes	yes	yes	yes	yes	yes
N	50,413	50,412	50,413	101	100	101
R^2	92.62%	88.92%	98.85%	95.98%	86.53%	97.43%
Bid Side						
Sample	ALL			$\Delta \leq 100$		
Model	(1.3)	(2.3)	(3.3)	(1.4)	(2.4)	(3.4)
Δ	0.03***			0.02***		
Std. Err.	0			0		
$\ln(\Delta)$		332.07***			0.54***	
Std. Err.		0.68			0.02	
$\sqrt{\Delta}$			6.86***			0.23***
Std. Err.			0			0
Constant	yes	yes	yes	yes	yes	yes
N	50,413	50,412	50,413	101	100	101
R^2	97.74%	82.63%	98.88%	97.00%	86.67%	97.86%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

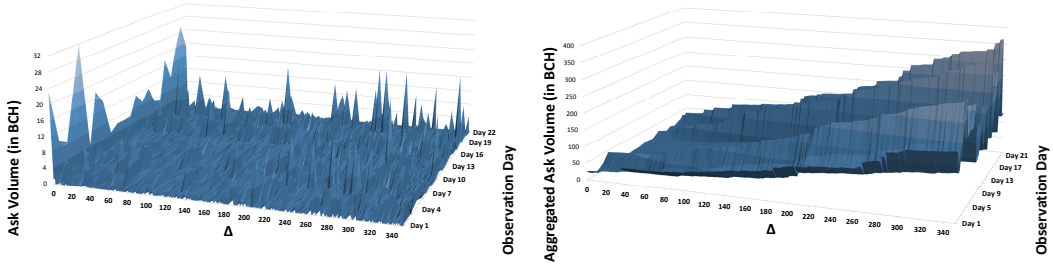
5.4.3 The Bitcoin Cash Order Book

Figure 5.6 and Figure 5.7 show the average ask and bid volume and the aggregated average ask and bid volume per observation day of the BCH/USD limit order book. We find that the shape of the order book is rather symmetrical between the bid and ask side. Compared to the BTC/USD aggregated order book however, more volume seems to be located away from the bid-ask spread resulting in a steeper slope of the order book. Similar to the previous section, we compute three different models to explain the slope of the order book as a function of Δ. We use the method of ordinary least squares (OLS) to compute the fit of three different models imposing a linear, logarithmic or square-root relation between the increase in Δ and the aggregated average order book volume:

1. $Aggregated\ Average\ LOB\ Volume_{BCH/USD}(\Delta) = \beta_0 + \beta_1 \times \Delta$
2. $Aggregated\ Average\ LOB\ Volume_{BCH/USD,}(\Delta) = \beta_0 + \beta_1 \times ln(\Delta)$
3. $Aggregated\ Average\ LOB\ Volume_{BCH/USD}(\Delta) = \beta_0 + \beta_1 \times \sqrt{\Delta}$

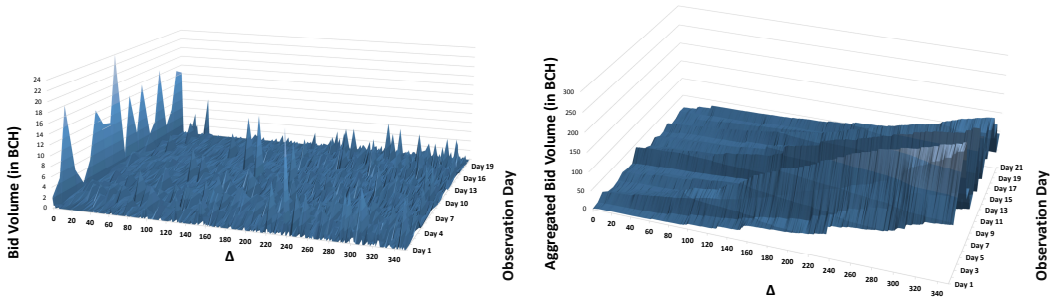
Our results are presented in Table 5.6. In line with our previous findings for the BTC/USD order book, we find the logarithmic model to possess the worst explanatory power. Comparing the linear to the square-root model, we find that the implied relation depends on the number of ticks taken into account. While both models do not differ much in their explanatory power, the square-root relation appears to better explain the increase in order book volume for both the bid and ask side if all levels of Δ are taken into account. Considering only the first 100 ticks closest to the bid-ask spread, the linear model outperforms the square-root model by three (ask side) and five (bid side) percentage points respectively.

Motivated by the volume pattern which emerged in the BTC/USD order book, we compute the average order book volume across observation days dependent on Δ (Figure 5.8). Again, we find large peaks at round figures for Δ, especially at multiples of 100 at both sides of the BCH/USD order book. Surprisingly, even larger peaks emerge at 500, 1,000 and 1,500 ticks away from the best price. It is important to note that these peaks represent average values and are unlikely to be outliers as they can be observed consistently across our observation period. We further find that the average shape of the order book appears to have a maximum away from the bid-ask spread, which also has been observed in stock markets by Bouchaud, Mézard, and Potters (2002). In the BCH/USD LOB, the average volume generally increases in Δ between 0 and approximately 250 ticks before it slowly declines from 250 ticks onward. Bouchaud, Mézard, and Potters (2002) propose an analytical approximation to compute the average order book, concluding that "[T]he shape of the average order book therefore reflects the competition between a power-law flow of limit orders with a finite lifetime, and the price dynamics that removes the orders



Notes: The figure shows both the average volume (left) and the average aggregated volume (right) dependent on the distance Δ to the best ask price for the first 350 ticks on a daily basis across the sample period. We derive the data by computing the average volume at each tick per day based on snapshots of the order book taken every 10 minutes. Accordingly, each single data point in the above chart represents the average value of 144 observations. Most of the volume is available directly at the spread, however, a pattern of slightly higher volume manifests surrounding definite numbers (100, 200 and 300 ticks away from the best ask price) emerge.

FIGURE 5.6: Ask Volume and Aggregated Ask Volume of the BCH/USD Limit Order Book relative to the Distance (Δ) towards the best Ask Price

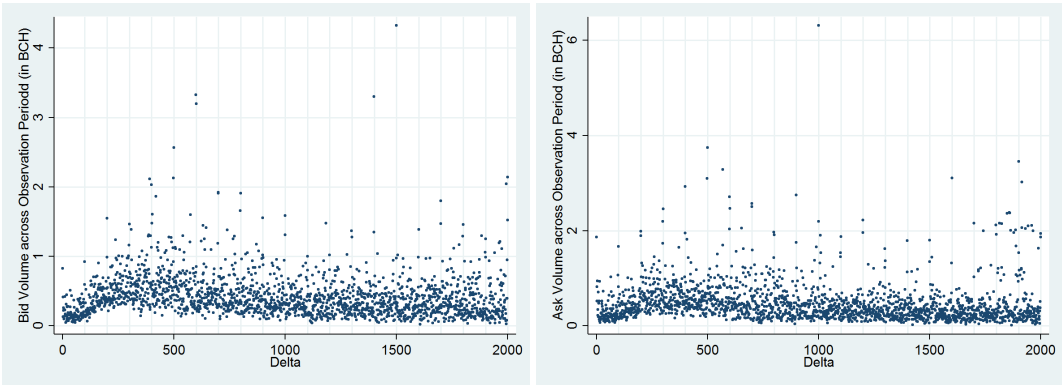


Notes: The figure shows both the average volume (left) and the average aggregated volume (right) dependent on the distance Δ to the best bid price for the first 350 ticks on a daily basis across the sample period. We derive the data by computing the average volume at each tick per day based on snapshots of the order book taken every 10 minutes. Accordingly, each single data point in the above chart represents the average value of 144 observations. Most of the volume is available directly at the spread, however, a pattern of slightly higher volume manifests surrounding definite numbers (100, 200 and 300 ticks away from the best ask price) emerge.

FIGURE 5.7: Bid Volume and Aggregated Bid Volume of the BCH/USD Limit Order Book relative to the Distance (Δ) towards the best Bid Price

close to the current price. These effects lead to a universal shape which will presumably hold for many different markets,[...]" (Bouchaud, Mézard, and Potters, 2002, p.11). The authors further indicate that preliminary results show that the same behavior can be observed in futures markets.

We provide empirical evidence that a hump away from the current mid point can also be observed in cryptocurrency markets. However, Bouchaud, Mézard, and Potters (2002) do not report volume peaks at round figures, making a significance test for our findings inevitable. To analyze this phenomenon, we perform statistical tests across the cryptocurrency order books for Bitcoin, Bitcoin Cash and Ethereum in Section 5.4.5.



Notes: The left (right) figure shows the aggregated bid (ask) volume relative to the best bid (ask) price across the observation period for the BCH/USD limit order book for the first 2,000 ticks. Volume directly at the spread is omitted due to scaling. Volume peaks occur at round numbers at both sides of the BCH/USD limit order book.

FIGURE 5.8: Average Ask and Bid Volume of the BCH/USD Limit Order Book relative to the Distance Delta (Δ) towards the Best Bid or Ask Price

TABLE 5.6: Slope of the Aggregated Order Book for BCH/USD

Notes: The table shows the slope of the average order Bitcoin Cash/US Dollar LOB across our sample period. The slope is derived from a linear regression for different model specifications. Model specifications (1.1)-(3.2) consider the slope of the supply side. Respectively Model specifications (1.3)-(3.4) show the empirical results for the demand side. Each observation has a high level of confidence, as it denotes the average aggregated volume of the order book Δ ticks away from the best price. We measure the volume every ten minutes continuously across our sample period.

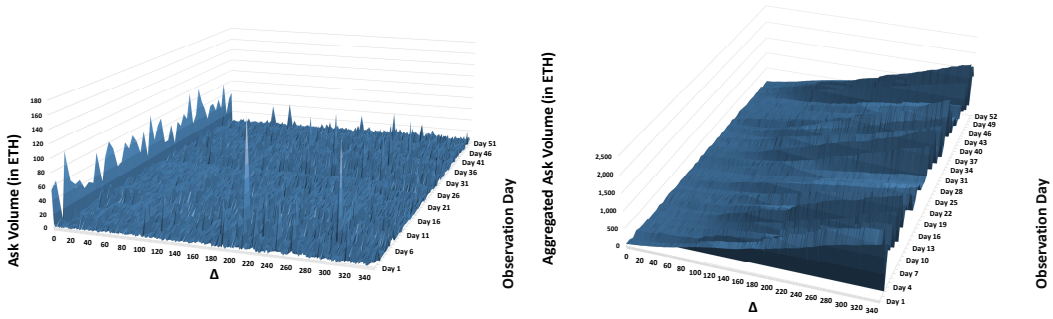
Ask Side						
Sample	ALL			$\Delta \leq 100$		
Model	(1.1)	(2.1)	(3.1)	(1.2)	(2.2)	(3.2)
Δ	0.18***			0.28***		
Std. Err.	0			0		
ln(Δ)		775.22			7.89***	
Std. Err.		2.24			0.02	
$\sqrt{\Delta}$			28.28***			3.34***
Std. Err.			0			0
Constant	yes	yes	yes	yes	yes	yes
N	15,178	15,177	15,178	101	100	101
R^2	95.74%	88.71%	99.92%	99.33%	81.98%	96.33%
Bid Side						
Sample	ALL			$\Delta \leq 100$		
Model	(1.3)	(2.3)	(3.3)	(1.4)	(2.4)	(3.4)
Δ	0.23***			0.18***		
Std. Err.	0			0		
ln(Δ)		922.55***			5.12***	
Std. Err.		3.21			0.27	
$\sqrt{\Delta}$			34.41***			2.17***
Std. Err.			0.02			0.05
Constant	yes	yes	yes	yes	yes	yes
N	15,178	15,177	15,178	101	100	101
R^2	98.22%	84.50%	99.47%	99.43%	78.64%	94.36%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.4.4 The Ethereum Order Book

The internal cryptocurrency of Ethereum (ETH) can be traded on cryptocurrency exchanges against fiat money or other cryptocurrencies. In this section we focus specifically on the ETH/USD limit order book to obtain comparability with the Bitcoin and Bitcoin Cash order book, i.e. one tick (Δ) resembles a buy or sell price one US Dollar cent lower or higher than the current best bid or ask price in the ETH/USD LOB.

Figure 5.9 and Figure 5.10 show the volume and the aggregated ask volume dependent on the distance towards the best price for the bid and ask side respectively. We find both sides of the order book to behave almost symmetrical. Interestingly, vol-

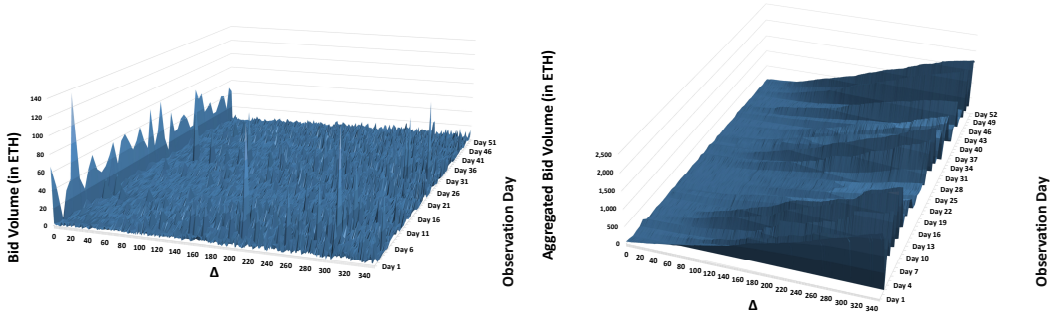


Notes: The figure shows both the average volume (left) and the average aggregated volume (right) dependent on the distance Δ to the best ask price for the first 350 ticks on a daily basis across the sample period. We derive the data by computing the average volume at each tick per day based on snapshots of the order book taken every 10 minutes. Accordingly, each single data point in the above chart represents the average value of 144 observations. Most of the volume is available directly at the spread, however, a pattern of slightly higher volume manifests surrounding definite numbers (100, 200 and 300 ticks away from the best ask price) emerge.

FIGURE 5.9: Ask Volume and Aggregated Ask Volume of the ETH/USD Limit Order Book relative to the Distance (Δ) towards the best Ask Price

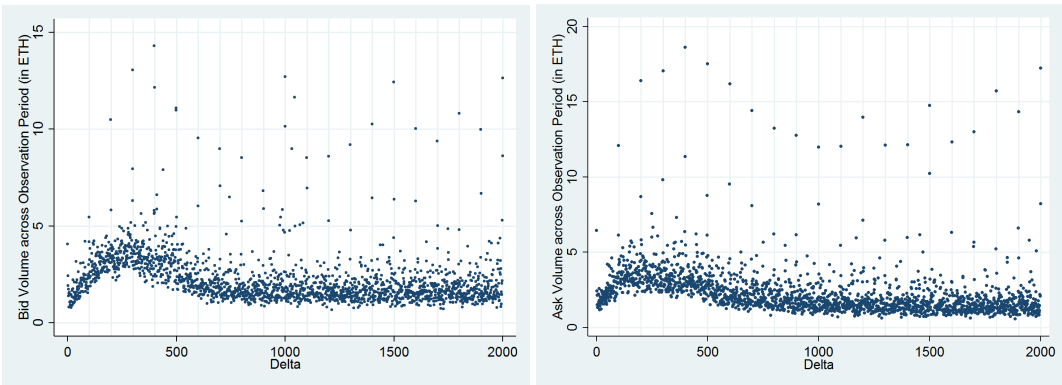
ume peaks at round figures seem to be omnipresent in the ETH/USD order book, supporting our hypothesis that this observation follows a pattern and is also observable across different assets. To get a better picture of the phenomenon, we compute the average volume at each Δ across the observation period (Figure 5.11). For both sides of the order book, we find the available volume to be of magnitudes higher when Δ is a multiple of 100 than at Δs surrounding those numbers. We also find that – analogous to the Bitcoin Cash order book– the average shape of the ETH/USD order book appears to have a maximum away from the best bid or ask price, once more supporting the applicability of the model proposed by Bouchaud, Mézard, and Potters (2002) in a different market.

We further find that the volume peaks seem to mimic the average shape of the order book, i.e. only considering the volume at multiples of 100, we also find a maximum away from the best ask or bid price. This is especially observable for the ask side of



Notes: The figure shows both the average volume (left) and the average aggregated volume (right) dependent on the distance Δ to the best bid price for the first 350 ticks on a daily basis across the sample period. We derive the data by computing the average volume at each tick per day based on snapshots of the order book taken every 10 minutes. Accordingly, each single data point in the above chart represents the average value of 144 observations. Most of the volume is available directly at the spread, however, a pattern of slightly higher volume manifests surrounding definite numbers (100, 200 and 300 ticks away from the best ask price) emerge.

FIGURE 5.10: Bid Volume and Aggregated Bid Volume of the ETH/USD Limit Order Book relative to the Distance (Δ) towards the best Bid Price



Notes: The left (right) figure shows the aggregated bid (ask) volume relative to the best bid (ask) price across the observation period for the ETH/USD limit order book for the first 2,000 ticks. Volume directly at the spread is omitted due to scaling. Volume peaks occur at round numbers at both sides of the ETH/USD limit order book.

FIGURE 5.11: Average Ask and Bid Volume of the ETH/USD Limit Order Book relative to the Distance Delta (Δ) towards the Best Bid or Ask Price

the ETH/USD order book (Figure 5.11). However, we find it difficult to explain this observation. A behavioral explanation could be that there exists a group of "lazy" investors that does not care about incremental tick sizes and only considers a less granular price grid, but still places orders based on a trade off between execution probability and time priority. We discuss and test this hypothesis in Section 5.4.5.
Finally, we take a look at the slope of the ETH/USD order book. Analogous to the analysis of the BTC/USD and the BCH/USD LOB, we compare the fit of three models supposing three different functional forms of the relation between the average aggregate order book volume and Δ:

1. $Aggregated\ Average\ LOB\ Volume_{ETH/USD}(\Delta) = \beta_0 + \beta_1 \times \Delta$
2. $Aggregated\ Average\ LOB\ Volume_{ETH/USD,}(\Delta) = \beta_0 + \beta_1 \times ln(\Delta)$
3. $Aggregated\ Average\ LOB\ Volume_{ETH/USD}(\Delta) = \beta_0 + \beta_1 \times \sqrt{\Delta}$

Similar to the Bitcoin and Bitcoin Cash LOB, the logarithmic model yields the worst fit, while the linear model has the highest explanatory power across all samples. The mediocre performance of both the logarithmic and square-root model hints that the slope of the ETH/USD LOB might be better explained by a convex function.[11]

5.4.5 The Lazy Investor Hypothesis

Based on the discovery of unusual high order book volume at specific Δs away from the best price, which we find consistently across all observed cryptocurrency LOBs during the graphical analysis, we formulate the following null hypothesis:

The Lazy Investor Hypothesis There exists a group of cryptocurrency investors, which disregard the full granularity of the price grid leading to a higher average limit order book volume at Δ ticks away from the best price, if Δ is a multiple of 100.

To our knowledge, this anomaly has not yet been examined in the literature before. We test the statistical significance by constructing a binary variable I_Δ which equals one, if the distance towards the best price is a multiple of 100 and zero otherwise. We compute the coefficient of I_Δ using OLS based on Equation 5.1:

$$Limit\ Order\ Book\ Volume_\Delta = \beta_0 + \beta_1\Delta + \beta_2 I_\Delta + \epsilon,\ where \tag{5.1}$$
$$I_\Delta = \begin{cases} 1,\ if\ \Delta\ =\ 100\, mod(0) \\ 0,\ else \end{cases}$$

To take a linear decline of the average order book volume into account, we include Δ as a control variable in our regression model. If there exist "lazy" investors, our

[11] We allow for a convex or concave shape and directly compare the Ethereum, Bitcoin and Bitcoin Cash LOB in Section 5.4.6.

TABLE 5.7: Slope of the Aggregated Order Book for ETH/USD

Notes: The table shows the slope of the average order Ethereum/US Dollar LOB across our sample period. The slope is derived from a linear regression for different model specifications. Model specifications (1.1)-(3.2) consider the slope of the supply side. Respectively Model specifications (1.3)-(3.4) show the empirical results for the demand side. Each observation has a high level of confidence, as it denotes the average aggregated volume of the order book Δ ticks away from the best price. We measure the volume every ten minutes continuously across our sample period.

Ask Side						
Sample	ALL			$\Delta \leq 100$		
Model	(1.1)	(2.1)	(3.1)	(1.2)	(2.2)	(3.2)
Δ	1.17***			2.07***		
Std. Err.	0			0		
$\ln(\Delta)$		10,901.41			56.21***	
Std. Err.		27.86			3.28	
$\sqrt{\Delta}$			270.40***			24.10***
Std. Err.			0.18			0.71
Constant	yes	yes	yes	yes	yes	yes
N	35,395	35,394	35,395	101	100	101
R^2	98.61%	81.22%	98.44%	99.28%	74.96%	92.14%
Bid Side						
Sample	ALL			$\Delta \leq 100$		
Model	(1.3)	(2.3)	(3.3)	(1.4)	(2.4)	(3.4)
Δ	2.20***			1.54***		
Std. Err.	0			0.02		
$\ln(\Delta)$		19,811.12***			41.24***	
Std. Err.		57.26			2.60	
$\sqrt{\Delta}$			500.71***			17.82***
Std. Err.			0.46			0
Constant	yes	yes	yes	yes	yes	yes
N	35,395	35,394	35,395	101	100	101
R^2	99.81%	77.18%	97.12%	98.49%	71.91%	89.95%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

estimated coefficient for β_2 should deviate significantly from zero. The results for each cryptocurrency are presented in Table 5.8. We find that the coefficient for our constructed variable is indeed highly significant for both sides of the BTC/USD, the BCH/USD, and the ETH/USD LOB, indicating that the LOB volume at specific levels relative to the best price appears to be dependent on whether the level is a multitude of 100, i.e. a round figure.

Key Finding 1: Based on our analysis of the limit order books of BTC, ETH, and BCH, our findings support the lazy investor hypothesis. The empirical results are hard to bring in line with the concept of rational agents, yet the pattern could be explained by a human preference for round figures dominating the desire to achieve the best possible price.

The observed volume peaks also imply a market stabilizing purpose related to price movements: If the price moves into the direction of a volume peak, the peak acts as a dam that curtails heavy price movements, e.g. the price can not drop lower than 100 ticks of the current price without the volume at 100 ticks away from the current price being matched against incoming orders first.[12] Based on our findings, future attempts to explain the shape of the average order book should try to take this effect into account.

TABLE 5.8: Volume peaks in Cryptocurrency Limit Order Books

Notes: This table shows the regression results for the average volume in the Bitcoin, Bitcoin Cash and Ethereum limit order book. I_Δ represents a dummy variable, which equals one if the distance towards the best price (Δ) can be divided by 100 without remainder.

	BTC/USD		**BCH/USD**		**ETH/USD**	
	Ask Side	Bid Side	Ask Side	Bid Side	Ask Side	Bid Side
I_Δ	0.28***	0.35***	1.58***	1.51***	8.11***	9.01***
Std. Err.	0	0	0.04	0.03	0.08	1.36
Δ	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
N	50,412	50,412	15,177	15,177	35,394	35,394
R^2	31.05%	6.00%	12.05%	16.06%	23.58%	2.79%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

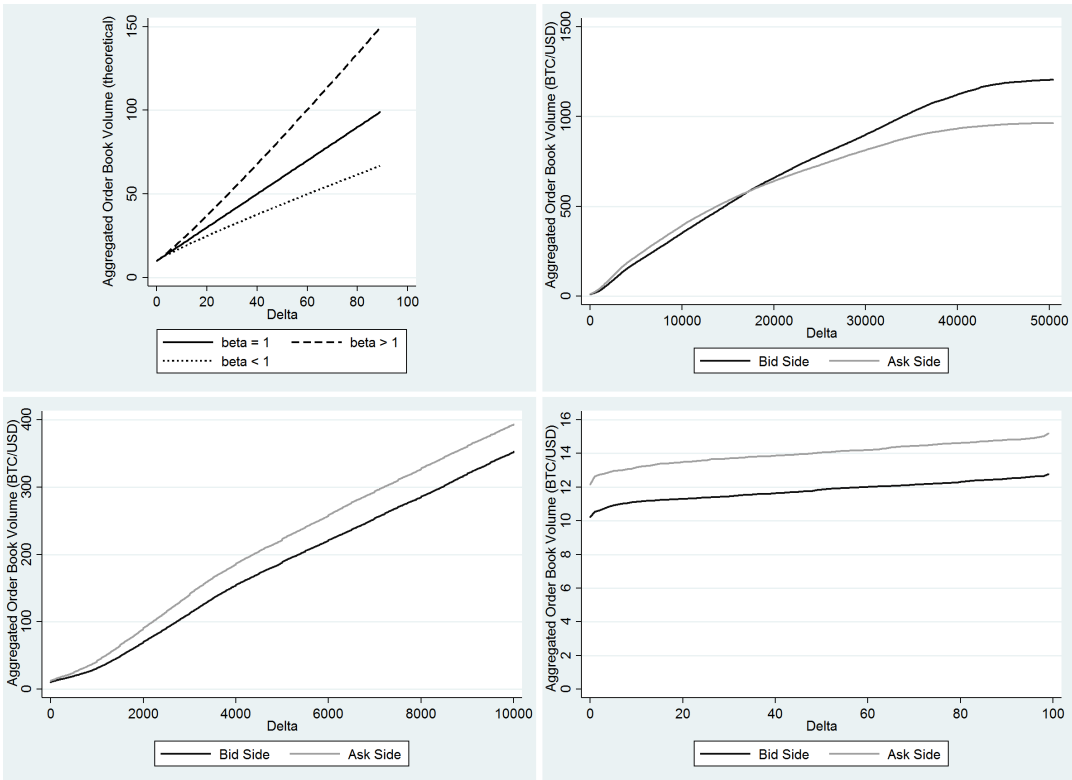
[12]Note that this is a hypothetical case. It is likely that some market participants will cancel their limit order before the price reaches the volume peak. Our finding would also induce that the probability of price movements of increasing magnitude does not decrease steadily.

5.4.6 Convexity of the Order Book Slope

Motivated by our finding that the functional form of the demand and supply curves slightly vary across cryptocurrencies (see Table 5.5, 5.6 and 5.7) we test, whether cryptocurrency order books have a concave or convex demand and supply curve.[13] Table 5.5, 5.6 and 5.7 equally show that we can rule out a logarithmic relation. Thus, we test for convexity by allowing the exponent to vary in a polynomial model, which can be regarded as an extension to the proposed square-root relation model:

$$Aggregated\ LOB\ Volume_{\Delta} = LOB\ Volume\ at\ the\ spread + \Delta^{\beta} \tag{5.2}$$

By applying the natural logartihm to Formula 5.2 we can estimate $\hat{\beta}$ using the



Notes: The upper left picture shows the theoretical aggregated order book volume implied by Equation 5.2 dependent on the β-parameter. The following pictures show the actual aggregated order book volume of the BTC/USD order book derived from our data. Each data point at the respective tick (Δ, Delta) is calculated as the average order book volume at that tick derived from 10 minute snapshots across 94 observation days, i.e. each data point is an average of roughly 13,500 observations. Considering all tick levels (upper right picture), the total aggregated order book volume of the BTC/USD currency pair appears to be concave, implying a β-parameter of less than one.

FIGURE 5.12: Convexity/Concavity of the Aggregated BTC/USD Limit Order Book

[13]The aggregated average order book volume of the ask (bid) side gives the available total volume available at this price. This means that the aggregated average volume can be interpreted as the demand (supply) curve of the respective cryptocurrency.

method of OLS. Consequently, $\hat{\beta} < 1$ ($\hat{\beta} > 1$) indicates a concave (convex) demand or supply curve. The results are presented in Table 5.9 and show that Bitcoin and Bitcoin Cash demand and supply curves are concave while the Ethereum supply and demand curve can be characterized as slightly convex. The direction of our results do not change when we restrict our sample to the first 100 ticks away from the best price. However, it is worthwhile to note that the measured $\hat{\beta}$-coefficients are close to one, justifying a linear approximation in empirical analysis without losing much of the explanatory power.

Key Finding 2: We conclude that it is reasonable to approximate the slope of the order book linearly.

5.5 Informativeness of Order Book Characteristics

In the previous sections we mainly examine static limit order book characteristics. In this section, we are focusing on dynamic relations of the LOB and investigate a potential link between the slope of the order book and cryptocurrency returns. We also hypothesize non-mechanic connections between the order book slope and trading activity.
Næs and Skjeltorp (2006) find a negative link between the order book slope and volume, volatility, and the correlation between volume and volatility in Norwegian stock markets and show that the slope can be regarded as a proxy for disagreement among investors. We follow the approach of Næs and Skjeltorp (2006) in that we also investigate three groups of models relating to the order book slope ($SLOPE_{i,t}$):

1. Price change$_{i,t}$ = $f(SLOPE_{i,t}, ...)$
2. Number of trades$_{i,t}$ = $f(SLOPE_{i,t}, ...)$
3. Correlation(Price change$_{i,t}$, Number of trades$_{i,t}$) = $f(SLOPE_{i,t}, ...)$

The first set of models examines the relation between the price change of a cryptocurrency and the slope of the order book. In an efficient market, price changes occur, when new information about the value of the underlying assets emerge. Observing a link between price changes and the slope of the order book indirectly suggests that the volume of the order book, i.e. the aggregated supply and demand, possibly contains information about the value of the underlying asset.
The second set of models considers the relation between the number of trades and the slope of the order book. Each trader has to choose between buying directly or placing a limit order. We are interested in whether traders consider the existing volume in the order book when placing a new order. A link between the slope of the order book and trading activity could shed light on the dilemma a trader faces when deciding at what price level he should place a new order.

TABLE 5.9: Concavity/Convexity of Aggregated Cryptocurrency Limit Order Books

Notes: We test the concavity (convexity) of the average aggregated volume of of the Bitcoin, Bitcoin Cash and Ethereum LOB across our sample period. Our initial equation equals: $Aggregated\ LOB\ Volume_{\Delta} = LOB\ Volume\ at\ the\ spread + \Delta^{\beta}$. We estimate β using OLS and transforming the initial equation: $ln(Aggregated\ LOB\ Volume_{\Delta} - Volume\ at\ the\ spread) = \beta * ln(\Delta)$. Consequently, $\hat{\beta} < 1$ ($\hat{\beta} > 1$) indicates a concave (convex) demand or supply curve.

Ask Side						
Sample	ALL			$\Delta \leq 100$		
Currency (*/USD)	BTC	BCH	ETH	BTC	BCH	ETH
β-Coeff.	0.64***	0.86***	1.04***	0.17***	0.74***	1.18***
Std. Err.	0	0	0	0.01	0	0.01
Constant	no	no	no	no	no	no
N	50,412	15,177 14	35,394	100	100	100
R^2	99.85%	99.94%	99.94%	88.53%	99.82%	99.50%
p-Value ($H_0: \beta > 1$)	<0.01	<0.01	>0.99	<0.01	<0.01	>0.99
Bid Side						
Sample	ALL			$\Delta \leq 100$		
Currency (*/USD)	BTC	BCH	ETH	BTC	BCH	ETH
β-Coeff.	0.65***	0.87***	1.08***	0.13***	0.59***	1.09***
Std. Err.	0	0	0	0.01	0.01	0.01
Constant	no	no	no	no	no	no
N	50,412	15,177	35,394	100	100	100
R^2	99.71%	99.96%	99.99%	71.66%	99.26%	99.50%
p-Value ($H_0: \beta > 1$)	<0.01	<0.01	>0.99	<0.01	<0.01	>0.99

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The third set of models examines the link between the correlation between price change and the number of trades ("volume-volatility relation") and the slope of the order book. The volume-volatility relationship is well documented for different financial markets including stocks, currencies, oil futures and other derivatives (Foster, 1995, Fung and Patterson, 1999, Sarwar, 2003 and Næs and Skjeltorp, 2006). The relation can be explained, assuming that new information is not incorporated in prices directly but over time. However, the relation could also be explained, in the case that traders do not agree on the impact of new information.[14] Studying the impact of the order book slope on this relation could yield insight into what causes the volume-volatility relation.

5.5.1 Measuring the Slope of the Limit Order Book

We define the average aggregated LOB volume up to a price level of Δ ticks away from the best price for the ask side of the order book as:

$$Avg.\,aggr.\,LOB\,Volume_{\Delta,ask,t,j} = \sum_{i=1}^{\Delta}\left(\frac{1}{N}\sum_{n=1}^{N} VOL_{i,n,ask,t,j}\right) \tag{5.3}$$

$VOL_{i,n,ask,t,j}$ gives the *ask* side LOB volume of currency j at a price level $Ask\,price + i$ during the n-th order book snapshot at day t. As we capture a snapshot of the LOB every ten minutes, N amounts to 144 snapshots per day.
In order to create our slope variable, we estimate the following regression model using OLS:

$$ln(Avg.\,aggr.\,LOB\,Volume_{\Delta,ask,t,j}) = \beta_0 + \beta_{ask,t,j} \times ln(Ask\,Price_{\Delta}) \tag{5.4}$$

$\hat{\beta}_{ask,t,j}$ is the estimated elasticity $\frac{\partial q}{\partial p}$ of the aggregated LOB volume with respect to price and resembles our slope measure $SLOPE_{ask,t,j}$. A higher value of $\hat{\beta}_{ask,t,j}$ resembles a steeper order book. We compute the β-parameter for both sides of the LOB, for each currency and each day. The resulting slope measure can be represented in matrix notation:

$$\boldsymbol{SLOPE}_{ask} = \begin{bmatrix} \beta_{ask,t=1,BTC} & \beta_{ask,t=1,BCH} & \beta_{ask,t=1,ETH} \\ \beta_{ask,t=2,BTC} & \beta_{ask,t=2,BCH} & \beta_{ask,t=2,ETH} \\ ... & ... & ... \end{bmatrix} \tag{5.5}$$

We repeat the above steps analogous for the bid side to compute the average slope of the LOB. We receive our final slope measure by computing the average daily slope of the LOB:

[14] This argumentation closely resembles the second interpretation of the widely used illiquidity measure proposed by Amihud (2002).

$$\boldsymbol{SLOPE} = \frac{\boldsymbol{SLOPE}_{ask} + \boldsymbol{SLOPE}_{bid}}{2} \tag{5.6}$$

We compute ***SLOPE*** considering two different sets of data: ***SLOPE*100** is calculated using only the first 100 ticks away from the best bid or ask price, thus representing the slope of the "inner" order book, whereas ***SLOPE*10000** includes price levels of up to 10,000 ticks – i.e. up to 100 USD – away from the best bid or ask price, capturing information potentially hidden in the depths of the order book (the "deeper" slope).

5.5.2 Order Book Slope and Volatility

In efficient markets, price jumps occur through the arrival of new information that impacts the value of an asset. If all investors agree on the impact of the new information, price adjustments happen without trading. Investors cancel existing limit orders and place new limit orders around the new equilibrium price level considering the updated information state. However, if there is no consensus among investors about the price impact of the new information, trading takes place until all investors that do not agree on the price impact sold or bought their assets, thereby moving the price to it's new equilibrium, where consensus is reached again.
We capture daily price jumps by computing the daily volatility of each cryptocurrency according to the following formula:

$$Volatility_{i,t} = \left| \frac{LOB\ Mid\ price_t}{LOB\ Mid\ price_{t-1}} - 1 \right| \tag{5.7}$$

In general, volatility is expected to be higher in less liquid assets, as the supply and demand side of the LOB is not thick enough to fulfill large market orders without executing limit orders at deeper price levels, moving the mid price in consequence. Spontaneous demand or supply of large quantities lead to larger price changes in illiquid assets, i.e. we would expect the *SLOPE* of illiquid assets to be more gentle *a priori* imposing a positive relationship between *SLOPE* and liquidity. When investigating the link between the order book slope and volatility, we also control for illiquidity, as illiquidity and volatility are likely correlated.
We include trading activity in our regression by deriving the number of trades per day based on Equation 5.8. We include a scaling factor in Equation 5.8 to make the number of trades comparable across days and cryptocurrencies (see Table 5.2):

$$N.of\ Trades_{i,t} = \#Recorded\ Trades_{i,t} \times \frac{1440}{Recorded\ Minutes_{i,t}} \tag{5.8}$$

We also compute the average trade size per day for each cryptocurrency based on Equation 5.9:

$$Trade\ Size_{i,t} = \frac{Trading\ Volume_{i,t}}{N.of\ Trades_{i,t}} \tag{5.9}$$

We further compute the average bid-ask spread ($Spread_{i,t}$) from our data and include the market capitalization in it's logarithmic form ($ln(MCAP)_{i,t}$). The variables $ln(MCAP)_{i,t}$ and $Spread_{i,t}$ are closely tied and proxy for liquidity. We estimate the following linear model:

$$\begin{aligned} Volatility_{i,t} = \beta_0 + \beta_1 SLOPE100_{i,t} + \beta_2 SLOPE10000_{i,t} + \beta_3 N.of\,Trades_{i,t} \\ + \beta_4 Trade\,Size_{i,t} + \beta_5 ln(MCAP)_{i,t} + \beta_6 Spread_{i,t} + c_i + u_{i,t} \end{aligned} \quad (5.10)$$

Using fixed-effects regressions, we get rid of time-constant unobserved effects c_i. Table 5.10 shows the results for six different model specifications. We find that the slope of the inner order book has a significant positive effect on volatility in Model 1, however this effect diminishes, when controlling for the number and size of trades.

Key Finding 3: We conclude that the slope of the inner LOB can not explain return variation.

Models 4-6 consider the LOB volume of the first 10,000 ticks and reveal a positive link between the deeper order book slope and return variation. The effect remains significant at the 10% level across all three model specifications. This finding indicates that a steeper order book slope can be associated with higher volatility. This finding seems to be contradictory at first, as a steeper slope should prevent large price jumps as the LOB holds enough volume to serve large market orders without shifting prices too much. However, as we control for the trading activity and liquidity, the volatility increase due to a steeper slope is likely caused by information, where investors agree upon the price impact, and consequently adjust their active limit orders.

Key Finding 4: Our findings indicate that a steeper slope can be associated with more "non-trading" volume, i.e. investors adjust their limit orders based on new unambiguous information more actively, when there is more LOB volume close to the spread.

This finding could be explained by a lower execution probability of a limit order *ceteris paribus*, when the slope is steep. A steep slope indicates that a lot of the LOB volume is centered around the spread, reducing the execution probability of limit orders deeper in the LOB. Traders observe this reduced execution probability and adjust their limit orders accordingly. Through this channel, limit orders deeper in the LOB are still relevant for the price formation process, even though they do not affect the price mechanically.

We further find that more trades and larger average trade size lead to higher volatility. The effect remains significant when controlling for $ln(MCAP)$ and $Spread$. Arguably, trading activity is often interpreted as a sign of liquidity and we would expect a decrease in volatility, when trading activity is high. However, we control for illiquidity by including $ln(MCAP)$ and $Spread$. Our results hint a very nervous cryptocurrency market environment, where many hectic trades are executed in a short period of time leading to huge price jumps.

TABLE 5.10: Volatility and the Slope of the Order Book

Notes: This table shows the fixed effects regression results for Equation 5.10. Volatility is measured as the daily absolute return of a cryptocurrency. Coefficients of *N. of trades*, *SLOPE100* and *SLOPE10000* are multiplied by 10^5 for better readability. The coefficient of *Trade size*, *ln(MCAP)* and *spread* is multiplied by 10^3.

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
SLOPE100	0.47***	0.36	0.40			
	(2.65)	(1.57)	(1.48)			
SLOPE10000				9.52*	9.43*	9.50*
				(1.72)	(1.71)	(1.66)
N. of trades		0.04***	0.04***		0.05***	0.05***
		(3.86)	(3.61)		(5.20)	(4.10)
Trade size		0.73***	0.71***		0.71***	0.73***
		(4.04)	(3.82)		(4.05)	(3.90)
ln(MCAP)			-2.87			2.73
			(−0.31)			(0.35)
Spread			-1.63			-3.19
			(−0.10)			(−0.2)
Month fixed effects	yes	yes	yes	yes	yes	yes
$N \times T$ (currency-days)	231	186	186	231	186	186
R^2	11.20%	29.20%	29.20%	9.60%	29.30%	29.40%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

5.5.3 Order Book Slope and Trading Activity

In this section, we examine the link between the order book slope and trading activity, measured by the number of trades per day. We compute four different models

based on Equation 5.11. Our results are presented in Table 5.11.

$$N.of\ Trades_{i,t} = \beta_0 + \beta_1 SLOPE100_{i,t} + \beta_2 SLOPE10000_{i,t} + \beta_3 Trade\ Size_{i,t} + \beta_4 ln(MCAP)_{i,t} + \beta_5 Spread_{i,t} + c_i + u_{i,t} \quad (5.11)$$

Key Finding 5: We find that the order book slope is significantly positive related to trading activity. Interestingly, this relation changes sign and becomes negative when considering the slope measure computed from the first 10,000 ticks instead of restricting the data used for computing the slope to the first 100 ticks.

With regard to causality, we argue that it is reasonable to assume that an investor inspects the order book before placing an order and not afterwards, assuming his goal is to make an educated trading decision. If this is the case, it is plausible that the slope of the order book affects trading activity and not vice versa. Næs and Skjeltorp (2006) find the same astounding result in Norwegian stock markets and conclude that the order book slope contains different information based on the depth of the order book used to compute the slope. Our result confirms their findings and shows that a sign change can be observed in cryptocurrency markets as well, even by using an alternative approach to calculate the slope.
We further find that the average trade size does not seem to influence trading activity and that the number of trades is higher when the spread increases. A narrow bid-ask spread generally reduces trading costs which should amplify trading *per se*. However, upon closer inspection we find that the spread in our data equals the minimum tick size of 0.01 USD in 42.02% of all observations, whereas the average spread is 0.50 USD, indicating that when the spread deviates from the minimum tick size, the deviation is quite large. Such a change in the spread could be interpreted by market participants as a trading signal, which would explain our empirical results.

5.5.4 Order Book Slope and the Volume-Volatility Relation

Next, we examine the interplay between the volume-volatility relation and the slope of the order book. The volume-volatility relation is a well known phenomenon observed across many markets, describing the empirical observation of high price volatility coupled with high trading volume which has been confirmed by a variety of studies (see Karpoff, 1987).
In our data, we find a positive correlation between volatility and daily trading volume of 19.19% and a positive correlation between volatility and the number of trades of 26.44%, indicating that the volume-volatility relation can be observed in cryptocurrency markets as well. We are interested in the cause of this relationship and different theoretical models have been proposed mainly focusing on market efficiency, the informedness of traders and speculative trading (Glosten and Harris,

TABLE 5.11: Trading Activity and the Slope of the Order Book

Notes: This table shows the fixed effects regression results for Equation 5.11. The number of trades (*N. of trades*) is the dependent variable.

Model	**(1)**	**(2)**	**(3)**	**(4)**
SLOPE100	12.03***	4.16**		
	(8.73)	(2.23)		
SLOPE10000			67.92	-67.36*
			(1.43)	(−1.71)
Trade size		-43.95		-123.90
		(−0.34)		(−0.96)
ln(MCAP)		24,162.50***		32,670.30***
		(4.03)		(6.77)
Spread		4,531.50***		5,536.40***
		(4.41)		(5.45)
Month fixed effects	yes	yes	yes	yes
$N \times T$ (currency-days)	186	186	186	186
R^2	55.20%	63.40%	36.50%	62.90%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

1988, Foster and Viswanathan, 1995). We test, whether the order book slope, i.e. the availability of potential trading volume can help to explain this relation. Næs and Skjeltorp (2006) find that the slope of the order book is significantly negatively related to the volume-volatility relation, concluding that a stronger volume-volatility relation is associated with a flat slope of the order book. However, the authors do not directly interpret their results. Similar to Næs and Skjeltorp (2006), we investigate this relationship by computing the daily correlation coefficient $Corr(N.of\,trades_{i,t}, |R_{i,t}|)$ measured over a month between the number of trades and the absolute return.[15] Næs and Skjeltorp (2006) argue that the number of trades is the crucial component of trading volume. Regressing $Corr(N.of\,trades_{i,t}, |R_{i,t}|)$ on our slope measure using a fixed-effects regression model, we only find weak support for the negative relationship between the volume-volatility relation and the slope of the order book. We compute the volume-volatility relation alternatively by directly incorporating daily trading volume: $Corr(Trading\,Volume_{i,t}, |R_{i,t}|)$ and estimate the following regression model:

$$\begin{aligned} Corr(Trading\,Volume_{i,t}, |R_{i,t}|) = \beta_0 + \beta_1 SLOPE100_{i,t} + \beta_2 SLOPE10000_{i,t} \\ + \beta_3 Trade\,Size_{i,t} + \beta_4 ln(MCAP)_{i,t} + \beta_5 Spread_{i,t} + c_i + u_{i,t} \end{aligned} \quad (5.12)$$

Key Finding 6: Our results indicate that there is a significant amount of information in the order book and the slope of the order book should be considered in theoretical models trying to explain the cause of the volume-volatility relation. We find that the volume-volatility relation seems to be stronger when the slope of the order book is steeper, which contradicts the results of Næs and Skjeltorp (2006).

The effect appears to be robust across model specifications and can be observed for both slope measure specifications and increases in magnitude, when controlling for market capitalization. Further, the explanatory power in all models presented in Table 5.12 is approximately 60% and does not vary much between models. It is noteworthy that the first and the fourth model already explain 60.80% and 59.00% of the variation in the volume-volatility relation.

5.5.5 A Note on Causality

While it is generally difficult to derive causality in economics, it is especially challenging in a market microstructure setting. However, some conjectures with respect to causality can be derived economically. To further increase the robustness of our results, we perform additional causality tests on our slope measure, trading activity and volatility for different sub periods. Granger causality test results are depicted in Table 5.13 and Table 5.14. We find that the null can be rejected at the 5% level for all

[15] We do not adjust for day-of-week effects as proposed by Næs and Skjeltorp, 2006 as cryptocurrencies can be traded at any time.

TABLE 5.12: The Volume-Volatility relation and the Slope of the Order Book

Notes: This table shows the fixed effects regression results for Equation 5.12. The dependent variable, $Corr(Trading\ Volume_{i,t}, |R_{i,t}|)$ is the daily correlation coefficient measured over a month between USD trading volume and the absolute return. Coefficients of *SLOPE100* and *SLOPE10000* are multiplied by 10^4 for better readability. The coefficient of *Spread* and *Trade size* has been multiplied by 10^2.

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
SLOPE100	0.37***	0.37***	0.44***			
	(3.68)	(2.98)	(3.57)			
SLOPE10000				5.60*	5.97*	5.61*
				(1.90)	(1.81)	(1.90)
Spread	0.71	0.84		1.61**	1.36	
	(0.89)	(0.98)		(2.10)	(1.62)	
Trade size		−0.03			0.11	
		(−0.28)			(−1.08)	
ln(MCAP)			−0.02			0.07**
			(−0.43)			(2.14)
Month fixed effects	yes	yes	yes	yes	yes	yes
$N \times T$ (currency-days)	231	186	231	231	186	231
R^2	60.80%	61.50%	60.70%	59.00%%	60.30%%	59.10%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

tests performed on the granger-causal relationship between number of trades and the slope of the order book (Table 5.13), indicating a bidirectional granger-causal relationship between the two variables.[16] Examining the causal relationship between volatility and the slope of the order book (Table 5.14) reveals a less unambiguous picture. We can not reject bidirectional Granger causality between volatility and slope changes for Bitcoin at the 10% level and for Bitcoin Cash at the 1% level, while we find evidence for unidirectional causality from volatility to slope changes for Ethereum.

[16]For most of the tests shown in Table 5.13 the null hypothesis can even be rejected at a 1% significance level.

TABLE 5.13: Linear Granger Causality Test Results on Trades

Notes: This table shows Granger causality test results between the number of trades and order book slope changes. Lag lengths are set with the Akaike (1974) information criterion adjusted by the number of observations, considering the first five lagged days. Sig denotes the marginal significance level of the computed χ^2-statistic used to test the zero restrictions implied by the null hypothesis of Granger noncausality. Slope indicates the number of ticks considered for estimating the order book slope measure.

			H_0: No. of trades do not cause slope changes.		H_0: Slope changes do not cause No. of trades.	
Lags	N	Slope	χ^2	Sign	χ^2	Sign
Panel A: Bitcoin (April 2018 – August 2019)						
5	18	100	20.77	0.00	28.72	0.00
5	18	10,000	118.83	0.00	24.96	0.00
Panel B: Bitcoin Cash (May 2018 – August 2019)						
4	13	100	196.17	0.00	11.33	0.02
4	13	10,000	13.15	0.01	46.19	0.00
Panel C: Ethereum (April 2018 – August 2019)						
5	15	100	66.30	0.00	15.94	0.01
5	15	10,000	80.31	0.00	15.67	0.01

TABLE 5.14: Linear Granger Causality Test Results on Volatility

Notes: This table shows Granger causality test results between volatility ($|R_{i,t}|$) and order book slope changes. Lag lengths are set with the Akaike (1974) information criterion adjusted by the number of observations, considering the first five lagged days. Sign denotes the marginal significance level of the computed χ^2-statistic used to test the zero restrictions implied by the null hypothesis of Granger non-causality. Slope indicates the number of ticks considered for estimating the order book slope measure.

			H_0: Volatility does not cause slope changes.		H_0: Slope changes do not cause volatility.	
Lags	N	Slope	χ^2	Sign	χ^2	Sign
Panel A: Bitcoin (April 2018 – August 2019)						
5	22	100	10.27	0.07	41.96	0.00
2	51	10,000	12.32	0.00	6.62	0.04
Panel B: Bitcoin Cash (May 2018 – August 2019)						
5	14	100	25.50	0.00	234.72	0.00
5	14	10,000	22.54	0.00	17.65	0.00
Panel C: Ethereum (April 2018 – August 2019)						
1	68	100	8.54	0.00	0.01	0.75
1	68	10,000	2.49	0.12	0.23	0.63

5.6 Conclusion

This study examines the statistical properties of cryptocurrency limit order books. Our descriptive analysis reveals that the secondary market for cryptocurrencies is shaped by high cancellation rates and a preference for limit orders over market orders. These findings are valid for both the buy and sell side. Our subsequent empirical analysis reveals six key findings:

Building upon our in-depth analysis of the LOBs of three major cryptocurrencies, we find evidence supporting our hypothesis that a group of "lazy" investors disregard the full granularity of the price grid, which is reflected in volume peaks at certain price levels distant to the best price (**Key Finding 1**). Testing the empirical fit of different explanatory models for the slope of the order book, we find that a linear approximation of the slope of the order book is reasonable without losing much explanatory power (**Key Finding 2**). Employing three different models, we gain empirical evidence on the interplay between the slope of the order book, price changes and trading activity. We compute two slope measures (the "inner" slope and the "deeper" slope) different in the range of data used for estimation. The empirical analysis further reveals that the inner slope can not explain return variation, while the deeper slope seems to contain information about cryptocurrency returns (**Key Finding 3**). This finding indicates that limit orders in the depths of the order book – even though having a low probability to be executed and not being mechanically linked to price changes – are still relevant for the price formation process. This finding suggests that traders incorporate the whole state of the order book when buying or selling cryptocurrencies (**Key Finding 4)**. This explanation is also practically reasonable, as the state of the limit order book is visible to traders at any given time. In addition, we find that the inner slope of the order book has a significant positive effect on trading activity. However, this relation changes sign when considering the deeper slope of the order book (**Key Finding 5**). This phenomenon has also been observed in stock markets by Næs and Skjeltorp (2006). Our results point into the same direction and show that this anomaly can be observed in cryptocurrency markets as well. Moreover, we find a positive relationship between trading volume and volatility, confirming the volume-volatility to be prevalent in cryptocurrency markets as well. Surprisingly, we find the relation to be weaker when the slope of the order book is steeper (**Key Finding 6**). This finding is significant across six different model specifications and contradicts the empirical results presented by Næs and Skjeltorp (2006).

We further find that the spread equals the minimum tick size of 0.01 USD in 42.02% of the time in our data, raising questions about the perception of the endogeneity of the spread by traders in today's markets. This issue emerges from the finite granularity of the price grid and has also been discussed by Biais, Hillion, and Spatt (1995).

Chapter 6

Heuristics in Cryptocurrency Limit Order Placement

Exchanges and trading platforms allow us to analyze the behavior of interacting market participants. From a scientific point of view, the main advantage of financial markets is that all market participants trade the same asset with the aim to maximize profits, incentivizing rational behavior. This unique setting allows to study not only economic and financial theories, but also theories of human behavior. In this study, we show that a power-law used to describe the distribution of limit order prices in stock markets can be extended to cryptocurrency markets, despite the very different market frameworks. We hypothesize that cryptocurrency traders fall back to heuristics when placing limit orders, provide a straightforward model extension that accounts for this behavior, and show that our model fits the empirical data better than the vanilla power-law model proposed in the literature.

6.1 Introduction

In this study, we examine the probability of incoming orders in a limit order market. Today, most of security trading is arranged via electronic order matching by exchanges operating a limit order book (LOB). The flow of incoming limit orders yields insights into market dynamics at the fine-granular level and therefore receives particular interest from academia. Zovko and Farmer (2002) and Bouchaud, Mézard, and Potters (2002) find a striking behavioral pattern while observing the placement of limit orders in stock markets. Bouchaud, Mézard, and Potters (2002) state that the distribution of incoming limit order prices depends on the distance towards the best available price and can be described by a *universal* power-law. Further studies provide empirical evidence that supports the validity of the proposed power-law for varying stocks and time frames (see Potters and Bouchaud, 2003, Mike and Farmer, 2008 and Cont, Stoikov, and Talreja, 2010).

We extend the current state of scientific knowledge by showing that the *universal* validity of the power-law can be generally extended to cryptocurrency markets. Accompanied by the increasing popularity and adoption of cryptocurrencies,

secondary markets for cryptocurrencies emerge. Those trading platforms operate LOBs in the same way as traditional stock exchanges and similar trading rules apply.[1] However, there exist some considerable differences between traditional exchanges and cryptocurrency trading platforms as well. Cryptocurrency trading platforms typically operate globally and without a break, while traditional (national) exchanges only provide services during localized business hours. Further, traded volume in stock markets is of magnitudes higher than cryptocurrency trading volume and important characteristics of market participants likely differ as well between these market places, e.g. investment horizon, trading strategy or location. Those market conditions raise doubts on the unconditional transferability of the power-law to cryptocurrency markets.

Using cryptocurrency limit order flow data from a major cryptocurrency trading platform, we find that order placement is increased when the relative distance towards the best price equals an (positive) integer. We explain our finding by traders using a heuristic when placing limit orders. Supposing that traders reduce the complexity of order placement by not considering the full granularity of the price grid, we propose a straightforward extension of the power-law relation and show that the extended model fits the empirical distribution of incoming limit order prices. The appeal of this extension lies in it's simplicity, which reflects the simplification made by traders during limit order placement. The existence of a substantial amount of traders that rely on a simple heuristic when placing limit orders might indicate that the cryptocurrency market is still an emerging market, where inefficiencies exist. To our knowledge, no previous study that focuses on stock markets detects this placement behavior. The remainder of this investigation is structured as follows: Section 6.2 provides the theoretical background of the proposed power-law in limit order markets. We motivate our hypothesis in Section 6.3. We derive our theoretical model and provide empirical results in Section 6.4. Section 6.5 concludes.

6.2 Statistics and Distribution of Incoming Limit Order Prices

In a limit order market, traders submit a limit order to buy or sell a quantity of an asset for a specific price. Limit orders at the highest bid or lowest sell price are matched against incoming market orders while other limit orders remain in the LOB until they are either executed when the current price reaches their price level at a later point in time or canceled by the trader. Similar to Bouchaud, Mézard, and Potters (2002), we denote Δ as the absolute difference measured in ticks on the US Dollar price grid between the current best price and the price of an incoming limit order:

$$\Delta = |\textit{best available price} - \textit{limit order price}|$$

[1] Notably, most cryptocurrency trading platforms operate a LOW that follows a first-come, first serve principle and a price-time priority.

Note that Δ can be computed for bid and ask limit orders similarly. An impatient trader chooses Δ to be close to zero, thereby increasing the likelihood of quick order execution.[2] However, the time advantage of placing a limit order close to the current bid or ask price runs in opposition to the risk of having the order executed at an unfavorable price. Hence, each trader faces a trade-off when placing a limit order. Zovko and Farmer (2002) state that the choice of placing a limit order also depends on the individual goal of a trader and his trading strategy making order placement a complex task. Based on the assumption that traders differ in their expectations of future returns, time horizon and risk aversion, Chiarella, Iori, and Perelló (2009) show that heterogeneous trading rules impact the limit order flow. Hence, the distribution of incoming limit order prices is not *a priori* clear.
By analyzing stock LOBs from the Paris Bourse, Bouchaud, Mézard, and Potters (2002) detect that the probability of a limit order arriving at Δ can be described by a *universal* power-law of the following form:[3]

$$\rho(\Delta) \sim \frac{1}{\Delta^{1+\mu}}$$

The advantage of power-law distributions lies in their simple representation and their prevalence in real world data. Power-laws can be found across a wide range of complex economic relations that are influenced by many independent factors. In fact, the well-known pareto distribution (Pareto, 1964), which is widely applied in economics, is a power-law.
Using stock order flow data of three listed stocks, Bouchaud, Mézard, and Potters (2002) estimate that $\hat{\mu} = 0.6$. While subsequent studies validate the power-law distribution, disagreement persists about the true value of μ.

6.3 Limit Order Placement

We gather meta data by tracking the limit order flow via the application programming interface of a large cryptocurrency exchange. Our data originates from four cryptocurrencies traded against the US Dollar (USD) in April and June 2018, namely Bitcoin (BTC), Ethereum (ETH), Bitcoin Cash (BCH) and Litecoin (LTC).[4] In total, we gather $9,625,526$ BTC/USD, $34,905,938$ ETH/USD, $17,827,541$ BCH/USD and $33,467,649$ LTC/USD limit order prices with a maximum of 500 ticks away from the current best price. Supplemental information about the cryptocurrencies in our sample is given in Table 6.1. Figure 6.1 shows the empirical distribution of incoming

[2] $\Delta = 0$ implies a limit order at the best bid price or ask price.

[3] Note that the proposed power-law is scale-invariant, i.e. if Δ is multiplied by a constant c we would derive the following direct proportionality: $\rho(c\Delta) = \frac{1}{(c\Delta)^{1+\mu}} = c^{-(1+\mu)}\rho(\Delta) \propto \rho(\Delta)$, where $\propto$ denotes direct proportionality.

[4] BTC, ETH, BCH and LTC combined account for roughly 80% of the total cryptocurrency market capitalization.

TABLE 6.1: Descriptive Statistics

Notes: This table contains some supplemental information about the four cryptocurrency pairs BTC/USD, ETH/USD, BCH/USD and LTC/USD in our sample.

	BTC/USD	**ETH/USD**	**BCH/USD**	**LTC/USD**
Initial price (USD)	8186.01	524.35	1262.54	140.92
Final price (USD)	6180.03	413.99	665.51	79.49
Tick size (USD)	0.01	0.01	0.01	0.01
Avg. daily transaction volume (thsd.)	9.31 BTC	107.21 ETH	21.04 BCH	202.25 LTC
Avg. daily transaction volume (USD mn.)	77.63	67.23	20.90	19.98
Total # limit orders[1] (mn.)	9.63	34.91	17.83	33.47
# limit ask orders (mn.)	5.16	19.55	9.73	17.99
# limit bid orders (mn.)	4.46	15.35	8.10	15.48

[1] Numbers only consider limit orders with $\Delta \leq 500$ that were recorded from April to June 2018 (with gaps).

limit order prices dependent on the distance towards the best price (Δ) in logarithmized form up to a maximum distance of $\Delta_{max} = 500$ for BTC, ETH, BCH and LTC. The depicted values are aggregated across the bid and ask side.
From visual inspection, the power-law seems to fit the distribution of limit order prices quite well considering incoming BTC/USD and LTC/USD limit orders. However, the distribution of incoming limit orders for ETH/USD reveals a deviating picture and shows that while most orders are placed close to the best price, a local maximum exists at about 100 to 200 ticks away from the current best price. This local maximum can be observed in the BCH/USD order placement as well and is even more pronounced.[5] Nevertheless, the observed cryptocurrencies consistently show that the probability of order placement diminishes when moving away from the best price, as implied by the power-law, leading to our first hypothesis:

Hypothesis A ($H_{0,A}$): The distribution of incoming limit order prices in cryptocurrency markets follows a power-law like the order flow in stock markets.

Analyzing Figure 6.1 reveals that the distribution of incoming limit orders exhibits peaks which can be observed across all four currencies to a varying degree. They are most eminent in the limit order flow of the BTC/USD currency pair. We suppose that these peaks do not occur at random but follow a simple rule originating in a traders' heuristics used to reduce the complexity of the order placement decision process, which leads to our second hypothesis:

Hypothesis B ($H_{0,B}$): The probability of an incoming limit order is increased, when the distance towards the best price (Δ) divided by 100 is a positive integer.

[5] While it is not the focus of this study, we suppose that the humps are related to the liquidity of the respective asset.

Note that an incoming limit order at $\Delta = 100$ refers to a limit order placed at exactly 1.00 USD away from the best price. To illustrate our hypothesis, we include vertical lines in Figure 6.1 highlighting the associated values for $\Delta = [100, 200, ..., 500]$. We find the respective vertical lines to be located exactly at the peaks of the distribution across all observed cryptocurrencies.

6.4 Empirical Results

To empirically fit the power-law proposed by Bouchaud, Mézard, and Potters (2002) and in order to empirically test it's *universal* character by applying it in cryptocurrency markets ($H_{0,A}$), we estimate the value of the parameter μ. We do so by transforming the proposed power-law equation by taking the natural logarithm at both sides:

$$P(\Delta) = \frac{c}{\Delta^{1+\mu}}$$

$$\Rightarrow \underbrace{ln\left(P(\Delta)\right)}_{Y} = \underbrace{ln(c)}_{\beta_0} + \underbrace{(-1-\mu)}_{\beta_1} \cdot \underbrace{ln(\Delta)}_{X_1}$$

$P(\Delta)$ denotes the frequency of a limit buy or limit sell order arriving Δ ticks away from the current best price. As indicated by the brackets, the transformed equation can be substituted to derive an equation that is linear in its parameters. Using the method of ordinary least squares, this property allows us to compute an estimator for μ by estimating the value of $\hat{\beta}_1$, as $\hat{\mu} = -\hat{\beta}_1 - 1$.

Motivated by our graphical analysis of order placement behavior (Figure 6.1) and $H_{0,B}$, we extend this model by including a factor λ and a function $D_{\mathbb{N}}(\Delta)$ that depends on the value of Δ and has the value 1, if $\frac{\Delta}{100} \in \mathbb{N}$ and zero otherwise. We propose the following expression describing the distribution of incoming limit order prices:

$$P(\Delta) = \frac{c \cdot e^{\lambda \cdot D_{\mathbb{N}}(\Delta)}}{\Delta^{1+\mu}}$$

$$\Rightarrow \underbrace{ln\left(P(\Delta)\right)}_{Y} = \underbrace{ln(c)}_{\beta_0} + \underbrace{(-1-\mu)}_{\beta_1} \cdot \underbrace{ln(\Delta)}_{X_1} + \underbrace{\lambda}_{\beta_2} \cdot \underbrace{D_{\mathbb{N}}(\Delta)}_{X_2}$$

As shown above, the extended model can be transformed by applying the logarithm at both sides as well, allowing us to estimate its parameters analog to the vanilla power-law model. We measure $D_{\mathbb{N}}(\Delta)$ by creating a dummy variable, indicating whether the distance of an incoming limit order price divided by 100 is a natural number, i.e. whether the price of an incoming limit order i is $1, 2, ..., 5$ USD away from the best price. ϵ denotes the error term. Our final empirical model is shown in

Formula 6.1:

$$ln(P(\Delta))_i = \beta_0 + \beta_1 ln(\Delta)_i + \beta_2 D_{\mathbb{N}}(\Delta)_i + \epsilon_i, \; where \tag{6.1}$$
$$D_{\mathbb{N}}(\Delta) = \begin{cases} 1, \; if \; \frac{\Delta}{100} \in \mathbb{N} \\ 0, \; else \end{cases}$$

Our empirical results are shown in Table 6.2. We find $\hat{\beta}_1$ and $\hat{\mu}$ respectively to be

TABLE 6.2: Fitted Power Law Results

Notes: This table provides regression results for Equation 6.1. Note that we estimate μ by computing $\hat{\mu} = -(\hat{\beta}_1 + 1)$. To capture potential measurement errors, we allow a narrow interval of $[\mathbb{N} - 0.05; \mathbb{N} + 0.05]$ for which the dummy variable $D_{\mathbb{N}} = 1$. Data was recorded from April to June 2018 (with gaps). The exponent of the denominator of the power-law is denoted as $1 + \mu$, i.e. a negative value > -1 for $\hat{\mu}$ is not surprising.

Currency (*/USD)	**BTC**	**BTC**	**ETH**	**ETH**	**BCH**	**BCH**	**LTC**	**LTC**
$\hat{\mu}$	-0.66***	-0.65***	-0.26***	-0.25***	-0.78***	-0.77***	0.85***	0.85***
	(-26.70)	(-28.95)	(-31.42)	(-31.82)	(-6.88)	(-7.06)	(-55.77)	(-56.00)
$D_{\mathbb{N}}(\Delta)$		0.31***		0.23***		0.22**		0.25**
		(7.96)		(3.04)		(2.03)		(2.36)
Constant	yes	yes	yes	yes	yes	yes	yes	yes
Adj. R^2	58.80%	63.38%	66.40%	66.94%	8.49%	9.06%	86.17%	86.30%
N	500	500	500	500	500	500	500	500

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, t-statistics in parenthesis.

statistically significant at the 1%-level across all cryptocurrencies and model specifications, indicating that the postulated *universal* nature of the power-law can be extended to cryptocurrency markets as well, thereby strongly boosting it's external validity.

Interim Result A: Therefore, we can not reject $H_{0,A}$.

The explanatory power of the ordinary power-law of up to 86.17% (LTC/USD) for cryptocurrencies indicates that the power-law may represent a fundamental characteristic of competitive limit order markets and should hence be considered an asset-independent feature in the fields of market microstructure. As cryptocurrency markets did not yet exist, when the power-law distribution was first discovered, our findings provide empirical evidence that the power-law distribution is an inherent property of competitive markets.
Moreover, our results show that the estimated parameter $\hat{\mu}$ varies between -0.78 to 0.85, which is considerably below the estimated value in stock markets, ranging from 0.6 to 1.5 (see Bouchaud, Mézard, and Potters, 2002 and Zovko and Farmer, 2002).

A low value for μ implies that more limit orders are placed away from the current price in cryptocurrency markets. This finding suggests that cryptocurrency traders expect larger USD price jumps in cryptocurrency markets, which would be in line with the general perception of cryptocurrencies being more volatile than stocks.
We further find that the coefficient λ of our constructed variable $D_{\mathbb{N}}(\Delta)$ is statistically significant and of the same magnitude across all four cryptocurrencies.

Interim Result B: Hence, we can not reject $H_{0,B}$.

The results show that the probability of a specific limit order price is higher, when the limit order price is exactly $1.00, 2.00, ..., 5.00$ USD away from the current best price. We suggest that this occurrence is of behavioral nature and traders prefer rounded numbers when setting the price level of limit orders, i.e. they follow heuristics when placing limit orders. Obviously, traders seem to base their decision on the relative distance towards the best price rather than the best price itself. E.g., our results suggest that a trader does not contemplate about "buying at 50 USD" but rather about "buying at 5 USD below the current price".[6] Including $D_{\mathbb{N}}$ in our regression model also increases the explanatory power by 0.13 to 4.58 percentage points.

Key Finding: We argue that while the simple form of the power-law is appealing in explaining the limit order price distribution, the existence of behavioral preferences of traders makes it necessary to account for specific values of the relative distance towards the best price in future theoretical models.

To our knowledge, the existence of peaks in order price frequencies has not been observed by studies focusing on highly efficient stock markets.
Figure 6.2 compares the fit of the *universal* power-law proposed by Bouchaud, Mézard, and Potters (2002) with the fit of our extended model exemplary for incoming BTC/USD limit orders. Figure 6.2 demonstrates that solely using the power-law to describe the distribution of limit order prices leads to a slight underestimation of $P(\Delta)$ – especially close to the best price – which can be reduced by accounting for peaks.

6.5 Conclusion

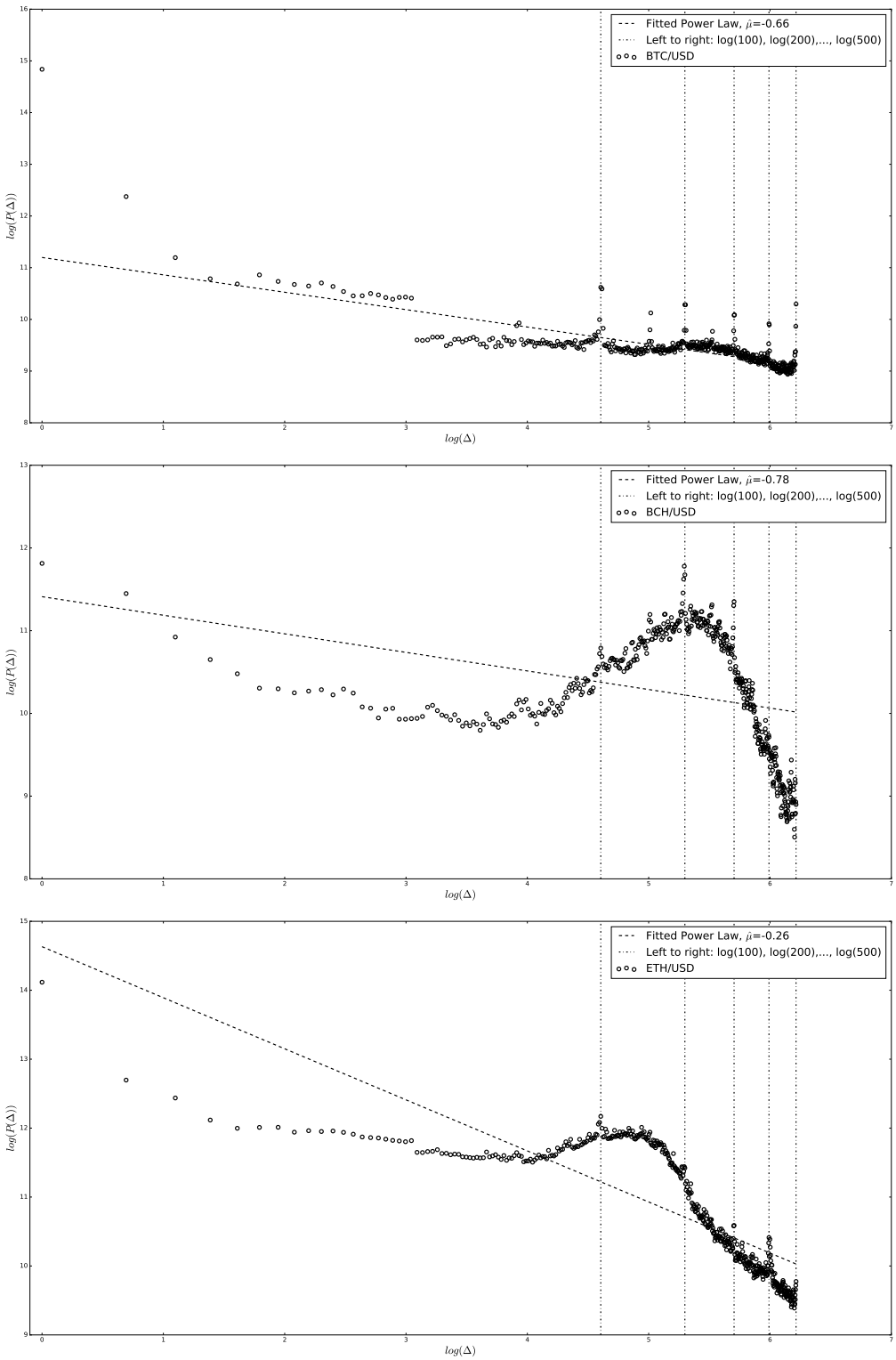
In this study, we examine the limit order placement of four major cryptocurrencies and compute the relative distance of incoming limit order prices towards the best price at arrival. We show that limit order placement in cryptocurrency markets can

[6]Preferences of financial agents for rounded numbers have been documented in the literature before, e.g. Corwin (2003) finds that the issue yield in seasoned equity offers is related to underwriter pricing conventions such as price rounding and pricing relative to the bid quote. Further links between numeric fluency and human preferences have been documented by Kettle and Häubl (2010).

be described by a power-law that was first discovered in the stock market literature. Further, we find empirical evidence that traders prefer integers when considering how far from the best price they place their limit order. This preference for integers is likely the result of applied heuristics during limit order placement. Traders resort to heuristics when dealing with the complexity of limit order placement and reduce the complexity by disregarding the full granularity of the price grid.

We suggest a straightforward extension of the power-law approach to describe the distribution of limit order prices and show that accounting for the observed behavioral regularity yields a better empirical fit than the plain power-law approach suggested in the literature. Given the statistical significance of our finding, we suggest that future models trying to explain limit order placement should take this behavioral regularity into account.

Further studies are necessary to better understand the factors that influence the limit order placement process in cryptocurrency markets over time. Our data gives hints to the fact that the observed phenomenon of an increased probability for incoming limit orders at certain numbers might be scale-invariant in that more incoming limit orders are placed at 0.1 USD than at 0.11 USD (0.5 USD than at 0.49 USD) away from the best price. However, our extended model needs to be tested in other markets and time frames. We would expect that it can show an improved explanatory power, especially in illiquid and emerging markets. Finally, the occurrence of humps in the limit order price distribution in our data indicates a potential link between limit order prices and liquidity. This complex relation needs further research and is crucial to fully understand and describe order placement behavior.



Notes: This figure shows the cumulative distribution of Δ of incoming orders as a function of Δ for BTC/USD, BCH/USD, ETH/USD and LTC/USD. We take the logarithm at both sides. $\Delta \leq 500$.

FIGURE 6.1: Incoming Limit Orders and Fitted Power-Law

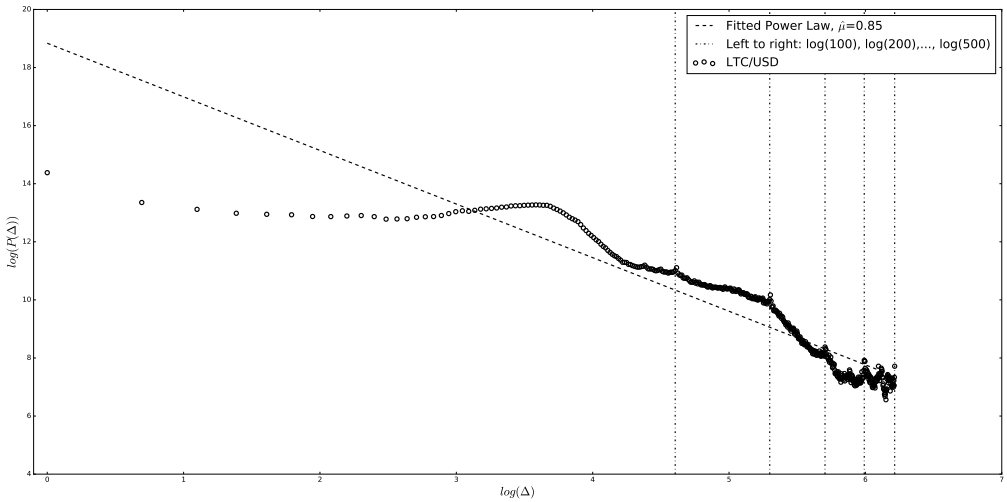
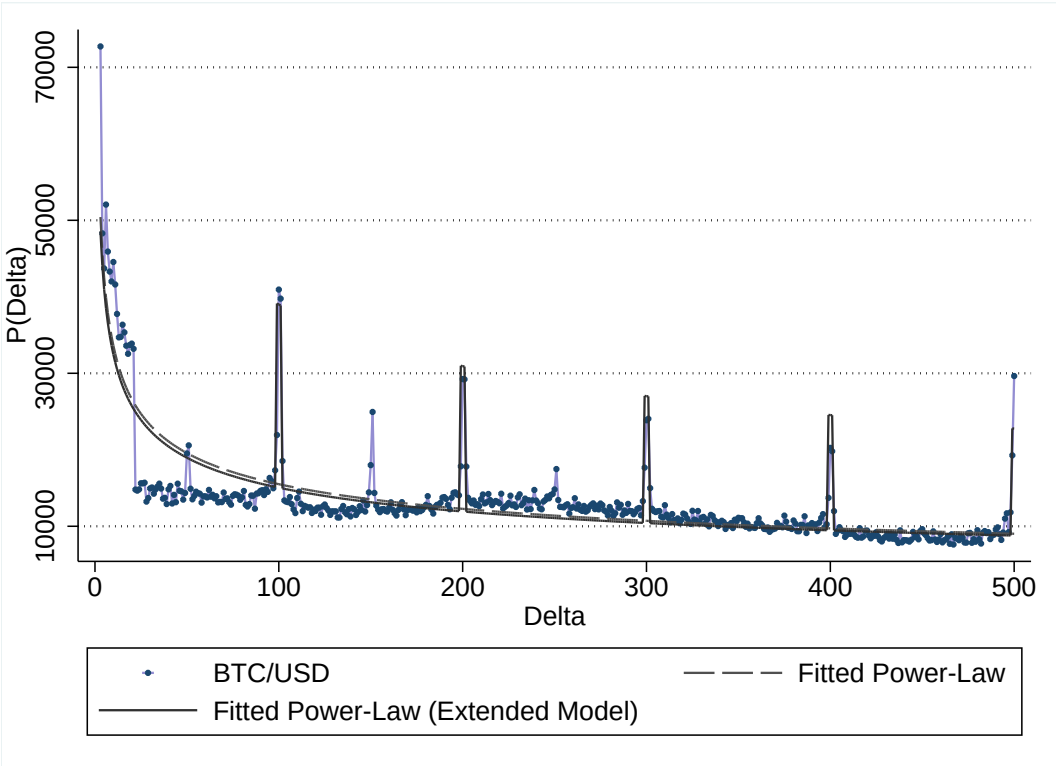


FIGURE 6.1: (continued)



Notes: We set an interval of $[\mathbb{N} - 0.01; \mathbb{N} + 0.01]$ for which the dummy variable $D_{\mathbb{N}} = 1$. Note the decreasing magnitude of the peaks in the extended model caused by the parameter λ, which also fits the empirical data. The predicted peaks decrease because λ is incorporated multiplicative rather than additive in the extended model. We estimate $\hat{\lambda}$ as the β-coefficient of the dummy variable $D_{\mathbb{N}} = 1$.

FIGURE 6.2: Fitted Power-Law and Extended Model for BTC/USD

Chapter 7

Reaktionen der Kryptowährungsmärkte auf die COVID-19-Pandemie

Chapter 7 has been published as a journal article:
Häfner, David, Jianan He, and Dirk Schiereck (2020). "Reaktionen der Kryptowährungsmärkte auf die COVID-19-Pandemie". In: *Corporate Finance* 05-06, pp. 141–144, ISSN 2198-8889.

Chapter 8

Concluding Remarks

This dissertation explores the behavior and interaction of market participants in different market conditions and is focused on empirically analyzing the efficiency of various financial market segments and the behavior of market participants under competition and their reaction to exogenous shocks and imperfect information.

In Chapter 2, we examine the development of corporate social responsibility and its compatibility with efficient investments. By reviewing 74 leading articles published between 1984 and 2016, we provide a review on the scientific literature and highlight the importance and effects of socially responsible and sustainable investments from three different perspectives: the investor level, the company level and the portfolio level. We provide an in-depth analysis of the motivation, behavior and demographics of socially responsible investors. We further show that the motivation of the management, financial inducement and exogenous influence – e.g. publicity or consumer behavior – can impact the extent to which companies act responsible. Portfolio implications are focused on the financial effects, i.e. risk and return of socially responsible investments. We group the existing literature and empirical findings geographically, as the regulatory framework and government support is likely to shape the investment environment, e.g. Henke (2016) finds that European funds have higher responsibility scores. We discuss the historic development of socially responsible investments and gather empirical results considering the financial performance of socially responsible investment funds over time. It remains an open question, whether responsible investment funds perform better or worse than unrestricted funds as evidence for both directions is provided in the literature. However, we do not find clear evidence that social responsible investments struggle, which – considering the restricted investment universe – is quite remarkable. However, further empirical studies, including more recent data, are necessary to determine the impact of social responsibility on financial performance.

In Chapter 3, we empirically analyze the impact of a regulatory change in European financial markets on information asymmetry, idiosyncratic risk and liquidity. The introduction of MiFID II in January 2018 included extensive changes for financial markets, with the superordinate aim to increase transparency. The most drastic change concerns the market for investment research provided by financial analysts.

We discuss, how the literature defines the role of analysts in financial markets and highlight, how investment research is conducted in the EU prior to MiFID II, and compare the regulatory framework with the US. Using data of 1,281 EU firms and 1,646 US firms, we employ a difference-in-difference approach to determine the effect of MiFID II on a firm's analyst coverage, liquidity, stock trading volume, the bid-ask spread and its idiosyncratic risk. We empirically show that MiFID II had a significant impact on European financial markets as the overall bid-ask spread and the idiosyncratic risk increased. We hypothesize that a financial analyst experiences increased competition since the implementation of MiFID II, which should have a verifiable effect on the bid-ask spread. We provide empirical evidence that an increase in analyst coverage reduces the bid-ask spread of a stock, emphasizing the informational role of financial analysts. By estimating the effect of analyst coverage on the bid-ask spread (while controlling for liquidity), we find evidence that financial analysts affect the level of asymmetric information of a stock. This effect is amplified by MiFID II.

In Chapter 4, we empirically analyze the structure and performance of ICOs. We show that ICOs share many similarities with IPOs and are currently generally performed by young and entrepreneurial companies due to their low direct costs. Gathering data from 175 ICOs, we empirically analyze the indirect costs of an ICO by computing the underpricing during ICOs. We find that the underpricing phenomenon – defined as a positive return at the first trading day of a stock – can be observed during an ICO as well and is even more apparent than during IPOs. There exist several explanatory approaches for the existence of IPO underpricing, of which the majority is centered on information asymmetry between the issuing firm, the underwriting bank and investors. We discuss these hypotheses and their transferability onto the ICO setting. Our results suggest that the high level of ICO underpricing is related to a higher degree of information asymmetry between the issuing firm and investors or between investors, as information about ICOs is scarce oftentimes, with a white paper as the sole source of information. Therefore, investors might demand underpricing as compensation for participating in an ICO with little prior knowledge about the risk and return profile. However, we cannot rule out behavioral explanatory approaches and further research is necessary to fully understand the cause for the high level of underpricing observed during ICOs. We further find that the long-term performance of new cryptocurrencies generally remains below the performance of established cryptocurrencies. This anomaly has been documented in the stock markets as well (see Ritter and Welch, 2002).

In Chapter 5, we analyze statistical properties of cryptocurrency limit order books. First, we motivate the importance of market microstructure as a prerequisite for efficient capital allocation. Next, we discuss the scope of application for cryptocurrencies and their disruptive potential. We gather data from one of the largest cryptocurrency exchanges and reconstruct the limit order books for three different cryptocurrencies (Bitcoin, Ethereum and Bitcoin Cash). We find two distinctive features of the aggregated limit order book volume known as the slope of the order book: The slope varies over time and can have linear, concave or even convex properties. We test the empirical fit for different shapes of the slope of the order book and find that a linear slope is generally a sufficient approximation and justifiable with respect to model simplicity. We further stumble upon an anomaly in the distribution of limit order book volume. We observe volume peaks in the limit order book at specific price levels relative to the best price. This anomaly is detectable across all analyzed limit order books. We hypothesize that this pattern is created by a group of investors, which do not consider the entire granularity of the price grid. This leads to an accumulation of available volume at specific price levels. We coin this the "lazy investor hypothesis" and develop a straight-forward regression model to test the empirical significance of this finding and find the respective estimated regression coefficient to be highly significant at both the bid and the ask side across all limit order books.
Finally, we test, whether there is information stored in the order book slope. Similar empirical work has been done by Næs and Skjeltorp (2006), who study stock limit order books. We investigate three groups of models, using the slope of the order book to explain price changes, trading volume and the correlation between price changes and trading volume. We compute the slope on a daily basis averaged across the bid and ask side of the limit order book and show that the slope of the limit order book can help to explain variation in the dependent variables. Our findings also indicate that limit orders placed far away from the best price still appear to be relevant for the price formation process. This finding is interesting, as there is no plausible mechanical link explaining this relationship.
In Chapter 6, we focus on the limit order placement behavior of cryptocurrency traders. There have been some seemingly universal statistical laws discovered in the stock market literature that describe the placement of limit orders. Each trader has to decide at which price level he or she issues a limit order. Each trader faces a trade-off regarding the execution probability of a limit order, time-priority and the threat of having the limit order executed at an unfavorable price. The decision on the price level further depends on the individual trading strategy, liquidity preferences and other factors, making limit order placement a complex process. This makes the emergence of patterns in the arrival of limit orders even more surprising. Zovko and Farmer (2002) and Bouchaud, Mézard, and Potters (2002) show that the probability of an incoming limit order can be described by a *universal* power-law in stock markets. We document that this power-law can be observed in cryptocurrency markets

as well. Moreover, we find empirical evidence that cryptocurrency traders seem to prefer certain price levels and more limit orders are issued at price levels exactly 1.00, 2.00, ... , 5.00 USD away from the best price. We hypothesize that the probability of an incoming limit order is increased, when the distance towards the best price divided by 100 is an integer, indicating that traders apply heuristics when placing limit orders. We extend the power-law model to account for this occurrence and show that our extended model better fits the empirical data than the plain model.
In Chapter 7 we analyze the impact of a radical change of the macroeconomic environment on the market for cryptocurrencies. The outbreak of the coronavirus in 2019/2020 had a sudden impact on all industries and caused unprecedenced changes to the market environment across all asset classes around the globe. In this chapter we focus on the impact on the cryptocurrency market and the potential of cryptocurrencies as a save investment haven in times of a crisis. We find that cryptocurrencies with a high market capitalization seem to be affected by the overall depreciation at the beginning of the outbreak and – in line with Corbet et al. (2020) – their suitability as a save investment haven is therefore questioned. However, smaller cryptocurrencies seem to be affected less by the virus outbreak and might be more appropriate for a crisis-resistant "save haven" portfolio. However, the full extent of COVID-19 and its impact on cryptocurrency markets is not yet foreseeable and further research needs to be conducted to understand the ways in which a virus outbreak affects this emerging asset class.
After completion of this thesis, our findings concerning the six broader research questions derived in Chapter 1 can be summarized as follows:

- **Research Question 1**: Is the concept of market efficiency compatible with social responsibility?

  Result: Market efficiency is compatible with social responsibility mainly due to non-monetary benefits (e.g. consumer reactions, personal values or publicity). Further, empirically evidence suggests that funds engaging in socially responsible investments do not perform significantly worse than unrestricted funds.

- **Research Question 2**: How do participants in highly efficient stock markets react to an exogenous shock in market design (regulatory change)?

  Result: The net positive effect of regulation on financial markets remains an open question and side effects of a regulatory change are tough to predict. In our empirical analysis however, we specifically focus on MiFID II and show that this regulatory change led to an increase in the average bid-ask spread and idiosyncratic risk across EU stock markets. We also find that the effect of

analyst coverage on the bid-ask spread is enhanced since the implementation of MiFID II.

- **Research Question 3**: How do investors cope with information asymmetry?

  Result: While investors do not have the capabilities to obtain all price-relevant information, they do not seem to be deterred from information asymmetry. They rather seek compensation for engaging in uncertain and thereby risky investments. We show that ICOs – which are typically surrounded by a high level of information asymmetry between investors and the issuing company – exhibit high initial returns at the first trading day. We compare our results to the IPO process and find that ICO underpricing is higher than IPO underpricing, supporting the hypothesis that the underpricing is caused by information asymmetry and can be regarded as a compensation for investors.

- **Research Question 4**: How does the market microstructure shape market dynamics?

  Result: We group market dynamics into three distinct categories: Price volatility, trading volume and the volume-volatility relation. We try to connect the market microstructure and the market development by computing daily averages of certain microstructure features. Analyzing the market for cryptocurrencies, we find that the slope of the order book – representing the aggregated limit order book volume – has an impact on all three of the categories. The market microstructure seems to significantly impact the broader market development and should be considered in asset pricing models. Traders can observe the state of the limit order book at any given point in time and seem to consider its shape during their investment decision.

- **Research Question 5**: How do traders behave when placing limit buy and sell orders in a competitive market setting?

  Result: We find that limit order placement is a complex task and the solution is determined by individual time and price preferences and the respective trading strategy. We find empirical evidence suggesting that traders resort to heuristics when determining limit order prices. By disregarding the granularity of the price grid, traders reduce the complexity and find a solution more easily. This behavioral regularity has not been documented in the literature before and we provide an extended model that takes this finding into account.

Our empirical results show that our theoretical model can explain the distribution of incoming limit order prices better than existing models.

- **Research Question 6**: How does the cryptocurrency market react to an exogenous shock (market crisis)?

  Result: We find that smaller cryptocurrencies exhibit return patterns that separate them from cryptocurrencies with a high market capitalization. Trading volume seems to be a predictor for returns of small cryptocurrencies. This relation is enhanced since the exogenous shock caused by COVID-19. Further, the interaction term between return and trading volume has a significant negative effect on future returns of large cryptocurrencies. While all cryptocurrencies suffered from the initial macroeconomic shock caused by the COVID-19 pandemic, the market for smaller cryptocurrencies seems to be more resilient to exogenous shocks.

In conclusion, this dissertation provides an in-depth analysis of the complex interplay between financial market agents in different market settings. We find that information asymmetry plays a crucial role across markets, time periods and geographical regions and provide empirical evidence about the extent to which information disparity can affect asset prices and liquidity. Throughout this dissertation, we show that investors act responsible despite no direct monetary incentive (Chapter 2), rely on financial analysts to obtain price-relevant information and are affected by regulatory changes (Chapter 3), demand a premium when not having full information during an investment decision (Chapter 4), use the limit order book as a source of information (Chapter 5) and resort to heuristics when solving complex tasks (Chapter 6). Chapter 7 finally showcases, how sensitive the market equilibrium can be to a sudden change in the market setting and discusses the adjustments made by investors.

Bibliography

Akaike, Hirotugu (1974). "A new look at the statistical model identification". In: *IEEE transactions on automatic control* 19.6, pp. 716–723.

AMF (2018). *MiFID II: Impact of the new tick size regime*. URL: https://www.amf-france.org/en_US/Publications/Lettres-et-cahiers/Risques-et-tendances/Archives?docId=workspace%5C%3A%5C%2F%5C%2FSpacesStore%5C%2F4ee6cbf6-c425-4537-ab74-ef249b9d316d (visited on 01/15/2019).

Amihud, Yakov (2002). "Illiquidity and stock returns: cross-section and time-series effects". In: *Journal of Financial Markets* 5.1, pp. 31–56.

Barber, Brad, Reuven Lehavy, Maureen McNichols, and Brett Trueman (2001). "Can investors profit from the prophets? Security analyst recommendations and stock returns". In: *The Journal of Finance* 56.2, pp. 531–563.

Bessler, Wolfgang and Matthias Stanzel (2007). "Qualität und Effizienz der Gewinnprognosen von Analysten". In: *Kredit und Kapital* 40.1, pp. 89–129.

Biais, Bruno, Pierre Hillion, and Chester Spatt (1995). "An empirical analysis of the limit order book and the order flow in the Paris Bourse". In: *The Journal of Finance* 50.5, pp. 1655–1689.

Bouchaud, Jean-Philippe, Marc Mézard, and Marc Potters (2002). "Statistical properties of stock order books: empirical results and models". In: *Quantitative Finance* 2.4, pp. 251–256.

Casey, Jean-Pierre and Karel Lannoo (2009). *The MiFID Revolution*. Cambridge University Press.

Chen, Qi, Itay Goldstein, and Wei Jiang (2006). "Price informativeness and investment sensitivity to stock price". In: *The Review of Financial Studies* 20.3, pp. 619–650.

Chiarella, Carl, Giulia Iori, and Josep Perelló (2009). "The impact of heterogeneous trading rules on the limit order book and order flows". In: *Journal of Economic Dynamics and Control* 33.3, pp. 525–537.

Chung, Kee H., Thomas H. McInish, Robert A. Wood, and Donald J. Wyhowski (1995). "Production of information, information asymmetry, and the bid-ask spread: Empirical evidence from analysts' forecasts". In: *Journal of Banking & Finance* 19.6, pp. 1025–1046.

Citigate Dewe Rogerson (2019). *11th Annual IR Survey*. URL: https://citigate%5C-dewe%5C-rogerson.com/uk-companies-among-hardest-hit-decline-

analyst-coverage-according-citigate-dewe-rogersons-11th-annual-ir-survey/ (visited on 10/28/2019).

Cont, Rama (2001). "Empirical properties of asset returns: stylized facts and statistical issues". In: *Quantitative Finance* 1.2, pp. 223–236.

Cont, Rama, Sasha Stoikov, and Rishi Talreja (2010). "A stochastic model for order book dynamics". In: *Operations Research* 58.3, pp. 549–563.

Copeland, Thomas E. and Dan Galai (1983). "Information effects on the bid-ask spread". In: *The Journal of Finance* 38.5, pp. 1457–1469.

Corbet, Shaen, Charles Larkin, and Brian Lucey (2020). "The contagion effects of the covid-19 pandemic: Evidence from gold and cryptocurrencies". In: *Finance Research Letters* 35, p. 101554.

Corwin, Shane A. (2003). "The determinants of underpricing for seasoned equity offers". In: *The Journal of Finance* 58.5, pp. 2249–2279.

De Bondt, Werner F. M. and Richard H. Thaler (1990). "Do security analysts overreact?" In: *The American Economic Review*, pp. 52–57.

De Vries, Alex (2018). "Bitcoin's growing energy problem". In: *Joule* 2.5, pp. 801–805.

Deutsche Börse (2017). *Die Auswirkungen von MiFID II auf die Verfügbarkeit von Research*. URL: https://www.deutsche-boerse.com/resource/blob/166524/ab0a3455fe3aab17051126128607e577/data/Market-Trends-MiFID-II-Research.pdf (visited on 11/04/2019).

Deutscher Investor Relations Verband (2017). *Mögliche Auswirkungen von MiFID II für Emittenten*. URL: https://www.dirk.org/dirk_webseite/static/uploads/171006-MV_KS_Praesentation_Bommer_MiFID.pdf (visited on 11/04/2019).

Durnev, Artyom, Randall Morck, Bernard Yeung, and Paul Zarowin (2003). "Does greater firm-specific return variation mean more or less informed stock pricing?" In: *Journal of Accounting Research* 41.5, pp. 797–836.

Easley, David, Nicholas M. Kiefer, Maureen O'Hara, and Joseph B. Paperman (1996). "Liquidity, information, and infrequently traded stocks". In: *The Journal of Finance* 51.4, pp. 1405–1436.

Easley, David, Maureen O'Hara, and Joseph Paperman (1998). "Financial analysts and information-based trade". In: *Journal of Financial Markets* 1.2, pp. 175–201.

ESMA (2018). *Amendment to Commission Delegated Regulation (EU) 2017/588 (RTS 11)*. URL: https://www.esma.europa.eu/press-news/consultations/amendment-commission-delegated-regulation-eu-2017588-rts-11#TODO (visited on 01/15/2019).

European Commission (2014). "Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on Market Financial Instruments and Amending Directive 2002/92/EC and Directive 2011/61/EU". In: *Official Journal of the European Union* 173, pp. 349–496.

Fama, Eugene F. (1970). "Efficient capital markets: A review of theory and empirical work". In: *The Journal of Finance* 25.2, pp. 383–417.

Fang, Bingxu, Ole-Kristian Hope, Zhongwei Huang, and Rucsandra Moldovan (2019). "The Effects of MiFID II on Sell-Side Analysts, Buy-Side Analysts, and Firms". URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3422155 (visited on 11/04/2019).

Ferrarini, Guido and Eddy Wymeersch (2006). *Investor protection in Europe: corporate law making, the MiFID and beyond*. Oxford University Press.

Foster, Andrew J. (1995). "Volume-volatility relationships for crude oil futures markets". In: *Journal of Futures Markets* 15.8, pp. 929–951.

Foster, F. Douglas and S. Viswanathan (1995). "Can speculative trading explain the volume–volatility relation?" In: *Journal of Business & Economic Statistics* 13.4, pp. 379–396.

French, Kenneth R. and Richard Roll (1986). "Stock return variances: The arrival of information and the reaction of traders". In: *Journal of Financial Economics* 17.1, pp. 5–26.

Fung, Hung-Gay and Gary A. Patterson (1999). "The dynamic relationship of volatility, volume, and market depth in currency futures markets". In: *Journal of International Financial Markets, Institutions and Money* 9.1, pp. 33–59.

Givoly, Dan and Josef Lakonishok (1979). "The information content of financial analysts' forecasts of earnings: Some evidence on semi-strong inefficiency". In: *Journal of Accounting and Economics* 1.3, pp. 165–185.

Gleason, Cristi A. and Charles M. C. Lee (2003). "Analyst forecast revisions and market price discovery". In: *The Accounting Review* 78.1, pp. 193–225.

Glosten, Lawrence R. (1987). "Components of the bid-ask spread and the statistical properties of transaction prices". In: *The Journal of Finance* 42.5, pp. 1293–1307.

Glosten, Lawrence R. and Lawrence E. Harris (1988). "Estimating the components of the bid/ask spread". In: *Journal of Financial Economics* 21.1, pp. 123–142.

Glosten, Lawrence R. and Paul R. Milgrom (1985). "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders". In: *Journal of Financial Economics* 14.1, pp. 71–100.

Grossman, Sanford J. and Joseph E. Stiglitz (1980). "On the impossibility of informationally efficient markets". In: *The American Economic Review* 70.3, pp. 393–408.

Henke, Hans-Martin (2016). "The effect of social screening on bond mutual fund performance". In: *Journal of Banking & Finance* 67, pp. 69–84.

Holste, Bjoern and Christoph Gallus (2019). *Sind Krypto-Währungsmärkte Fair?(Are Crypto-Currency Markets Fair?)* Working Paper.

Johnsen, Edward J. and John Grady (2017). *SEC provides relief to US firms attempting to comply with EU MiFID II's research "unbundling" provisions*. URL: https://www.dlapiper.com/en/us/insights/publications/2017/10/sec-relief-to-us-firms-attempting-to-comply/ (visited on 01/10/2019).

Karpoff, Jonathan M. (1987). "The relation between price changes and trading volume: A survey". In: *Journal of Financial and Quantitative Analysis* 22.1, pp. 109–126.

Kelly, Bryan and Alexander Ljungqvist (2012). "Testing asymmetric-information asset pricing models". In: *The Review of Financial Studies* 25.5, pp. 1366–1413.

Kettle, Keri and Gerald Häubl (2010). "Numeric fluency and preference". In: *NA – North American Advances* 37, pp. 150–152.

Kim, Oliver and Robert E. Verrecchia (1994). "Market liquidity and volume around earnings announcements". In: *Journal of Accounting and Economics* 17.1-2, pp. 41–67.

Lippiatt, Todd and Michael Oved (2018). *The Two Token Waterfall*. URL: `https://framework.factora.io/#/` (visited on 02/07/2019).

Mantegna, Rosario N. and H. Eugene Stanley (1999). *Introduction to econophysics: correlations and complexity in finance*. Cambridge University Press.

Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti (2019). *Bitcoin liquidity*. Working Paper. Massey University and Auckland University of Technology.

Mike, Szabolcs and J. Doyne Farmer (2008). "An empirical behavioral model of liquidity and volatility". In: *Journal of Economic Dynamics and Control* 32.1, pp. 200–234.

Morck, Randall, Bernard Yeung, and Wayne Yu (2000). "The information content of stock markets: why do emerging markets have synchronous stock price movements?" In: *Journal of Financial Economics* 58.1-2, pp. 215–260.

Næs, Randi and Johannes A Skjeltorp (2006). "Order book characteristics and the volume–volatility relation: Empirical evidence from a limit order market". In: *Journal of Financial Markets* 9.4, pp. 408–432.

Nakamoto, Satoshi (2008). *Bitcoin: A peer-to-peer electronic cash system*. Working Paper.

Nofer, Michael, Peter Gomber, Oliver Hinz, and Dirk Schiereck (2017). "Blockchain". In: *Business & Information Systems Engineering* 59.3, pp. 183–187.

O'Hara, Maureen (1997). *Market microstructure theory*. Wiley.

O'Hara, Maureen (2015). "High frequency market microstructure". In: *Journal of Financial Economics* 116.2, pp. 257–270.

Pareto, Vilfredo (1964). *Cours d'économie politique*. Vol. 1. Librairie Droz.

Piotroski, Joseph D. and Darren T. Roulstone (2004). "The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices". In: *The Accounting Review* 79.4, pp. 1119–1151.

Potters, Marc and Jean-Philippe Bouchaud (2003). "More statistical properties of order books and price impact". In: *Physica A: Statistical Mechanics and its Applications* 324.1-2, pp. 133–140.

Preece, Rhodri (2019). *MiFID II: One Year On*. URL: https://www.cfainstitute.org/-/media/documents/survey/cfa-mifid-II-survey-report.ashx (visited on 10/29/2019).

Preis, Tobias, Sebastian Golke, Wolfgang Paul, and Johannes J Schneider (2006). "Multi-agent-based order book model of financial markets". In: *EPL (Europhysics Letters)* 75.3, pp. 510–516.

Ritter, Jay R. (1987). "The costs of going public". In: *Journal of Financial Economics* 19.2, pp. 269–281.

Ritter, Jay R. and Ivo Welch (2002). "A review of IPO activity, pricing, and allocations". In: *The Journal of Finance* 57.4, pp. 1795–1828.

Roll, Richard (1988). "R2". In: *The Journal of Finance* 43.3, pp. 541–566.

Sarwar, Ghulam (2003). "The interrelation of price volatility and trading volume of currency options". In: *Journal of Futures Markets: Futures, Options, and Other Derivative Products* 23.7, pp. 681–700.

SEC (2017). *Response of the chief counsel's office division of investment management*. URL: https://www.sec.gov/divisions/investment/noaction/2017/sifma-102617-202a.htm (visited on 01/14/2019).

Stoll, Hans R. (1989). "Inferring the components of the bid-ask spread: Theory and empirical tests". In: *The Journal of Finance* 44.1, pp. 115–134.

Thomsen, Steen and Frederik Vinten (2014). "Delistings and the costs of governance: a study of European stock exchanges 1996–2004". In: *Journal of Management & Governance* 18.3, pp. 793–833.

Wallmeier, Martin (2005). "Analysts' earnings forecasts for DAX100 firms during the stock market boom of the 1990s". In: *Financial Markets and Portfolio Management* 19.2, pp. 131–151.

Wolfson, Rachel (2018). *A first for Manhattan: $30M Real Estate Property Tokenized With Blockchain*. URL: https://www.forbes.com/sites/rachelwolfson/2018/10/03/a-first-for-manhattan-30m-real-estate-property-tokenized-with-blockchain/#1910a2d48957 (visited on 02/07/2019).

Womack, Kent L. (1996). "Do brokerage analysts' recommendations have investment value?" In: *The Journal of Finance* 51.1, pp. 137–167.

Zhu, PengCheng, Vijay Jog, and Isaac Otchere (2014). "Idiosyncratic volatility and mergers and acquisitions in emerging markets". In: *Emerging Markets Review* 19, pp. 18–48.

Zovko, Ilija and J. Doyne Farmer (2002). "The power of patience: a behavioural regularity in limit-order placement". In: *Quantitative Finance* 2.5, pp. 387–392.