



Master Thesis Economics and Business - Financial Economics

# Underpricing in Initial Coin Offerings: An Empirical Study on Ex-ante Uncertainty

Máté Czumbel 476902mc

Supervisor: Haikun Zhu

Date final version: 21-04-2020

## Abstract:

This thesis aims to establish the determinants of underpricing in initial coin offerings (ICO). It is hypothesized that the lower the uncertainty about the future value of the venture before the ICO (ex-ante uncertainty), the lower the underpricing. In a sample of 582 exchange listed ICOs, I show that underpricing is significant and positive. Furthermore, I find that information conveyance proxies are only partially significant determinants of underpricing, whereas proxies related to quality signaling efforts of the token issuer are significantly related to underpricing. This thesis contributes to existing research in the field of cryptocurrencies, by empirically studying the relationship between uncertainty and underpricing in ICOs.

**Keywords:** *initial coin offering, underpricing, ex-ante uncertainty, information asymmetry, cryptocurrency, fundraising, token sale, initial returns*

The views stated in this thesis are those of the author and not necessarily those of Erasmus School of Economics or Erasmus University Rotterdam.

# Table of Contents

1	Introduction .....	2
2	Literature Review .....	3
2.1	Theoretical Background .....	3
2.1.1	Cryptocurrencies and Blockchain.....	3
2.1.2	Initial Coin Offerings .....	4
2.1.3	The Underpricing Phenomenon.....	6
2.1.4	Uncertainty in Initial Coin Offerings.....	7
2.1.5	Underpricing in Initial Coin Offerings.....	9
2.2	Hypotheses .....	11
3	Methodology .....	13
3.1	Measure of Underpricing.....	13
3.2	Existence of Underpricing.....	14
3.3	Regression models.....	14
3.4	Independent Variables of Interest.....	15
3.5	Control Variables .....	16
4	Data .....	17
4.1	Data Collection and Processing.....	17
4.2	Summary Statistics .....	18
5	Results .....	20
5.1	Underpricing in ICOs .....	20
5.2	Information Conveyance .....	21
5.3	Quality Signaling.....	23
5.4	Multicollinearity.....	25
6	Discussion and Limitations .....	25
6.1	General Discussion.....	25
6.2	Limitations and Future Research.....	28
6.3	Implications of Results .....	27
7	Conclusion.....	29
8	References .....	30
	Appendix A .....	34
	Appendix B .....	35
	Appendix C .....	36
	Appendix D .....	37

# 1 Introduction

One of the most novel fundraising methods for blockchain-based startups is an initial coin offering (ICO), where the underlying cryptocurrency is offered to the general public for the first time. As in the case of initial public offerings of stocks (IPO), ICOs can be subject to an inefficiency known as underpricing, where the average offer price in the public sale is lower than the market trading price. Underpricing is a common phenomenon in IPOs and there is some evidence that ICOs are also characterized by positive underpricing, however conclusive empirical findings on the determinants of underpricing in ICOs are scarce.

IPO literature on underpricing focuses on the theory that high ex-ante uncertainty about the future value of the company results in a lower offer price, which is then corrected in the first day of trading. According to uncertainty based underpricing theories, the higher the uncertainty, the higher the underpricing. ICOs are characterized by asymmetric information between investors and issuers, therefore there is inherent ex-ante uncertainty regarding the future value of the venture. This thesis aims to answer the following research question:

*Can ex-ante uncertainty explain underpricing in Initial Coin Offerings?*

This thesis contributes to existing studies by examining the empirical relationship between proxies for ex-ante uncertainty and underpricing in ICOs. Underpricing represents an inefficiency that affects both the issuer and the investors, therefore the importance of this thesis is threefold. Firstly, finding the determinants of underpricing in ICOs can allow issuers to reduce underpricing and maximize the efficiency of fundraising. Secondly, investors may be able to more effectively identify mispriced issues by observing signals of issuers. Thirdly, understanding the determinants of ICO underpricing can provide important insight for regulators on the operation of the ICO markets and provide guidance on forming future cryptocurrency legislation.

The thesis consists of the following sections. Section 2 presents the theoretical background and empirical findings of previous literature on cryptocurrencies, initial coin offerings and underpricing, and presents the hypotheses. Section 3 and 4 discuss the methodology and the data sample used in the analyses. Section 5 presents the results of the analyses, whereas Section 6 discusses the findings, limitations and implications of the results. Section 7 concludes.

## 2 Literature Review

### 2.1 Theoretical Background

#### 2.1.1 Cryptocurrencies and Blockchain

Cryptocurrencies are one of the most innovative fintech developments of recent years, sparking interest in several fields of research, including economics. Bitcoin, the first completely decentralized digital currency was created following the 2008 global financial crisis, when the instability of the existing financial system was highlighted by the need for government bailouts. The value of bitcoin has skyrocketed in the past decade, amidst concerns about excessive volatility, reaching an all-time high of around 20,000 USD on December 17, 2017.<sup>1</sup> The spreading popularity of bitcoin jump-started the development of alternative cryptocurrencies and crypto tokens, so much so that as of January 2020, there are 2,388 coins listed on coinmarketcap.com (CMC).<sup>2</sup>

The main contribution of bitcoin to the cryptocurrency world is the blockchain concept. Blockchain is defined as a distributed, digital ledger that allows the verification of transactions in a peer-to-peer network (Yermack, 2017). The main advantage of a distributed network is that there is no need for a central authority to verify transactions. This is possible due to the chronological chain of all previous block of transactions (i.e. the blockchain), which is shared and stored by all system users. The shared ledger of transactions means that altering previous records is not possible without breaking the blockchain, because at any given moment the history is stored on all nodes. The reconciliation of entries is carried out by participants of the network, who are rewarded for their computing power in a process known as mining.<sup>3</sup> Blockchain technology has facilitated the advancement of numerous promising projects and has already been incorporated into various applications within payment systems, trade and logistics (Cong & He, 2019). Monetary authorities, such as the Dutch Central Bank, closely monitor crypto developments and are actively involved in their own cryptocurrency technological innovations enabled by blockchain technology in order to gain understanding of the possibilities and future potential (De Nederlandsche Bank, 2018).

---

<sup>1</sup> Price data retrieved from coinmarketcap.com.

<sup>2</sup> Coinmarketcap.com is a price data aggregator website and it is considered to be a reliable source of market data for cryptocurrencies (Benedetti & Kostovetsky, 2018).

<sup>3</sup> The in-depth technical explanation of the blockchain technology is outside of the focus of this paper, more details are available in the original Bitcoin whitepaper: <https://bitcoin.org/bitcoin.pdf>.

A promising application of the blockchain technology is the concept of smart contracts. A smart contract is a protocol that executes payment contracts automatically and digitally, as soon as the terms of the contract are met (Cong & He, 2019). The advancement of blockchain has given a new push to the implementation of smart contracts in the form of innovative projects such as Ethereum. Ethereum is a distributed platform based on blockchain that offers a flexible interface for implementing smart contracts. Ethereum raised \$16 million in the initial coin offering (ICO) of its cryptocurrency in 2014. Ether has since become the second largest cryptocurrency by market capitalization after bitcoin.<sup>4</sup>

The expansion of Ethereum gave rise to a large wave of ICOs by providing an adaptable platform for the implementation of smart contracts. ICOs are cryptographic token sales for startups or projects within the blockchain ecosystem, where investors are offered tokens of the venture for the first time in exchange for funds. Since not all tokens are classified as currencies, initial token sale is a more complete term than initial coin offering. Nevertheless, for simplicity, I will continue to use these terms interchangeably in this paper, as the term ICO is used more commonly in the cryptocurrency world. Although not all ICOs are based on the Ethereum blockchain, Ethereum brought about the possibility of token sales without the need to develop a standalone blockchain protocol. The first ICO was held in 2013 and there have been more than 6,000 offerings globally since then, attracting considerable interest from individual and institutional investors alike (Howell, Niessner, & Yermack, 2019).

### 2.1.2 Initial Coin Offerings

ICOs represent an innovative way of early-stage fundraising for start-ups active in the blockchain ecosystem (Adhami, Giudici, & Martinazzi, 2018). Ventures can raise funds via a number of different ways. Traditional ways of equity financing include initial public offerings (IPO) and venture capital (VC), whereas crowdfunding is a way of raising funds without giving up ownership or control of the business. The academic literature predominantly places ICOs between VCs and crowdfunding in terms of characteristics (Chod & Lyandres, 2018). ICOs, similarly to VCs and crowdfunding are held by ventures in the early stages of product development, often without a working prototype and aim to raise money from a large number of investors. Unlike in equity financing, investors in ICOs in most cases do not receive a share of equity in the venture. The

---

<sup>4</sup> Market capitalization retrieved from [coinmarketcap.com](https://coinmarketcap.com), amount raised retrieved from [icodrops.com](https://icodrops.com).

cryptographic tokens received in return for the investment give access to the future service or product of the venture that will eventually be developed, but do not give rights to future profits of the venture.<sup>5</sup> This characteristic of ICOs means that investors take on high risks for limited compensation.

Initial coin offerings, similarly to other ways of fundraising, have their respective advantages and disadvantages. According to Kaal (2018), low entry barriers and the borderless nature of ICOs enable it to be one of the most efficient ways of fundraising. Howell et al. (2019) outline the advantages of ICOs, focusing on lower transaction costs and improved liquidity. The relatively simple fundraising process, the low transaction costs along with the high liquidity due to continuous trading mean that ICOs are very cost effective, especially considering the fact that no control has to be given up of the venture. On the other hand, ICOs also have their respective disadvantages, especially from the point of view of the prospective investors (Kaal, 2018). Since the investors do not receive equity in the venture in exchange for their investment, they have no control of the future strategy of the firm. The investors in most cases invest only in a promise to have access to a future product or service. Since in most countries, ICOs are not required to be registered with the local securities authority, there is limited transparency and a general lack of regulation regarding disclosure requirements. As such, there is very high uncertainty about the future performance and profitability of the venture, since there is rarely any business history available.

Commonly, there are three different token types: currency, utility and security tokens. Currency tokens are a medium of exchange and can be used for payment, whereas utility tokens allow investors to use the services of the start-up after the eventual project launch. A security token is essentially a token of ownership in the venture that entitles the owner to claims on profits, similarly to traditional securities (Zetsche, Buckley, Arner, & Föhr, 2017). For this reason, ICOs are often compared to IPOs in literature. Thijn (2018) provides a comprehensive comparison between IPOs and ICOs. Important similarities include, raising funds in a public sale from a large number of investors, eventual exchange listing and the phases of the fundraising process. Thijn (2018) also points out several contrasting differences between the IPO and the ICO processes. The strict regulatory environment and complicated procedure makes IPOs lengthy and expensive to set up,

---

<sup>5</sup> Adhami et al. (2018) describe an example of an ICO in detail, the MobileGo project.

whereas ICOs are easy and cheap to launch as regulation is weak, and there is no need for underwriters.

The mixed nature of ICOs raises the dilemma whether tokens sold in an ICO should be treated as a security or not. Lyandres, Palazzo, and Rabetti (2018) attempt to respond to this question by analyzing ICO characteristics and post-ICO performance. The findings of this research give empirical evidence that determinants of ICO success are consistent with that of the IPO literature and the theories that explain IPO characteristics are partially applicable to ICOs. The paper concludes that in essence, tokens sold as part of an ICO behave like securities and most differences are attributable to a difference in the institutional set up between cryptocurrency and traditional equity markets. Howell et al. (2019) establish that prior to the launch of an ICO, the seller faces similar design choices as an issuer of an IPO. As opposed to Lyandres et al. (2018), Howell et al. (2019) cannot conclude that cryptocurrency tokens should be treated as securities. Most tokens are classified as utility tokens and as such, they are not classified as financial assets, but as a product or commodity. Howell et al. (2019) also give an overview of the current regulatory environment, which highlights the inconsistencies regarding the treatment of ICOs.

### 2.1.3 The Underpricing Phenomenon

An important topic of discussion regarding exchange listed issues is underpricing. According to Ibbotson (1975), an issue is underpriced, if the offer price of the stock in the IPO is lower than the closing price of the first trading day. This represents an unfavorable situation, as the stock was offered below market price, which results in an inefficiency for the issuer. The literature exploring the determinants of underpricing in ICOs is limited. There are only a handful of research articles that study underpricing in ICOs, however few of them explore empirically the factors that are related to underpricing and even those that do, are restricted to a small dataset, therefore it is useful to consult the IPO literature on this topic. Underpricing in IPOs is a common occurrence and has been subject to thorough research, resulting in several theories that attempt to explain the reason behind positive first-day returns (Baron, 1982; Beatty & Ritter, 1986; Rock, 1986). Information asymmetry and ex-ante uncertainty are related terms in the explanations behind underpricing in IPOs. Uncertainty about the future value of a company after the issue is referred to as ex-ante uncertainty. This type of uncertainty commonly stems from asymmetric information between different parties of a transaction. Asymmetric information exists when one party in a transaction

has superior information either in terms of quantity or quality than others. According to Akerlof (1970), market participants have incentives to try to sell low-quality goods as high-quality goods, by exploiting uninformed buyers who possess less information, therefore asymmetric information creates uncertainty about the true quality of a good. Löfgren, Persson, and Weibull (2002) argue that buyers are interested in reducing uncertainty about the true quality of the good by acquiring more information and that sellers also aim to reduce uncertainty about their goods by signaling their quality.

In Rock's (1986) model of underpricing, riskier issues are more underpriced. The study of Beatty and Ritter (1986) finds evidence that underpricing in IPOs is positively related to ex-ante uncertainty, by constructing a model with two variables of interest that act as proxies for uncertainty about the values of firms prior to the issue. In this model, IPO issuers need to lower the offer price in order to be able to attract investors and fight against uncertainty. As a result, issuers can effectively reduce the level of underpricing by signaling quality to investors to reduce uncertainty. Welch (1989) argues that the true value of the issue can be effectively signaled by the issuers of IPOs by underpricing the issue, which forces out the low-quality issuers. In a similar way, the offer price and the percentage of retained stock is a way for issuers to signal true firm value (Grinblatt & Hwang, 1989). In the theoretical model of Grinblatt and Hwang (1989), IPO issuers are able to generate a much larger interest in their issue by underpricing it below its true value and thus eventually reaching a higher market price than without initial underpricing. Similarly to the offer price, retention is also identified in the model as way of signaling quality to investors, since by withholding a larger fraction of shares, issuers signal their expectations of larger future cashflows.

#### 2.1.4 Uncertainty in Initial Coin Offerings

Information asymmetry is arguably a major source of uncertainty among ICO investors. Howell et al. (2019) discuss the parallels between ICOs and other popular forms of raising capital and provide a thorough explanation of the ICO process, discussing advantages, importance and regulation. Empirically, they show that reducing information asymmetry, such as through the existence of a whitepaper is associated with liquidity and trading volume of ICOs and to some extent success is associated with social media engagement. Zetzsche et al. (2017) underline information asymmetry as a key problem regarding ICOs. They identify whitepapers as an important source of information

for investors, however there is varying level of information available in whitepapers of different ventures. The survey-based findings of Yadav (2017) reiterate this finding, as cryptocurrency investors actively look to whitepapers when making investment decisions. Purchasing tokens in an ICO is a high-risk investment with a lot of uncertainty, due to the unregulated nature, the lack of investor protection and the relatively large number of questionable projects or scams they represent (Zetzsche et al., 2017).

Boreiko and Sahdev (2018), Davydiuk, Gupta, and Rosen (2019), Chod and Lyandres (2018) all identify possible ways that issuers can effectively reduce information asymmetry regarding the quality of their ventures. Boreiko and Sahdev (2018) show that successful ICOs are differentiated by signaling their quality to investors more effectively, which allows them to reduce uncertainty. By classifying ICOs into groups, they find that ICOs that reach their hardcap are more likely to have a whitepaper, to be selling an Ethereum based token and have previous venture capital support. Furthermore, they establish that successful ICOs have a significantly higher number of twitter followers and a higher rating on the ICO aggregator website icobench.com. Chod and Lyandres (2018) develop a theoretical model, in which they predict that high quality ICOs are only present in the marketplace, if there is a way for them to effectively reduce information asymmetry and differentiate themselves from low quality ICOs. Similarly to other literature, they also identify whitepapers as the most important information source. Although the theoretical model developed by Chod and Lyandres (2018) can provide a useful base for establishing the role of information asymmetry in ICOs, it is not supported by empirical data. In line with the empirical findings of Boreiko and Sahdev (2018) and the theoretical predictions of Chod and Lyandres (2018), Davydiuk et al. (2019) show that ICO success is strongly related to quality signaling. There is empirical evidence that various measures of success, such as exchange listing or total amount raised are significantly predicted by token retention and that higher proportion of token retention lowers asymmetric information.

The above findings show that it is difficult to accurately value the venture at the time of an ICO and that information asymmetry is the principal concern. Ventures launching ICOs are relatively new with no proven track record, share limited information and are not under the same regulatory scrutiny as ventures engaged in traditional equity financing. Weak regulation and varying quality of information create uncertainty around ICOs, which is one of the of the biggest concerns

regarding this type of financing. Consequently, issuers of tokens in ICOs should aim reduce asymmetry by communicating information about their ventures or signal their quality to reduce uncertainty regarding the future profitability of the venture.

#### 2.1.5 Underpricing in Initial Coin Offerings

The collection of studies that empirically investigate underpricing in ICOs establish varying levels of positive underpricing. Adhami et al. (2018) provide a detailed explanation of ICOs and using a sample of 253 offerings, empirically analyze the probability of token sale success based on certain variables. The authors show that the success of ICOs is more probable, when investors can access the source code and when issuers sell tokens pre-issue. The study also concludes that the mean level of underpricing in the sample used is 929.9%.

Benedetti and Kostovetsky (2018) also find evidence for heavy underpricing at an average level of 173% in their dataset of 416 ICOs of exchange listed tokens. In their extended sample of 887 ICOs that includes tokens that were never listed on an exchange (assumed return of -50%), the average value weighted return is 82%. Furthermore, they find that underpricing is reduced when the initial offer price is higher and when there is a Pre-ICO before the main offering. In the view of the authors, this is due to the fact that a low ICO price attracts higher demand from investors, which results in higher performance directly after the ICO. This finding is in line with the theoretical model based on IPO underpricing of Grinblatt and Hwang (1989). Benedetti and Kostovetsky (2018) also study the relationship between underpricing and information asymmetry, for which they use a Twitter social media variable as a proxy for the age of the venture, but fail to find an empirical relationship.

Lyandres et al. (2018) find 269% average underpricing defined as the log difference between the first day opening price and the ICO offer price. As opposed to Benedetti and Kostovetsky (2018), Lyandres et al. (2018) find significant negative relationship between social media variables and underpricing. They also establish that underpricing and long-term returns are strongly related in a negative way, meaning that positive initial returns reverse in the long-term providing evidence of the winner's curse theory.

Chanson, Gjoen, Risius, and Wortman (2018) examine the relationship between underpricing in ICOs and a collection of social media variables. Similarly to Benedetti and Kostovetsky (2018), Chanson et al. (2018) also fail to find evidence that social media variables, such as Twitter

followers and tweets or the age of the firm affect ICO underpricing. On the other hand, they show that if a token sale is oversubscribed i.e. the hardcap has been reached, underpricing is higher.

Momtaz (2019) finds positive underpricing albeit at a much lower level than other papers at an average of 15%. Momtaz (2019) also argues that as 40% of ICOs are overpriced, median underpricing can show a more accurate picture of the level of inefficiency at 3%. The paper also finds empirical evidence that there is a size effect related to ICO underpricing. While underpricing is small, but positive, aggregate money left on the table is negative in half of the years in the sample, meaning that large ICOs, as measured by market capitalization destroy value. An additional finding is that liquidity is positively related to underpricing, which can either be explained by market sentiment or that the issue is underpriced to create liquidity. Momtaz (2018) finds an average underpricing of around 8% and examines the effect of project quality and controls on underpricing. It is shown that team quality, market sentiment and the application of ERC20 standard are all significant determinants of underpricing in ICOs.

Similarly to Lyandres et al. (2018) and Howell et al. (2019), Thijn (2018) adds to the literature that draws parallels between ICOs and IPOs. Thijn (2018) provides a comparison between ICOs and IPOs and establishes that variables used in IPO literature to explain underpricing are also relevant for ICOs. In particular, the author establishes that information asymmetry theories developed for IPOs can be applied for ICOs through the use of proxy variables. Using a sample of 279 token sales that is underpriced at an average level of 102%, it is concluded that underpricing is significantly and negatively related to issue size, as measured by the amount raised in the issue. Furthermore, underpricing is positively related to market sentiment and to the trading volume on the listing day, which is an indication of the liquidity of the given ICO. These findings are in line with the findings of Momtaz (2018) and Momtaz (2019) who also establish a similar relationship, albeit with a different measure of issue size. On the other hand, Thijn (2018) does not find a significant relationship between underpricing and retained percentage of tokens, nor the binary variable of whether the fundraising target was reached, a relationship for which Chanson et al. (2018) do find evidence. In contrast with Benedetti and Kostovetsky (2018) who find that offer price and the existence of a pre-ICO are negatively related underpricing, Thijn (2018) only finds empirical support that pre-ICO is a significant determinant of underpricing.

Research examining the effect of regulation on ICO underpricing is inadequate, however IPO literature can provide a basis to understand the association between disclosure requirements of security issuance and underpricing. Shi, Pukthuanthong, and Walker (2012) find evidence from a cross country sample of IPOs that the level of disclosure requirements in securities laws is negatively related to the level of underpricing.

## 2.2 Hypotheses

Based on the theoretical background and the findings of previous literature presented in the previous section, it shows that underpricing in ICOs is a common phenomenon. However, the level of underpricing varies, depending on the definition of underpricing and the adjustment methods used. For this reason, I hypothesize the following:

*H1: ICOs are significantly underpriced on average in the data sample, even when adjusted for the overall performance of the cryptocurrency market.*

Based on the findings of previous literature, the most prominent explanations of underpricing in IPOs is related to ex-ante uncertainty, which is created by asymmetric information between issuers and investors. By reducing information asymmetry between issuers and investors, underpricing is expected to be lower, because the ex-ante uncertainty about the authenticity and the future performance of the project can be reduced. Thus, the following set of hypotheses relate to the level of asymmetric information regarding the venture. As such, underpricing is expected to be negatively related to information conveyance variables, such as whitepaper variables, type of the token protocol and country of legislation. The whitepaper of an ICO is one of the most important sources of information, by which issuers of tokens can reduce information asymmetry. The ERC20 protocol is based on the Ethereum blockchain and certain properties of the sale are predetermined prior to the ICO and cannot be altered by the issuer (Howel et al., 2018). Such a standardized protocol induces trust and conveys relevant information about the reliability of the ICO to the investors, therefore it is expected that using the ERC20 standard as a token protocol in an ICO reduces information asymmetry. Investors in token sales do not typically receive voting rights, therefore they face considerable uncertainty about the future of the venture. Investor's concerns about uncertainty partially stem from the lack of comprehensive regulation of ICOs (Kaal, 2018). Regulation that affects ICOs can reduce this uncertainty by introducing mandatory disclosure requirements, meaning that investors are subject to lower information asymmetry, therefore it is

expected that uncertainty is lower in countries with stricter regulation. In line with the above, I formulate the second set of hypotheses related to information asymmetry as follows:

*H2a: The level of underpricing is negatively related to the information content of the whitepaper.*

*H2b: The level of underpricing is lower, if the token uses an ERC20 token protocol.*

*H2c: The level of underpricing is lower, if the ICO is issued in a country with strict cryptocurrency legislation.*

The literature that examines the underpricing phenomenon in IPOs, identifies signaling the quality of the issue as a possible way to reduce underpricing. This is because by signaling the quality to investors, uncertainty is reduced. The signaling theory present in IPOs can also be applied to ICOs, however empirical evidence is scarce and contradicting. Thus, the following set of hypotheses relate to the ability of ICO issuers to signal their quality to investors. The offer price of an issue is a way suggested in IPO literature to effectively signal the true value of the issue. The ICO price can similarly be set to a level that attracts larger interest from investors. In addition, share retention is also an important signal of quality in IPO literature and similarly it represents a way for ICO issuers to signal quality as well, because by distributing a lower proportion of tokens to investors (or retaining more tokens), issuers communicate that their expectation about the future performance of the token price is positive. The ICO literature thus far has found conflicting empirical results regarding the relationship between underpricing and these two signaling methods, however according to the theoretical models, it is expected that in the case of ICOs, the offer price is negatively and the level of token distribution is positively related to the level of underpricing. According to Welch (1992), sequential IPOs are an efficient way to a price an issue, as the issuer signals its quality in the initial sale by underpricing and by gauging the interest of the investors, it can set a higher price in subsequent offerings. In a similar way, a pre-ICO can be thought of as a gauging mechanism, by which token issuers are able to adjust their offer price to reduce underpricing in the ICO. Based on the above, I hypothesize the following:

*H3a: The level of underpricing is negatively related to the offer price of the ICO.*

*H3b: The level of underpricing is positively related to the proportion of tokens distributed to investors.*

*H3c: The level of underpricing is lower, if there is a pre-ICO token sale.*

To control for other factors that are related to underpricing, I use the following control variables that are well-established in previous literature: trading volume, issue size, sentiment, crowdedness, ico duration, hardcap target reached.

### 3 Methodology

This chapter presents the methodology used to test the hypotheses and answer the research question of this paper. To test the hypotheses, a t-test is used to determine the significance of underpricing and two sets of OLS regression analyses are used to estimate the determinants of underpricing in ICOs. The following sections present the dependent and independent variables, the t-test and the regression models used in the analysis.

#### 3.1 Measure of Underpricing

The dependent variable in this study is the level of underpricing in ICOs. Following the norm in IPO literature (Loughran & Ritter, 2004), raw underpricing ( $UP_{raw}$ ) is defined in this paper as the initial return of the first day closing price ( $P_c$ ) compared to the offer price in the respective ICO ( $P_{ico}$ ):

$$UP_{raw} = \frac{P_c - P_{ico}}{P_{ico}} \quad (1)$$

Although using closing prices of the first trading day is common in IPO literature, other definitions of underpricing are used as well. Schultz and Zaman (1994) calculate underpricing using opening trading prices, therefore to ensure robustness, underpricing is also calculated using the opening price on the first trading day for the purposes of testing Hypothesis 1.

In this paper, raw underpricing is adjusted by the market return to obtain market-adjusted underpricing  $UP_{adj}$ . By adjusting underpricing to the market, it can be ensured that the initial returns on the investment in the ICO are not due to an overall positive performance in the market. In literature, market adjustment is carried out using the return of an alternative investment portfolio or a relevant market index. In the cryptocurrency world, bitcoin price is shown to be correlated with a large portion of other cryptocurrency prices (Ciaian, Rajcaniova, & Kancs, 2018; Yi, Xu, & Wang, 2018). For this reason and its relatively large market capitalization compared to other exchange listed cryptocurrencies, the return on bitcoin is often used as index of overall market performance. An alternative to Bitcoin is the CCI30 Cryptocurrency Index. The CCI30 is a market-

cap weighted index that tracks the performance of the top 30 cryptocurrencies as measured by market capitalization.<sup>6</sup> Although the index was only launched in the beginning of 2017, its value has been calculated retrospectively starting from 2015 January. In this paper, market-adjusted underpricing is calculated by subtracting the market return from raw underpricing, as follows:

$$UP_{adj} = UP_{raw} - R_{market} \quad (2)$$

where  $UP_{adj}$  is the market adjusted level of underpricing,  $R_{market}$  is the daily return of the market.

### 3.2 Existence of Underpricing

To test Hypothesis 1 and establish whether the mean level of ICO underpricing is positive and significantly different than zero, a one sample t-test is used. The t-test of means assumes normal distribution of the level of underpricing in the sample, and due to the large sample size, this assumption is satisfied. For the purposes of Hypothesis 1, underpricing is calculated using both opening and closing prices, and it is adjusted to bitcoin return and the return of the CCI30 Index.

### 3.3 Regression models

In order to test Hypotheses 2 and 3, a series of OLS regression models are used, where the dependent variable is underpricing. For the purposes of Hypothesis 2 and 3, underpricing is calculated using closing prices and it is adjusted by the return of the CCI30 index. Such a regression analysis is suitable in this case, because it can show the effect of the independent variables on the dependent variable. For the regression analyses, underpricing is further adjusted to satisfy normality by performing a natural log transformation. Following Thijn (2018), prior to the log transformation, the constant 2 is added to the market adjusted underpricing before the logarithm is taken, in order to deal with negative values in the untransformed underpricing variable.

Equation (3) shows the regression equation of model H2 used to test Hypothesis 2, where the dependent variable is underpricing, calculated as explained above. Model H2 shows the relation between underpricing and the variables that are related to the transmission of information from the venture to prospective investors, including control variables. Model H2:

---

<sup>6</sup> According to <https://cci30.com/>, the index covers over 92% of the market capitalization of exchange listed cryptocurrencies.

$$\begin{aligned}
UP_{adj} = & \alpha + \beta_1 \mathbf{whitepaper} + \beta_2 \mathbf{ERC20} + \beta_3 \mathbf{regulated} + \beta_4 \mathbf{duration} \\
& + \beta_5 \mathbf{tradingvolume} + \beta_6 \mathbf{amountraised} + \beta_7 \mathbf{sentiment} \\
& + \beta_8 \mathbf{crowdedness} + \beta_9 \mathbf{targetreached} + \epsilon
\end{aligned} \tag{3}$$

Equation (4) shows the regression equation of Model H3 used to test Hypothesis 3, where the dependent variable is underpricing, defined identically to underpricing in Model H2. Model H3 shows the relation between underpricing and variables that are related to the signaling actions of ventures about their quality, including control variables. Model H3.

$$\begin{aligned}
UP_{adj} = & \alpha + \beta_1 \mathbf{icoprice} + \beta_2 \mathbf{distributed} + \beta_3 \mathbf{preico} + \beta_4 \mathbf{duration} \\
& + \beta_5 \mathbf{tradingvolume} + \beta_6 \mathbf{amountraised} + \beta_7 \mathbf{sentiment} \\
& + \beta_8 \mathbf{crowdedness} + \beta_9 \mathbf{targetreached} + \epsilon
\end{aligned} \tag{4}$$

The OLS regression models and all statistical test are performed using the statistical analysis software STATA. The regression models are estimated using Huber–White (heteroscedasticity robust) standard errors. In order to control for variation in underpricing that is not attributable to the information conveyance or the quality signaling hypotheses, both models include a number of control variables. These variables, along with the variables of interest are presented in Appendix A and are explained in the following sections.

### 3.4 Independent Variables of Interest

*Whitepaper* is a continuous variable obtained by dividing the file size of the whitepaper in kilobytes by the number of pages in the document (kB per page). In a large number of ICO research papers, variables related to whitepaper in regression models are constructed as a binary variable based on the existence of a whitepaper of a given ICO. Constructing such a binary variable would not add value to the current analysis, as the sample used for this paper contains only ICOs, which issue tokens that are eventually listed on an exchange. ICOs of exchange listed tokens typically have a whitepaper. As a result, there would be very limited variation in a variable that simply captures the existence of a whitepaper. The length of the whitepaper and the size of the files are accessible, however both are one dimensional measures of the amount of information transmitted in the whitepaper. By inspecting whitepapers of several ICOs, it is evident that these documents lack a common form and the content of whitepapers are very different. Looking at the size of the

document per page on the other hand can give an indication of useable information conveyed in the whitepaper.

*ERC20* is a binary variable, which is equal to 1, if the given ICO issued a token using the ERC20 protocol.

*Regulated* is a binary variable, which is equal to 1, if the country of issuance of the ICO, is a country with regulation about disclosure requirements that applies to ICOs prior to the start date of the ICO. In order to construct this variable, analysis and publications on global cryptocurrency regulation were consulted (Global Legal Group, 2019; The Law Library of Congress, 2019; De Vauplan, 2018; Securities Commission Malaysia, 2020). An overview of countries that have introduced ICO specific regulation is presented in Appendix B.

*Icoprice* is the offer price of the ICO in USD. If the offer price was published in a currency other than USD, the exchange rate on the first day of the ICO was used to calculate the offer price in USD.

*Distributed* is the proportion of tokens distributed to investors in the respective ICO.

*Preico* is a binary variable, which is equal to 1, if a pre-sale preceded the ICO for the given token.

### 3.5 Control Variables

*Duration* is the number of days between the last and the first day of the ICO.

*Tradingvolume* is a continuous variable that captures the trading volume of the token on the first trading day in USD.

*Amountraised* is a continuous variable equal to the total funds raised in the ICO in USD.

*Sentiment* is the return of the CCI30 cryptocurrency index in the 30 days prior to the ICO start date.

*Crowdedness* is a continuous variable, calculated as the number of ICOs 30 days before and after the given ICO.

*Targetreached* is a binary variable, which is equal to 1, if the total funds raised in the ICO is larger than the hardcap fundraising goal of the ICO.

## 4 Data

### 4.1 Data Collection and Processing

The data were collected from three different ICO data aggregator websites that are all commonly used in the ICO literature: [icobench.com](http://icobench.com), [cryptocompare.com](http://cryptocompare.com) and [coinmarketcap.com](http://coinmarketcap.com). [Icobench.com](http://icobench.com) is one of the most widely used sources of ICO data in the literature (Lyandres et al., 2018). Its recognition is due to its large number of variables and extensive coverage. At the time of data collection (2019 June), [Icobench](http://icobench.com) aggregates data on more than 5,500 ICOs. A possible disadvantage of this large coverage is that [icobench](http://icobench.com) also covers ICOs that never complete, either due to fraud or low-quality venture. This however is only a limited concern in this paper, as only ICOs of exchange listed tokens are considered in the data sample. [Icobench](http://icobench.com) is used to collect variables related to the token sale, such as offer price, ICO dates, caps, raised amount, token distribution, country of issuance etc. [Cryptocompare.com](http://cryptocompare.com), similarly to [Icobench](http://icobench.com) is an aggregator website that collects a large number of variables, however only about 2,000 ICOs are covered. In this thesis, data from [cryptocompare.com](http://cryptocompare.com) is used to replace missing data in observations collected from [icobench.com](http://icobench.com), primarily the start and end dates of the ICO and token distribution variables. [Coinmarketcap.com](http://coinmarketcap.com) (CMC) is used for market price data of the issued tokens, such as opening and closing prices and trading volume. CMC is regarded as the most reliable source for historical token price data (Fenu, Marchesi, Marchesi, & Tonelli, 2018). Additional data sources are Wharton Research Data Services (WRDS) database for exchange rates and [cci30.com](http://cci30.com) for the CCI30 cryptocurrency index.

In order to construct the dataset, the following steps were taken. All ICOs listed on [icobench.com](http://icobench.com) were downloaded using a web scraper application. Similarly, a list of all cryptocurrencies listed on [coinmarketcap.com](http://coinmarketcap.com) was downloaded. As there is no standard identifier for ICOs, such as the ISIN for the issuance of securities, the two lists were merged using a set of custom merging variables. Several ICOs can have the same name and it is common that the same symbol is used by different tokens. The lists were first merged using the combination string of the name, the symbol and the website of the ICO as the merging variable. The remaining unmerged observations were merged using combinations of symbol and website or name and website. This merging procedure results in a list of exchange listed ICOs, for which a unique ICO identifier is created.

A set of custom Excel VBA macros were used to collect ICO data from icobench and market price data from CMC for tokens on the merged list. Using the unique ICO identifier created before, the ICO data is then merged with the daily historical price data from CMC for the first available date with closing price after the start of the ICO. The above procedure results in a sample of 783 ICOs. This sample is complemented with hand collected data from cryptocompare.com for token distribution variables and ICO dates. As the next step, observations that have missing values for the start or end date of the ICO or the ICO price are dropped. Furthermore, observations are dropped if there are inconsistencies regarding ICO dates, for example if the ICO end date precedes the ICO start date. Additionally, observations of ICOs of tokens that are not listed for at least 30 days on an exchange after the ICO are dropped from the sample. As a result, the final sample used in this paper consists of 582 observations.

## 4.2 Summary Statistics

Table 1 presents the summary statistics for underpricing (raw return without market correction and without logarithmic transformation), the independent variables of interest and the control variables. The sample used in this paper includes 582 ICOs in a period from August 2015 to April 2019. Mean underpricing is 79%, meaning that the ICOs on average, trade for 79% higher at the end of the first trading day when compared to the ICO offer price. Interestingly, while the mean underpricing is highly positive, the median underpricing is negative. The median value of 25% is an indication that a larger number of ICOs are overpriced than underpriced. The minimum value of raw underpricing is -100%, which occurs when the token of the given ICO loses all its value over the first trading day. Finding lower median underpricing than mean underpricing is common in ICO samples. Benedetti and Kostovetsky (2018), Thijn (2018), and Momtaz (2019) find a substantially lower median than mean underpricing. In line with these findings, Momtaz (2019) finds that the amount left on the table (calculated by multiplying underpricing by the amount raised in the ICO) is negative in nearly half of the ICOs.

On average, whitepapers are 115.9kB/page with a range between 3.3 to 1,502. The median of 54.41 kb/page indicates that the size per page is slightly skewed to the right. 88% of the ICOs in the sample are utilizing the ERC20 protocol and 16% of the ICOs were issued in country with financial regulation about disclosure requirements that applies to ICOs prior to the start date of the ICO. The mean ICO price is \$14.13, whereas the median price is \$0.12. The minimum and maximum values

show large range of offer prices, as the cheapest ICO was \$0.00001, whereas the most expensive ICO was offered at \$7554.6. The summary statistics regarding ICO price are comparable to ICO price statistics in the sample of other research papers. For example, Benedetti and Kostovetksy (2018) find an average ICO price of \$22.7 and a median of \$0.14 in a sample of 609 ICOs of listed tokens. On average, 54% of coins are distributed to investors, with the median percentage sold to investors at 55%. The nearly equal mean and median values suggest a symmetrical distribution. It is also noteworthy that the issuers of at least one ICO sold all coins to investors, retaining no coins for the team. At the other end of the spectrum, the lowest rate of distribution is 2% of coins being sold to investors. The statistics of the binary variable *Preico* show that 44% of ICOs conducted a pre-sale.

Table 1: Summary Statistics

	N	Mean	Median	SD	Min	Max
Underpricing (raw)	582	0.79	-0.25	4.30	-1.00	53.94
Whitepaper	558	115.90	54.41	177.04	3.30	1,502.14
ERC20	582	0.88	-	0.33	0	1
Regulated	582	0.16	-	0.37	0	1
Icoprice	582	14.13	0.12	313.26	0.00	7,554.60
Distributed	509	0.54	0.55	0.21	0.02	1.00
Preico	582	0.44	-	0.50	0	1
Sentiment	582	0.10	0.05	0.41	-0.79	1.16
Crowdedness	582	68.42	71.00	33.35	2.00	128.00
Amountraised	514	17.27	10.38	25.24	0.02	320.00
Duration	582	40.11	31.00	48.52	1.00	572.00
Targetreached	514	0.11	-	0.32	0	1
Tradingvolume	581	3.26	0.12	16.82	0.00	228.24

In terms of the control variables, only 11% of ICOs have managed to reach their fundraising target in the sample. Market sentiment is positive on average and the distribution appears to be relatively symmetric as mean and median values are close. Similarly, the distribution of crowdedness does not appear to be skewed, as mean and median number of ICOs within 30 days of the ICO is around 70. The minimum number of ICOs in the period is 2, whereas the maximum is 128. The most successful ICO raised \$320 million, while the least successful raised less than \$20,000. ICOs raised \$17million on average. The shortest ICO lasted 1 day, whereas the longest concluded in 572 days,

while on average, an ICO lasts about 40 days. The trading volume on the first trading day was \$3,260 on average, while the median volume was substantially lower at \$120, showing that there is positive skewness. In general, summary statistics show some trademark distinct characteristics of ICOs.

## 5 Results

The following section will present the empirical results of this paper. Firstly, the results of the t-test are shown for establishing the existence of underpricing in the sample for Hypothesis 1. Secondly, the results of the regression analyses for Hypothesis 2 are shown. Thirdly, the results of the regression analyses for Hypothesis 3 are presented.

### 5.1 Underpricing in ICOs

In order to test the first hypothesis and establish that ICOs are underpriced in the sample on average, a one sample t-test was conducted to investigate whether the mean underpricing is significantly different from zero. To ensure the robustness of the results, the t-test is performed on several measures of underpricing, which include calculating with closing and opening prices, as well as correction for market return. The large sample of size of 582 observations ensures that the normal distribution of the sample can be assumed for the purposes of the t-test. Table 2 presents the results of the t-test. The mean underpricing ranges from 0.593 (59.3%) when using opening prices and CCI30 index as market correction, to 0.785 (78.5%) when calculating underpricing with closing prices without market correction. It is evident that using opening prices of the first trading day to calculate underpricing results in lower underpricing than closing prices. Furthermore, market corrected underpricing measures result in lower mean underpricing when compared to raw measures. Nevertheless, even the lowest mean for the ‘Opening CCI’ measure is significantly different from zero at the 1% significance level. Overall, it can be concluded that underpricing is positive and significant in the sample, regardless of the underpricing measure used. Therefore, there is evidence to support Hypothesis 1, that ICOs are significantly underpriced on average in the sample.

Table 2: T-test Mean Underpricing

	N	Mean	SD	t-value	p-value
Closing raw	582	0.785	4.298	4.400***	0.000
Closing BTC	582	0.783	4.298	4.400***	0.000
Closing CCI	582	0.782	4.297	4.400***	0.000
Opening raw	582	0.596	3.548	4.050***	0.000
Opening BTC	582	0.594	3.548	4.050***	0.000
Opening CCI	582	0.593	3.547	4.050***	0.000
Closing CCI (ln)	582	0.699	0.649	26.000***	0.000

Note: The significance level of 1% is indicated \*\*\*.

## 5.2 Information Conveyance

Table 3 presents the results of the OLS regression analysis used to test the information conveyance hypothesis. Model 1-3 include the independent variables of interest separately, whereas Model 4 includes all three variables of interest in one model. Model 5 includes all variables of interest and control variables. The results show that the coefficient for *Whitepaper* is negative and significant in all models of the regression analysis, however significance reduces to the 5% level after the inclusion of controls in Model 5. Similarly, the coefficients for binary variables *ERC20* and *Regulated* are negative in all models, although coefficients for these two variables lack significance. Examining model 5, it can be seen that coefficients for all control variables are significant at least at the 10% level, except the coefficient for *Targetreached*. The explanatory power is low, as for Model 4 and Model 5 it is 1.2% and 17.5% respectively, meaning that the models explain 1.2% and 17.5% of the variation in underpricing.

The significant and negative coefficient of *Whitepaper* variable suggests that as the information content (as proxied by the size per page) of the whitepaper increases, the level of underpricing decreases, finding evidence for hypothesis H2a. 100kB per page increase in the *Whitepaper* variable is associated with a 3% decrease in underpricing.<sup>7</sup> In standardized terms, one standard deviation increase in the size per page of the whitepaper is related to a 7.53% decrease in

<sup>7</sup> The estimates of regression coefficients are interpreted as follows:  $\% \Delta y = 100(\exp(\beta) - 1)$ .

underpricing.<sup>8</sup> As increasing the information content of the whitepaper is a relatively simple effort from the issuers point of view, this finding suggest that the strong relationship between *Whitepaper* and underpricing is not only statistically, but also economically significant.

Table 3. Determinants of underpricing – information conveyance hypothesis

	(1)	(2)	(3)	(4)	(5)
<b>Whitepaper</b>	-0.00034*** (0.00372)			-0.00033*** (0.00550)	-0.00030** (0.01906)
<b>ERC20</b>		-0.07768 (0.40135)		-0.04994 (0.58636)	-0.00293 (0.97519)
<b>Regulated</b>			-0.08221 (0.26420)	-0.08414 (0.24677)	-0.06184 (0.42836)
Sentiment					0.27547*** (0.00011)
Crowdedness					-0.00317*** (0.00019)
Amountraised					0.00000*** (0.00843)
Duration					-0.00136* (0.09965)
Targetreached					0.14520 (0.14040)
Tradingvolume					0.00001*** (0.00081)
Constant	0.73889*** (0.00000)	0.76742*** (0.00000)	0.71235*** (0.00000)	0.79484*** (0.00000)	0.93406*** (0.00000)
Observations	558	582	582	558	495
R-squared	0.00880	0.00154	0.00216	0.01185	0.17490

Note: Dependent variable is market adjusted and log transformed underpricing. The significance level of 10%, 5% and 1% is indicated \*, \*\* and \*\*\* respectively. P-values are reported in parentheses. Standard errors are robust to heteroscedasticity.

<sup>8</sup> Beta coefficients based on standardized variables are reported in Appendix C.

Hypothesis H2b theorizes a negative relationship between the use of ERC20 protocol and underpricing. While the results show the expected negative relationship, the coefficient for ERC20 is insignificant, therefore there is not enough evidence to support hypothesis H2b. Hypothesis H2c posits that ICOs that are issued in country with strict cryptocurrency regulation are less underpriced. The coefficient of the variable *Regulated* indicates that the relationship between regulation and underpricing is negative, however a lack of significance for this coefficient in any of the models means that there is not enough evidence to support hypothesis H2c. In summary, the information conveyance hypothesis is partially accepted as the information content of whitepapers is significantly and negatively related to ICO underpricing.

### 5.3 Quality Signaling

Table 4 presents the models used to test the signaling hypothesis. Similarly to the models presented in Table 3, model 1-3 include the independent variables of interest separately, whereas Model 4 includes all three variables of interest in one model. Model 5 includes all variables of interest and control variables. The results show that the coefficient for *Icoprice* is negative and significant at the 1% level in all models of the regression analysis. The coefficient for *Distributed* is positive and insignificant in model 2 and model 4, however the coefficient becomes significant at the 5% level as controls are added in model 5. The coefficient of *Preico* is negative and significant at the 1% in all three models. The coefficient of the control variable *Duration* is insignificant in model 5, however all other control variables have significant coefficients at least at the 10% level. The R<sup>2</sup> of model 5 shows that 20.25% of the variation in underpricing is explained by the model.

Hypothesis H3a states that the level of underpricing is negatively related to the offer price of the ICO. Given the negative and significant coefficient of *Icoprice* in all models, there is enough evidence to support hypothesis H3a. An increase in the ICO price by one USD is associated with a decrease of 0.008% in underpricing, while one standard deviation increase in ICO price amounts to a 4.57% decrease in underpricing. According to hypothesis H3b, the level of token distribution is positively related to underpricing, meaning that it is expected that the higher proportion of coins are distributed to investors in an ICO, the higher the underpricing. The coefficient for *Distributed* is positive and significant in model 3, therefore there is evidence to support Hypothesis H3b. One standard deviation increase in the proportion of tokens distributed to investors is related to an increase in underpricing of 9.89%. Hypothesis H3c theorizes that the level of underpricing is lower

if there is a pre-ICO token sale. This prediction cannot be rejected by the results of the regression analysis as the coefficient for the *Preico* variable is negative and significant in all models, meaning that an ICO with a pre-sale is 19.47% less underpriced, compared to an ICO with no pre-sale.

Table 4. Determinants of underpricing – signaling hypothesis

	(1)	(2)	(3)	(4)	(5)
<b>Icoprice</b>	-0.00002*** (0.00000)			-0.00004*** (0.00000)	-0.00008*** (0.00000)
<b>Distributed</b>		0.11214 (0.41688)		0.12856 (0.34097)	0.28159** (0.04853)
<b>Preico</b>			-0.33785*** (0.00000)	-0.30848*** (0.00000)	-0.21650*** (0.00014)
Sentiment					0.22169*** (0.00395)
Crowdedness					-0.00210** (0.01074)
Amountraised					0.00000* (0.09518)
Duration					-0.00110 (0.21166)
Targetreached					0.17123* (0.09809)
Tradingvolume					0.00001*** (0.00121)
Constant	0.69955*** (0.00000)	0.60587*** (0.00000)	0.84724*** (0.00000)	0.74011*** (0.00000)	0.73521*** (0.00000)
Observations	582	509	582	509	449
R-squared	0.00013	0.00150	0.06687	0.06320	0.20248

Note: Dependent variable is market adjusted and log transformed underpricing. The significance level of 10%, 5% and 1% is indicated \*, \*\* and \*\*\* respectively. P-values are reported in parentheses. Standard errors are robust to heteroscedasticity.

Based on these results, the quality signaling hypothesis cannot be rejected, as the ICO offer price, the proportion of coins distributed to investors and the existence of a pre-sale are all significantly related to ICO underpricing. Furthermore, the interpretation of the coefficients show that the results are also economically significant. The standardized coefficients show that even the weakest relationship between ICO price and underpricing shows economic significance. The relations of *Whitepaper* and *Distributed* to underpricing are comparable in strength, while the strongest relationship is between underpricing and the existence of a pre-sale.

#### 5.4 Multicollinearity

The OLS regression models above assume that there is no multicollinearity present between the independent variables of the models. Correlation between independent variables in OLS models will not ensure unbiased estimators, therefore it needs to be verified that the models are free of multicollinearity. A common approach to ensuring that the significance of the regression models is not influenced by multicollinearity is by estimating the variance inflation factor (VIF). The mean factor provided by the VIF test ranges between 1.05 and 1.07 for the model specifications presented in Table 3 and Table 4, which are well below the commonly used multicollinearity benchmark of 10. (Chatterjee & Hadi, 1986). The detailed results of the VIF test are presented in Appendix D. Considering the outcome of the VIF test, it can be established that there is no evidence of multicollinearity between the variables in the regression models.

## 6 Discussion and Limitations

### 6.1 General Discussion

The results of the analysis provide valuable insight on the determinants of underpricing in ICOs. From the results, it is evident that ex-ante uncertainty explains underpricing in an ICO setting. The following section discusses the results presented in the previous section.

The results of the t-test presented in the previous section show that ICOs in the sample are significantly underpriced and this underpricing is robust to the specification of underpricing. The definition of underpricing and the applicable market correction can be a source of ambiguity. Consequently, underpricing is calculated with a number of different specifications and in all cases, underpricing is positive and significantly different than 0. This finding ensures that the sample of ICOs is underpriced on average, providing a basis for the regression analysis. The findings of

Hypothesis 1 are within the findings of comparable literature on ICO underpricing. Momtaz (2018) and Momtaz (2019) find lower levels of mean underpricing at 8% and 15% respectively, whereas Benedetti and Kostovetsky (2018) and Adhami et al. (2018) find larger levels of average underpricing at 82% and 929%.

The first set of regression models provide partial support for the information conveyance hypothesis. The size per page measure of whitepaper information content is negatively related to underpricing. In a similar fashion, Florysiak and Schandlbauer (2019) find that the information content of ICO whitepapers is a significant determinant of underpricing, which is in line with the findings of this paper, although the measure used is different.

Although I find no evidence that the use ERC20 protocol is significantly related to underpricing, this finding still carries valuable information. The lack of significance is possibly due to the lack of variation in the data, as 88% of the ICOs in the sample use the ERC20 protocol. This is explained by the practice of issuing a token on the ERC20 protocol first, while the development of an own platform is ongoing (Fenu et al., 2018). As a result, there are a several tokens that exist both on the ERC20 protocol first and are subsequently transferred to their own platform, therefore the data can be subject to inconsistencies in the reported token platform.

Based on the results, the hypothesis that ICOs issued in a country with ICO specific regulation are less underpriced cannot be accepted. The regression results and the summary statistics lead to similar conclusions, as in the case of the ERC20 variable. There is little variation in the data for this binary variable and although the sign of the coefficient is negative as hypothesized, significance is low in all models. Cryptocurrency regulation has been on the agenda in many legal jurisdictions, however progress is slow in this area. For example, Russian legislators have proposed a “Digital Financial Assets” bill that would cover cryptocurrencies, however the adoption of the bill has been postponed several times (Alexandre, 2019). There are only a handful of countries that have ICO specific regulation and even countries that recognize cryptocurrencies specifically, often apply regulatory requirements only to issuers of security tokens. Besides ambiguous regulation, it is difficult to verify the true impact of introduced regulations on ICO issuers. China has officially banned ICOs as of 2017 September, however there is evidence that there is an ICO market in China despite this seemingly strict approach (Global Legal Group, 2019).

The second set of regression models provide support for the quality signaling hypothesis, as the coefficients of all three variables of interest are significantly related to underpricing and the signs of the coefficients are in line with the expectations. ICO price is negatively related to underpricing, meaning that on average more expensive issues are less underpriced. This finding is partially in line with the findings of Thijn (2018) who also finds a negative relationship between offer price and underpricing, but only in models that calculate underpricing using an extended time frame.

The token distribution of ICOs is also a significant determinant of underpricing, as shown by model 5 in Table 4. The positive coefficient can be interpreted in a way that ICOs that distribute a large proportion of tokens to investors are more underpriced on average. Felix and von Eije (2019) in their empirical analysis examine the relationship between underpricing and token retention – defined as the proportion of tokens not distributed by the issuer – however find no significant relationship.

The binary variable of the existence of a pre-sale is a significant determinant of underpricing. The negative coefficient implies that issuers that conduct a pre-sale prior to the ICO, experience lower underpricing on average. This finding is in line with the study of Benedetti and Kostovetsky (2018). As around half of the ICOs in the sample conduct a pre-ICO, there are no concerns regarding a lack of balance in the data. These findings underline that a pre-ICO can act as a gauging mechanism for issuers to more efficiently price their sales and can provide the basis for further research in the mechanics of pre-ICOs.

## 6.2 Implications of Results

The findings of this thesis provide some practical implications. The results show that ICO issuers can signal their quality and convey information to investors for a more efficient token sale. From the investor's point of view, the results highlight the ICO characteristics investors need to consider to identify underpriced issues. This finding however means that investors are susceptible to signals and information from issuers and this can be a concern in the ICO market, as it is difficult to ensure the validity of information. In a market with no formal investor protection policies in place, market feedback is the only penalty mechanism ensuring that fraudulent ICOs do not survive. This highlights the need for regulation, in the form of disclosure and registration requirements for ICOs, to protect investors and assure the evolution of the ICO market.

### 6.3 Limitations and Future Research

The results and discussions presented above should be evaluated, while keeping the limitations of this research in mind. The following section aims to present these limitations and suggests possible areas for future research. Concerns about data quality are inherent in cryptocurrency literature, as there is lack of standardized data sources and this issue has been highlighted by previous cryptocurrency literature. Concerns about data quality can be reduced by crosschecking data sources, however these attempts are often hindered by the lack of unique identifiers and merging difficulties. As a consequence, cross-checking is time consuming and often entails manual corrections, which become less feasible as the sample size increases. Due to the scope of this thesis and time limitations, cross-checking was limited to missing observations in order to ensure an appropriate sample size, however no cross-checking was performed in order to ensure the accuracy of data for non-missing observations.

Regarding future research, I identify three interesting areas to investigate further, as a possible follow-up to this paper. Firstly, the information content of whitepapers needs to be captured. The recent advancement of artificial intelligence and developments in the area of natural language processing can help analyze the contents of whitepapers and establish a relationship between underpricing and the true information content of whitepapers, instead of proxy variables. Secondly, this paper has found evidence that the existence of pre-ICOs is negatively related to underpricing. It would be valuable to explore this relationship further, by examining the price adjustments that are undertaken by issuers following a pre-ICO. This would provide additional basis for understanding the underpricing phenomenon in ICOs. Thirdly, the analysis on the effects of regulation on underpricing need to be revisited. ICO specific regulation is in its infancy, however as distributed ledger and blockchain systems become more widely adopted, it is expected that specific regulatory action will follow in a larger number of countries. As more ICO specific regulation is introduced in a widespread manner, the impact of regulation on the level of underpricing can provide important insight for regulatory bodies on the effectiveness of regulation.

## 7 Conclusion

This thesis examines the relationship between ex-ante uncertainty and underpricing in an ICO setting. Following the extensive theoretical and empirical research on IPOs, I hypothesize that the lower the uncertainty regarding the future value of the project, the lower the underpricing in the ICO. In more detail, I theorize that lower uncertainty – achieved by reducing information asymmetry or by engaging in quality signaling – is related to lower underpricing. To test the hypotheses, the level of underpricing in the sample is established first. Furthermore, a set of OLS regression analyses are conducted, to establish the empirical relationship between the proxies for uncertainty and underpricing. The results show that ICOs are significantly underpriced and that there is evidence to support the quality signaling hypothesis. On the other hand, there is only partial support for the information conveyance hypothesis, since no significant relationship is found between underpricing and ICO specific regulation. Therefore, a possible topic for future research is to further examine the relationship between regulation and underpricing in ICOs.

## 8 References

- Adhami, S., Giudici, G., & Martinazzi, S. (2018). Why do businesses go crypto? An empirical analysis of initial coin offerings. *Journal of Economics and Business*, *100*, 64–75. doi: 10.1016/j.jeconbus.2018.04.001
- Alexandre, A. (2019, July 9). Russia to Postpone Adoption of Bill on Digital Currencies. *Cointelegraph*. Retrieved from <https://cointelegraph.com/news/official-russia-to-postpone-adoption-of-bill-on-digital-currencies>
- Baron, D. P. (1982). A Model of the Demand for Investment Banking Advising and Distribution Services for New Issues. *The Journal of Finance*, *37*(4), 955–976. doi: 10.1111/j.1540-6261.1982.tb03591.x
- Beatty, R., & Ritter, J. (1986). Investment banking, reputation, and the underpricing of initial public offerings. *Journal Of Financial Economics*, *15*(1-2), 213-232. doi: 10.1016/0304-405x(86)90055-3
- Benedetti, H. E., & Kostovetsky, L. (2018). Digital Tulips? Returns to Investors in Initial Coin Offerings. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3182169
- Boreiko, D., & Sahdev, N. K. (2018). To ICO or not to ICO – Empirical analysis of Initial Coin Offerings and Token Sales. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3209180
- Chanson, M., Gjoen, J., Risius, M., & Wortmann, F. (2018). Initial Coin Offerings (ICOs): The role of Social Media for Organizational Legitimacy and Underpricing. *Thirty Ninth International Conference on Information Systems, San Francisco*. Retrieved from [https://www.semanticscholar.org/paper/Initial-Coin-Offerings-\(ICOs\)%3A-The-role-of-Social-Chanson-Gjoen/ffddfacc0b9a1b52f85f65525326c6b6d4eb3081c](https://www.semanticscholar.org/paper/Initial-Coin-Offerings-(ICOs)%3A-The-role-of-Social-Chanson-Gjoen/ffddfacc0b9a1b52f85f65525326c6b6d4eb3081c)
- Chatterjee, S., & A. S. Hadi. (1986). Influential observations, high leverage points, and outliers in linear regression. *Statistical Science*, *1*, 379-393.
- Chod, J., & Lyandres, E. (2018). A Theory of ICOs: Diversification, Agency, and Information Asymmetry. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3159528

- Ciaian, P., Rajcaniova, M., & Kancs, D. (2018). Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money*, 52, 173–195. doi: 10.1016/j.intfin.2017.11.001
- Cong, L. W., & He, Z. W. (2019). Blockchain Disruption and Smart Contracts. *The Review of Financial Studies*, 32(5), 1754–1797. doi: <https://doi-org.eur.idm.oclc.org/10.1093/rfs/hhz007>
- Davydiuk, T., Gupta, D., & Rosen, S. (2019). De-crypto-ing Signals in Initial Coin Offerings: Evidence of Rational Token Retention. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3286835
- De Vauplan, H. (2018). *Overview of ICO Regulations Around the World* [Presentation Slides]. *Kramer Levin Naftalis & Frankel LLP*. Retrieved from <https://www.kramerlevin.com/images/content/4/6/v2/46184/181115-presentation-ICO.pdf>
- Global Legal Group. , J. (Ed.). (2019). *Blockchain & Cryptocurrency Regulation 2020*. Retrieved from <https://www.globallegalinsights.com/practice-areas/blockchain-laws-and-regulations/>
- Felix, T. H., & von Eije, H. (2019). Underpricing in the cryptocurrency world: evidence from initial coin offerings. *Managerial Finance*, 45(4), 563-578. doi: 10.1108/MF-06-2018-0281
- Fenu, G., Marchesi, L., Marchesi, M., & Tonelli, R. (2018). The ICO phenomenon and its relationships with ethereum smart contract environment. *2018 International Workshop on Blockchain Oriented Software Engineering (IWBOSE)*. doi: 10.1109/iwbose.2018.8327568
- Florysiak, D., & Schandlbauer, A. (2019). The Information Content of ICO White Papers. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3265007>
- Grinblatt, M., & Hwang, C. Y. (1989). Signalling and the Pricing of New Issues. *The Journal of Finance*, 44(2), 393–420. doi: 10.1111/j.1540-6261.1989.tb05063.x
- Howell, S. T., Niessner, M., & Yermack, D. (2019). Initial Coin Offerings: Financing Growth with Cryptocurrency Token Sales. *The Review of Financial Studies*. doi: 10.1093/rfs/hhz131

- Ibbotson, R. (1975). Price performance of common stock new issues. *Journal Of Financial Economics*, 2(3), 235-272. doi: 10.1016/0304-405x(75)90015-x
- Kaal, W. A. (2018). Initial Coin Offerings: The Top 25 Jurisdictions and Their Comparative Regulatory Responses. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3117224
- Loughran, T., & Ritter, J. (2004). Why Has IPO Underpricing Changed over Time? *Financial Management*, 33(3), 5-37. Retrieved from <http://www.jstor.org.eur.idm.oclc.org/stable/3666262>
- Löfgren, K. G., Persson, T., & Weibull, J. W. (2002). Markets with Asymmetric Information: The Contributions of George Akerlof, Michael Spence and Joseph Stiglitz. *The Scandinavian Journal of Economics*, 104(2), 195–211. doi: 10.1111/1467-9442.00280
- Lyandres, E., Palazzo, B., & Rabetti, D. (2018). Are Tokens Securities? An Anatomy of Initial Coin Offerings. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3287583
- Momtaz, P. P. (2018). Initial Coin Offerings. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3166709
- Momtaz, P. P. (2019). The Pricing and Performance of Cryptocurrency. *The European Journal of Finance*, 1–14. doi: 10.1080/1351847x.2019.1647259
- De Nederlandsche Bank. (2018). *Roundtable Cryptocurrencies/ICO's*. Retrieved from [https://www.dnb.nl/en/binaries/Position Paper DNB Cryptocurrencies.pdf](https://www.dnb.nl/en/binaries/Position%20Paper%20DNB%20Cryptocurrencies.pdf)
- Rock, K. (1986). Why new issues are underpriced. *Journal of Financial Economics*, 15(1-2), 187–212. doi: 10.1016/0304-405x(86)90054-1
- Schultz, P. H., & Zaman, M. A. (1994). Aftermarket support and underpricing of initial public offerings. *Journal of Financial Economics*, 35(2), 199–219. doi: 10.1016/0304-405x(94)90004-3
- Securities Commission Malaysia. (2020). *Guidelines on Digital Assets*. Retrieved from <https://www.sc.com.my/api/documentms/download.ashx?id=dabaa83c-c2e8-40c3-9d8f-1ce3cabe598a>

- Shi, C., Pukthuanthong, K., & Walker, T. (2012). Does Disclosure Regulation Work? Evidence from International IPO Markets. *Contemporary Accounting Research*, 30(1), 356–387. doi: 10.1111/j.1911-3846.2012.01158.x
- The Law Library of Congress. (2019). *Regulation of Cryptocurrency Around the World*. Global Legal Research Center. Retrieved from <https://www.loc.gov/law/help/cryptocurrency/world-survey.php>
- Thijn, F. (2018). Underpricing in the Cryptocurrency World: Evidence from Initial Coin Offerings. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3202320
- Welch, I. (1989). Seasoned Offerings, Imitation Costs, and the Underpricing of Initial Public Offerings. *The Journal of Finance*, 44(2), 421–449. doi: 10.1111/j.1540-6261.1989.tb05064.x
- Welch, I. (1992). Sequential Sales, Learning, and Cascades. *The Journal of Finance*, 47(2), 695–732. doi: 10.1111/j.1540-6261.1992.tb04406.x
- Yermack, D. (2017). Corporate Governance and Blockchains. *Review of Finance*, 21(1), 7–31. doi: 10.1093/rof/rfw074
- Yi, S., Xu, Z., & Wang, G.-J. (2018). Volatility connectedness in the cryptocurrency market: Is Bitcoin a dominant cryptocurrency? *International Review of Financial Analysis*, 60, 98–114. doi: 10.1016/j.irfa.2018.08.012
- Zetsche, D. A., Buckley, R. P., Arner, D. W., & Ffhr, L. (2017). The ICO Gold Rush: Its a Scam, Its a Bubble, Its a Super Challenge for Regulators. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3072298

## Appendix A

### Overview of variables in regression models H2 and H3

Table A1.

	Variable	Measurement
H2	Whitepaper	Size of the whitepaper document in kB per page
	ERC20	Token issued uses ERC20 protocol (binary)
	Regulated	ICO was conducted in a regulated country (binary)
H3	Icoprice	Offer price in the ICO
	Distributed	Proportion of tokens distributed to investors in the ICO.
	Preico	Existence of a pre-sale prior to the ICO (binary)
Controls	Duration	Duration of the ICO in days
	Tradingvolume	Volume of trading on the first day of the ICO (thousands)
	Amountraised	Funds raised during the ICO (millions)
	Sentiment	30-day return of the CCI30 index, following Felix and von Eije (2019), this variable aims to capture the overall outlook in the cryptocurrency market.
	Crowdedness	Number of ICOs in the 30 days before and after the ICO calculated using a moving reference group with upper and lower inclusion limits based on the start date of the ICO.
	Targetreached	Funds raised are higher than the hard cap of the ICO (binary)

## Appendix B

### Regulatory overview of countries with ICO specific regulation

Table B1.

Country	Authority	ICO Regulation	Disclosure requirements	Start Date
Australia	Australian Securities and Investment Commission	Disclosure and licensing requirements if falls under Corporation Act	Register with authority	September 2017
Canada	Canadian Securities Administration	Security regulation if token classified as security token	Register with authority	August 2017
France	Autorite Des Marchés Financiers	Pacte Act that governs public token sales	Disclosure of relevant information for all ICOs	May 2019
Japan	Central Bank of Japan	Virtual Currency Regulation as part of Payment Services Act	ICO specific requirements	June 2019
Malaysia	Malaysia Securities Commission	ICOs are banned, IEOs heavily regulated	Token sales only allowed by licenced platforms	2020 April
Malta	Malta Financial Services Authority	Crypto specific Digital Innovation Framework	whitepaper submitted to authority	July 2018
South Korea	Financial Supervisory Service	Offering restrictions on ICOs of coins if in violation of securities law	ICOs penalized if in violation of law	September 2017
Switzerland	Swiss Financial Market Supervisory Authority	Security regulation if token classified as security token	Register with authority	September 2017
USA	Securities and Exchange Commission	Federal securities law if fall under securities	Register with authority	December 2017

## Appendix C

Standardized coefficients of regression models presented in Table 3 and 4

Table C1.

	(1)	(2)
Whitepaper	-0.07824	
Icoprice		-0.04678
Distributed		0.09435
Observations	495	449

Note: Dependent variable is market adjusted and log transformed underpricing. Insignificant independent variables and controls are omitted from this table. Model 1 includes variables of the information conveyance hypothesis of Model 5 in Table 3. Model 2 includes variables of the quality signaling hypothesis of Model 5 in Table 4.

## Appendix D

### Variance Inflation Factors

Table D1.

Variable	VIF	1/VIF	Variable	VIF	1/VIF
Crowdedness	1.1	0.905812	Crowdedness	1.13	0.884557
Amountraised	1.07	0.936646	Preico	1.13	0.887044
Sentiment	1.07	0.937303	Sentiment	1.1	0.906057
Regulated	1.05	0.948167	Distributed	1.07	0.938952
Duration	1.04	0.962661	Amountraised	1.05	0.949121
Whitepaper	1.03	0.969333	Duration	1.05	0.953722
Targetreached	1.03	0.969819	Targetreached	1.05	0.953754
Tradingvolume	1.03	0.971817	Tradingvolume	1.04	0.959717
ERC20	1.02	0.979143	Icoprice	1.03	0.970824
Mean VIF	1.05		Mean VIF	1.07	