Do Tokens Behave Like Securities?
An Anatomy of Initial Coin Offerings *

Evgeny Lyandres[†] Berardino Palazzo[‡] Daniel Rabetti[§]

March 2019

Abstract

We construct one of the most comprehensive datasets of Initial Coin Offerings (ICOs) to study the determinants of ICO success, post-ICO returns, longer-term token returns, volatility and liquidity, as well as the evolution of ICO-backed-ventures' social media activity and a measure of their productivity. To overcome the problem of the generally low quality of ICO data, we develop an ICO data quality measure that allows us to validate the robustness of our results using a subset of data with the highest quality. Most of our results for ICOs are consistent with empirical regularities known to characterize Initial Public Offerings (IPOs) of equities. We argue that some theories that were initially developed to explain empirical patterns in the IPO market may be even better suited to explain some of the empirical results that we obtain for ICOs. In addition, some of the discrepancies between our findings for ICOs and corresponding results for IPOs may be traced to differences in institutional settings. Overall, our results contribute to the debate about whether tokens issued in an ICO should be considered securities by showing that tokens tend to behave similarly to equities.

*We thank Katya Malinova and seminar participants at Boston University and Sun Yat Sen University, and participants of 2018 MJAC & CryptoCompare Blockchain Summit and 2019 Mont Tremblant Risk Management and Financial Innovation Conference. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Board or the Federal Reserve System. All errors are our own.

[†]Questrom School of Business, Boston University, lyandres@bu.edu.

[‡]Federal Reserve Board, dino.palazzo@frb.gov.

[§]London School of Economics, d.rabetti@lse.ac.uk.

1 Introduction

In the last two years, a new form of financing – Initial Coin Offering (ICO) – has emerged, fueled by developments in the blockchain technology and its applications. An ICO enables an entrepreneurial venture ("project" henceforth) to raise funds in exchange for cryptographically secured tokens that are intended to be the sole form of payment for the project's future products or services. As of 11/30/2018, over 5,500 entrepreneurial ventures attempted financing using an ICO, raising a staggering $US 30 billion.

Despite the recent popularity of ICOs, there are many open questions about this financing method. First, valuation of ICOs is far from straightforward and has barely been studied in the academic literature.[1] Second, the theoretical ICO literature is just starting to explore the reasons for projects preferring ICO financing to either traditional financing forms, such as angel and venture capital financing, or various types of crowdfunding (e.g., Catalini and Gans (2018), Chod and Lyandres (2018), Cong et al. (2018), Li and Mann (2018), Malinova and Park (2018), and Sockin and Xiong (2018)). Third, optimal regulation of ICOs is not well understood, and various jurisdictions have so far adopted vastly different regulatory approaches.[2]

Our paper contributes to the emerging empirical ICO literature along two dimensions. First, we try to overcome important limitations to using ICO-related data for making empirical inferences. Due to a mostly unregulated nature of ICOs and, especially, due to a completely decentralized ICO process, data regarding various ICO characteristics are scattered among a multitude of online sources, which aggregate various pieces of information regarding ICO characteristics, mostly by retrieving this information from ICO "white papers" and project websites at various points in time. For this reason, various data sources cover subsets of attempted ICOs and the degree of pairwise overlap in coverage varies widely, resulting in large differences in sample sizes used in existing studies.[3,4]

[1]A notable exception is Cong, Li, and Wang (2018), who propose a model of token pricing.

[2]For example, China and South Korea ban ICOs altogether, while many countries, such as Singapore and Switzerland, impose very lax regulatory standards (e.g., Kaal (2018)).

[3]A list of recent papers that examine empirically characteristics of projects attempting ICOs, determinants of ICO success, and regularities in post-IPO token trading include Adhami, Giudici, and Martinazzi (2018), Amsden and Schweizer (2018), Benedetti and Kostovetsky (2018), Bourveau, De George, Ellahie, and Macciocchi (2018), Davvydiuk, Gupta, and Rosen (2018), de Jong, Roosenboom, and van der Kolk (2018), Fisch (2019), Howell, Neissner, and Yermack (2018), Hu, Parlour, and Rajan (2018), Huang, Meoli, and Vismara (2018), Lee, Li, and Shin (2018), and Momtaz (2018a).

[4]As an example, to study the determinants of ICO first-day returns, Momtaz (2018a), Bourveau et al. (2018), Benedetti and Kostovetsky (2018), and Lee et al. (2018) use samples of 224, 365, 408, and 423 ICOs respectively. Our final sample includes

To overcome these difficulties, we obtain and compare data from no less than 10 ICO aggregator websites to build one of the most comprehensive datasets of ICOs, likely covering almost the entire ICO population. Our initial sample contains over 5,500 attempted ICOs and over 4,400 completed ICOs with information on some variables of interest. In addition, we collect the most comprehensive auxiliary data on the time-series evolution of social media activity of ICO-funded projects across various platforms, on the evolution of project-related code updates on the world's leading open-source platform, and on the composition of crypto wallets containing tokens issued in ICOs, which allows us to study determinants of the evolution of the public's interest in ICO projects, the evolution of output of ICO projects, and the evolution of investor participation in ICOs and in post-ICO trading.

The data quality of aggregator websites also varies widely, with large differences in ICO-related reported values across aggregators. This makes inferences obtained by combining data from various sources questionable. To overcome this issue, we provide the first systematic analysis of ICO data quality at both the source-variable level and at the ICO level. Not only we do show that aggregators vary in terms of their data quality – a hypothesis that has been proposed in existing papers, although never carefully examined empirically – but that there is also a substantial variation in the quality of data regarding particular ICO-related variables within the same data source. Our conclusion is that one needs to be careful in using only the most reliable pieces of data obtained from various sources and to verify the robustness of empirical results within a subsample of ICOs with the most reliable data. Bringing discipline to the data collection process and using the most reliable data turn out to be important, as our results overturn some of the conclusions in existing empirical studies of ICOs.

Second, and as importantly, we contribute to the discussion of optimal ICO regulation. One of the most relevant questions that needs to be answered by regulators is: Are tokens securities? On the one hand, since most tokens provide their holders with rights to future products, services, or platform participation, they are described as "utility tokens" as opposed to "security tokens", i.e., they may not be classified as securities according to the Howey test.[5] On the other hand, James Clayton, the SEC Chairman, stated that "the structures of initial coin offerings that [he has] seen promoted involve the offer and sale of securities

617 ICOs with first-day-return over a similar period (i.e., prior to May 2018) and 878 ICOs with first-day return data overall.

[5]According to the U.S. Supreme Court's decision in the case of SEC against Howey in 1946, a transaction is an investment contract if the following criteria are satisfied: 1) there is a monetary investment, 2) there is an expectation of profits from this investment, 3) the investment is in a common enterprise, and 4) any profit from it is a result of efforts of a third party.

and directly implicate the securities registration requirements and other investor protection provisions of [U.S.] federal securities laws".[6]

While the question whether tokens *are* securities may have different answers on the normative level, a policy-relevant question is: Do tokens *behave like* securities? Another way to pose this question is: Do characteristics of ICOs and projects' post-ICO financial and operating performance suggest that ICO investors behave similarly to investors in instruments that are commonly considered securities, such as equities?

To answer this question, we draw parallels between determinants of ICO success and post-ICO performance (i.e., tokens' short-term and longer-term returns, their liquidity and return volatility, as well as their post-ICO productivity) with the vast evidence on Initial Public Offerings (IPOs) of equity. Most of our empirical results for ICOs are consistent with empirical regularities known to characterize IPOs. Moreover, we argue that some theories that were initially developed to explain empirical patterns in the IPO market may be even better suited to helping interpret some of the empirical results that we obtain for ICOs. Interestingly, some of the discrepancies between our findings for ICOs and corresponding results for IPOs may be traced to differences in institutional settings between ICOs and IPOs, such as the absence of mandatory disclosure and the lack of underwriter certification in ICOs. Overall, our findings lend support to the proponents of the view that tokens should be considered and treated as securities.

Our first set of empirical results relates to determinants of ICO success, measured by the (absolute and relative to funding objective) amount raised in an ICO and whether the issued tokens are eventually listed on a crypto exchange. We find that an important determinant of the amount raised is the certification by large/institutional/venture capital investors, who buy tokens in a "presale", prior to the offering of tokens to the general public. This result is consistent with the positive effect of informed investors on IPO success established both theoretically ((e.g., Benveniste and Spindt (1989) and Welch (1992)) and empirically (e.g., Busaba, Benveniste, and Guo (2001) and Dunbar and Foerster (2008)).

The relative success of an ICO is negatively associated with the amount that the project is attempting to

[6]Public statement on Cryptocurrencies and Initial Coin Offerings by SEC Chairman Jay Clayton, from December 11, 2017; https://www.sec.gov/news/public-statement/statement-clayton-2017-12-11. The SEC has recently sued Gladius Network LLC, which conducted a successful ICO in December 2017 that raised $12.7 million, arguing that the ICO did not qualify for an exemption from registration requirements of U.S. federal securities laws. The case was settled in February 2019.

raise, reminiscent of the finding in the IPO literature that success is negatively related to the offering size (e.g., Hanley (1993), Dunbar (1998), and Dunbar and Foerster (2008)). Signaling, a possible explanation of the latter result (e.g., Miller and Rock (1985)), is likely to be especially relevant in the case of ICOs, which are often characterized by extreme information asymmetry, as projects attempting to raise funds via an ICO are typically very young, and many are in the pre-R&D stage. We also find that ICO success is inversely related to entrepreneurs' "skin in the game": The relative and absolute amount raised in an ICO and the likelihood of a token issued in an ICO being listed on a crypto exchange are decreasing in the proportion of tokens sold in the ICO.

Our second set of results concerns the evolution of project-related social media activity around ICOs, as measured by various forms of activity on the four most popular social media platforms used by ICO projects: Twitter, Reddit, Medium, and BitcoinTalk. In the absence of mandatory disclosure of information, such as that present in equity markets, voluntary disclosure and associated discussions on various social media platforms play an important role of mitigating the information asymmetry inherent in ICO-backed projects. Project-related social media activity peaks just before an ICO and tends to drop dramatically following ICO end. The growth in cumulative social media activity is over 70% in the second-to-last quarter before ICO end, and it drops to 10% in the quarter following ICO end. However, this drop in social media activity is smaller for ICOs that were successful in raising funds, at both the extensive and intensive margins, suggesting that the interest of the public is related to the perception of the project's ongoing success.

Our third finding concerns post-ICO evolution of projects' output. While ICO-backed ventures' production process is typically difficult to measure or even quantify, many projects rely on developing code on an open source code development platform, Github. We show that the production of code revisions, which we use as a measure of productivity, tends to slow down considerably following an ICO. This result is reminiscent of the decline in operating performance following IPOs (e.g., Jain and Kini (1994) and Chemmanur, He, and Nandy (2008)). Similar to social-media-based evidence, relative post-ICO reduction in code production is less dramatic for ventures able to raise (significant) funds during their ICO.

Fourth, for tokens that are listed on an exchange, we examine the distribution and determinants of short-term returns following ICO completion (ICO underpricing). Unlike IPOs, in which underpricing is typically corrected during (and measured by returns on) the first day of trading, it typically takes a few weeks for a completed ICO to be listed on an exchange, resulting in the majority of underpricing happening between the end of ICO and the beginning of trading, whereas ICO first-trading-day returns tend to be less informative. Similar to IPOs, ICOs are on average underpriced. Mean (median) equally-weighted return between ICO end and the start of trading exceeds 200% (100%), whereas mean first-day return is roughly 12%, in line with typical IPO underpricing. Smaller ICOs are underpriced more, a result that has a counterpart in the IPO literature (e.g., Beatty and Ritter (1986), Megginson and Weiss (1991), and Michaely and Shaw (1994)). One result that we obtain seems to be different from the corresponding empirical regularity in the IPO market. Contrary to the "partial adjustment" effect, i.e. a positive relation between IPO offer price revision and IPO first-day return (e.g., Hanley (1993) and Bradley and Jordan (2002)), we find a negative relation between the ICO first-day return and the return between ICO end and first trading day, suggesting that ICO investors tend to overreact to information revealed between ICO end and the start of trading.

Fifth, we examine longer-term post-ICO cumulative returns of listed tokens. Obvious data limitations preclude us from examining long-term – three-to-five-years – returns, as is common in the IPO literature. We focus on post-ICO returns for horizons ranging from one month to one year and find that these cumulative post-ICO returns are inversely related to ICO underpricing. This result is consistent with findings in the IPO literature (e.g., Ritter (1984), Ritter (1991), and Ofek and Richardson (2003)) and is likely due to valuation uncertainty surrounding ICOs combined with the impossibility of short selling (e.g., Ljungqvist, Nanda, and Singh (2006)).

Interestingly, we find a negative relation between longer-term post-ICO return and ICO size, inconsistent with the positive relation for IPOs (e.g., Brav and Gompers (1997), Carter, Dark, and Singh (1998), and Teoh, Welch, and Wong (1998)). A possible reason underlying this discrepancy is twofold. First, the positive relation for IPOs may be due to the association between IPO size and underwriter reputation, which is not applicable in the ICO setting. Second, a large amount raised in an ICO may be a sign of overvaluation, which is more plausible in the absence of underwriter certification of ICOs than in the case of IPOs, in which underwrites play a prominent role in value discovery.

Our final set of results concerns the liquidity of tokens and the volatility of their returns. While tokens tend to be considerably less liquid than equities (e.g., Howell et al. (2018)), cross-sectional determinants of token liquidity are similar to those for equities: liquidity is increasing in ICO underpricing and in the amount raised. Interestingly, token liquidity tends to be positively associated with contemporaneous social media activity surrounding the project. The reason is that social media activity is related to investor participation in the token market: The number of cryptographic wallets holding a token (a measure of investor participation) has a very strong association with token liquidity, while partially driving out the explanatory power of social media activity. The positive relation between investor participation and liquidity is also present in equities (e.g. Naes, Skjeltor, and Odegaard (2011) and Blankespoor, Miller, and White (2014)).

Tokens are extremely volatile. The average daily return volatility is 16-19%, compared with 2-3% for equities. Interestingly, token return volatility is strongly negatively related to ICO size, echoing IPO evidence (e.g., Loughran and McDonald (2013)). Consistent with asymmetric information theories, the volatility of tokens associated with projects with larger code revision activity prior to their ICO tends to be smaller.

The remainder of the paper is organized as follows. Section 2 carefully describes the sample construction, various data sources and their limitations, and the derivation of an ICO data quality measure. Section 3 reports summary statistics of the main variables of interest. Section 4 presents our empirical analysis, while Section 5 concludes.

2 Data

2.1 Data acquisition and sample construction

Information on ICOs is scattered among a multitude of documents ("white papers") and online sources, whose availability prior to 2017 is limited and whose quality tends to be poor (e.g., Boreiko and Sahdev (2018)). Several listing websites aggregate ICO information and make it publicly available only starting in early 2017. Unfortunately, the information offered by these aggregators suffers from a number of drawbacks. First, they do not cover the entire universe of ICOs. For example `www.icobench.com`, one of the most popular sources used in the literature so far (e.g., Lee et al. (2018), Huang et al. (2018), and Momtaz (2018a)), has about 50% coverage. Second, information regarding various ICO characteristics is often not up-to-date, contains errors, and some projects are duplicated within the same source with similar or even identical names.

To overcome data coverage issues, researchers sometimes use data coming from multiple sources. However, critical issues arise when attempting to match data across sources. First, there is no unique identifier for each project, making the process of matching across datasets difficult. Second, as many listed ICOs are traded on multiple exchanges – there are 94 exchanges in our sample – several different projects may have the same ticker, which may also coincide with IDs of non-listed projects. In many cases, matching by project name is not helpful because of variations in project names, misspellings, names unrelated to original projects, and outdated or incomplete names.

To mitigate these problems, we adopt a series of measures. We start by selecting ten most popular ICO project data sources, where popularity is measured using historical Alexa Traffic Rank as of 11/30/2018:[7] www.etherscan.io, www.coindesk.com, www.coingecko.com, www.cryptocompare.com, www.icobench.com, www.icodrops.com, www.icorating.com, www.icomarks.io, www.icodata.io, and www.foundico.com. In what follows, we omit www and website address extension when referring to various websites.

etherscan is the most popular source of ICO data. A possible reason is that besides disclosing information on key ICO-related and project-related variables, it provides blockchain transaction data for Ethereum-based ("ERC") tokens, which represent a large portion of tokens issued in ICOs.[8] coindesk is the second most popular source of ICO-related data, but it contains information on fewer variables. Its popularity is mostly due to its role as a source of crypto news and ICO analysis. coingecko, cryptocompare, and icobench tend to provide relatively high-quality data but are limited in the range of variables they cover. icodrops provides some of the most reliable data on the amount raised in an ICO, with sparser coverage of some other variables. icorating mostly provides data on auxiliary ICO variables, discussed below, whereas icomarks has good coverage of the number of tokens issued and offered to investors in an ICO, but lacks coverage of other important variables. The most popular aggregator websites are lo-

[7] Alexa Traffic Rank is one of the leading tools for measurement and comparisons of the popularity of websites, see https://www.alexa.com/siteinfo.

[8] Currently, Ethereum-based tokens are used in 90% of ICOs and are responsible for 75% of ICO proceeds. The top 10 platforms by number of ICOs are: Ethereum, Waves, Stellar, Neo, Scrypt, New Blockchain, Separate Blockchain, Nem, Bitshares and NXT.

cated in the United States and Western Europe. However, we include icodata and foundico to improve the geographical diversity of our sample. Both these sources focus on Pacific Asian and Eastern European ICOs and, albeit being less popular than other aggregators, provide information on a large number of listed ICOs.

Our initial sample contains 17,158 projects across all sources. Since many ICOs are covered by multiple ICO aggregators, we match data across various sources to generate a sample of uniquely identified ICOs. Projects' names and tickers cannot be reliably used for matching for aforementioned reasons, therefore we use projects' website addresses to resolve potential conflicts. A project may have several addresses that we can exploit for matching purposes, however, not all reported addresses are accurate, up-to-date, or even related to the project. Thus, we adopt a website address validation criterion to reduce discrepancies in our matching process. In particular, we match ICOs according to the following order of preference: 1) ICO website address as reported on www.coinmarketcap.com, which is the source of post-ICO price and volume data for tokens traded on exchanges, 2) project website address as reported on some of the aggregator websites, 3) addresses of accounts on the following social media sources: Twitter, Medium, Reddit, BitcoinTalk, LinkedIn, Slack, and Telegram. We validate each match ex-post using the name of the project, the ticker symbol, and some key variable values (e.g, the amount raised and the number of tokens issued in the ICO).

In addition to data provided by the 10 aggregator websites, we obtain information regarding auxiliary variables directly from ICO white papers and project websites. When variables concern a monetary amount (e.g., the amount raised in the ICO and the maximum amount to be raised in an ICO – ICO hardcap) – we convert these values to \$U.S. using exchange rates on the final day of the ICO.

Our final sample comprises 5,430 unique merged projects in 133 countries, carried out between 2013 and 11/30/2018. As evident from Figure 1, all but 33 ICO happened (i.e. had an end date) in 2017 or 2018. ICO activity peaked between September, 2017 and June, 2018 with over 100 ICO-funded projects each month that were able to raise money from investors. During this 10-month period, close to \$U.S. 20 billion were raised, with the peak of almost \$U.S. 6 billion in June, 2018, in which the blockchain project EOS ended its ICO, raising \$U.S. 4.2 billion.

Table 1 contains detailed descriptions of dependent and independent variables used in the empirical analysis. Table 2 summarizes the distribution of ICO data availability across various sources. Almost half of uniquely identified ICOs are covered by at most two aggregator websites, whereas 23% of ICOs are covered by five aggregators or more. In the empirical analysis, we further restrict our attention to a subsample of ICOs in which we have data on the number of tokens issued for sale and/or the amount raised in the ICO. We do so in an attempt to eliminate incomplete ICOs, which are those that are halted prior to offering tokens to investors, as opposed to completed but unsuccessful ICOs (i.e. those that fail to raise money), which we keep in the sample. We also eliminate ICOs that are still ongoing as of 11/30/2018, leaving us with the final sample of 4,411 ICOs for our empirical analysis.

2.2 Data and variables

2.2.1 ICO data and data quality

Unfortunately, there are substantial inconsistencies in the reported values of main ICO characteristics – the amount raised, hardcap, the number of tokens available for sale, and the overall number of tokens issued – across aggregators. Table 3 reports the number of observations for each of the above variables across the 10 sources. All four variables are available for some ICOs in 7 to 8 partially overlapping sources.

In cases in which the value of a particular variable (x) is available across multiple sources, we measure the "relative distance" of the value reported in a given source from consensus (mean) value of this variable across all sources, $\bar{x}$, defined as $\left|\frac{x_i-\bar{x}}{x_i+\bar{x}}\right|$ for source i. If x_i equals the average value, $\bar{x}$, then the relative distance of source i for variable x is zero. If x_i approaches zero or infinity, the relative distance approaches one. Consider as an example, the data available on the total amount raised by Blocklancer during its ICO. The reported values as of February 2019 are \$300,000 (`icobench`), \$5,475,789 (`coingecko`), \$4,420,000 (`cryptocompare`), \$10,000,000 (`icorating`), and \$258,850 (`icodata`), with the mean value across sources being \$4,286,874. The relative distance of the amount raised reported by `icobench` is 0.89, while the relative distance of the amount raised reported by `coingecko` is 0.12. This example, while extreme, illustrates quite a common occurrence in the data.

Table 3 also reports the average relative distance from the consensus value for each of the ten sources across ICOs in which data for a given variable is available across multiple sources. As evident from the table, the values of the amount raised in an ICO, as reported in etherscan, tend to be the closest to consensus, with an average relative distance of 0.068, albeit available for only 242 ICOs. On the other end of the spectrum, data from icodata are the farthest from consensus, with an average relative distance of 0.32. The largest disagreement among the sources is regarding the number of tokens for sale in an ICO, with the average distance across all sources being 0.132. Another interesting finding is that icobench, which is the leading source of ICO data used in the academic literature, belongs to the bottom half of data quality distribution for the amount raised, while being one of the best sources for the other variables. In addition, coingecko appears to be one of the top-quality data aggregators across multiple variables.

The main take-away from the examination of consistency of variables across data sources is that ICOs may differ in the quality of information regarding key variables characterizing them. To mitigate data quality concerns, we construct an ICO-level measure of data quality and perform subsample analyses using only ICOs with the highest-quality data. In constructing our data quality measure, we are guided by the following considerations. First, data quality of an ICO should be increasing in its coverage, namely the number of sources with available data on variables characterizing that ICO. Second, the measure should be increasing in the quality of available sources. Third, the measure should be decreasing in the amount of disagreement among the sources regarding the values of main ICO characteristics.

The first step in building our measure is to identify all sources that report a value for the main four ICO variables. As an example, consider all the data available for Bancor ICO, reported in Panel A of Table 4. The data for the amount raised are consistent across seven sources. We define the consistency of a variable for a given ICO as one minus the mean relative difference of this variable across all sources reporting it, which in the case of amount raised in Bancor ICO equals one. On the other hand, the consistency of values of hardcap is far from one, as there are two sources reporting values of this variable, which are inconsistent with each other. The relative distance for hardcap reported by cryptocompare (icodata) is 0.333 (1), resulting in consistency of $1 - (0.333 + 1)/2 = 0.333$. Moreover, interestingly, both values of hardcap are much lower than the amount raised in Bancor ICO. The reason is that Bancor ignored its stated hardcap and continued to accept funds even after it was reached. This is just one example of possible consequences of lack of regulation in the ICO market.

To account for the number of data sources with available information and for their quality, we first calculate for each available source a measure of quality given by the inverse of that source's mean deviation reported in Table 3. For example, the inverse of mean deviation of the amount raised reported in etherscan (icorating) is 1/0.068 = 14.7 (1/0.115 = 8.7). Then, we compute the relative quality of data coming from a given source for a given variable by dividing the inverse mean deviation by the highest inverse mean deviation across all sources reporting data for that variable. The relative quality of the amount raised data from etherscan is 14.7/14.7 = 1, while the relative quality of icorating is 8.7/14.7 = 0.591.

Next, we move the the level of an ICO and for each variable we take the sum of qualities of all sources reporting information on that variable for a particular ICO to obtain the total quality of that variable. We take into account the consistency of available data for a given ICO by multiplying the total quality of a given variable by its consistency. For, example, the total quality of token supply data for Bancor ICO is 1.533 and the average consistency of this variable across sources is 0.944, resulting in the adjusted quality measure of $1.553 \times 0.944 = 1.466$. The overall data quality for Bancor ICO is given by the simple average of adjusted quality values across the four variables, equaling 1.475.

Figure 2 describes the distribution of our ICO data quality measure and its association with the average number of sources reporting data on each of the four variables for a given ICO and the average consistency of these sources.[9] The figure shows that our data quality measure is clearly increasing in both the number of sources and in their consistency.

In the empirical analysis, in cases of discrepancy in values of a given variable across multiple sources, we use the value for which the sum of source qualities for that variable is the highest across all reported values for the variable. For example, in the case of token supply of Bancor ICO, we use 75,783,855, as this value comes from coingecko, which has the highest quality (0.609) among all sources reporting token supply for Bancor ICO. However, given the importance of carefully validating ICO data in light of its sometimes suspect quality, we perform all the empirical tests also on subsamples of ICOs with the highest data quality.

In addition to the main ICO characteristics (amount raised, hardcap, and tokens issued and offered for sale in an ICO), we define a set of binary variables characterizing projects and ICOs: the occurrence of a

[9]The average consistency is the simple average of ICO-variable-level consistency measure, reported, for example, in the last raw of Panel A of Table 4 for the case of Bancor ICO.

presale of tokens (i.e., sale of tokens to large/institutional/VC investors prior to offering of tokens to general public); the requirement for investors to register in advance in order to participate in the ICO (known as "whitelist"), and the presence of a "know your customer" (kyc) requirement, which obliges token buyers to prove their identity by providing passport, national ID, or driving license information. We also have information on team members involved in ICO-backed projects, as well as information on the type ("industry") of ICO-funded projects. We aggregate industries into five sectors: entertainment, business services, blockchain, other software, and finance. Finally, we use information on ICO location, available for 3,469 ICOs, and aggregate locations into five regions: Western Europe, Canada, and Australia; Eastern Europe; Asia; USA; and rest of the world.[10]

2.2.2 White paper-based variables

For 1,136 ICOs with available white papers, we obtain additional information from examining their contents with the goal of measuring white paper informativeness, which is likely to be inversely related to project opaqueness and to the degree of information asymmetry between ICO issuers and potential investors. Our first such measure is the number of unique words identified by natural language processing (nlp).[11] The second measure is the ratio of "technical" words out of all words appearing in a white paper, with the idea that more technical white papers are found in projects in more advanced stages of development.[12,13]

[10] We also have information on a project's legal form, on the availability of a "minimum viable product", on the presence of maximum and minimum token purchase requirements, on the intended use of ICO proceeds, on the presence of reward programs (aka bounty) and discounts (aka bonus), on the possibility of receiving tokens by means of solving a computationally difficult puzzle (aka mining), and on the presence of an escrow account. These additional variables tend to not be significantly associated with outcome variables in our empirical analysis.

[11] Natural language processing is focused on identifying common roots (e.g., "buy" and "buying"), while eliminating stop words (e.g., "a", "the", and "and").

[12] Technical words are common words used in the Blockchain and computer science white papers, as for instance "block", "node" and "ledger". We build a dictionary with 144 most frequent tech words extracted from words frequency in ICO white papers and based on several blockchain and computer science glossaries extracted from various websites and tech forums. Details are available upon request. See Florysiak and Schandlbauer (2018) for a more detailed analysis of ICO white paper contents.

[13] In addition to these variables, we obtain such white paper characteristics as page count, word count, image count, and .pdf file size, which tend to have lower explanatory power than the number of nlp words and the ratio of technical words.

2.2.3 Post-ICO token prices and returns

For post-ICO token price data we rely on www.coinmarketcap.com, which has become the standard source for researchers interested in measuring token performance post-ICO (e.g., Benedetti and Kostovetsky (2018), Lee et al. (2018), and Howell et al. (2018)). As of 11/30/2018, coinmarketcap has daily price data on 2,046 listed coins and tokens. We match market price data with our ICO sample using ICO website address, as explained above. This matching procedure allows us to identify 878 unique completed ICOs with available information on either the number of tokens for sale or the amount raised in the ICO, which end up being listed on at least one crypto exchange. We also collect data on the number of exchanges on which a token is listed.[14] These data are reported as of 11/30/2018, however we find that there is little variation in the number of exchanges on which a token is listed over time.[15]

One crucial variable that is used in construction of ICO initial returns is the average price paid by ICO investors for tokens sold in the ICO.[16] We compute the average ICO price as the ratio of amount raised at the ICO and the number of tokens issued in the ICO, as reported by coinmarketcap.[17]

2.2.4 Social media data

To examine the time-series evolution of the coverage of ICOs in social media, we rely on four most popular social media channels used by ICO projects: Twitter, Reddit, Medium, and BitcoinTalk.[18] Time-series data are extracted for all projects when the associated social media account is available, not suspended, and is

[14]The ten largest exchanges by trading volume in November 2018, can be found at the following addresses: www.binance.com, www.okex.com, www.hbg.com, www.digifinex.com, www.dobitrade.com, www.upbit.com, www.coinbene.com, www.bibox.com, www.zb.com, and www.hitbtc.com

[15]Tokens are usually listed within a few weeks of ICO end date. Listing on a crypto exchange tends to be costly, with fees that depend on project's reputation and token's potential liquidity, and often exceed $US 500,000. Therefore, projects usually use ICO proceeds to list tokens on exchanges, resulting in little variation in the number of exchanges on which tokens are traded throughout their lives.

[16]Most ICOs employ an accelerated pricing schedule, in which early (and presale) investors pay lower-than-average prices for issued tokens.

[17]We measure the number of tokens 7 days after the beginning of trading or after the appearance of the first observation of market cap (whichever happens later). The reason is that not all tokens reach exchanges immediately, and the number of tokens typically stabilizes within a week.

[18]These sources differ widely in content and format. For instance, Medium articles are usually well-written, have hundreds and even thousands of words, and usually focus on a project's description, solutions, milestones, achievements, and information useful for potential token buyers or token holders. Twitter tweets, on the other hand, are limited to 280 characters, are often written in a poor and abbreviated language, and are used for quick press releases and for sharing videos, photos, and additional content from other social and news channels. Other social media channels, which are less frequently used by ICO projects, include: Facebook, Telegram, VK, Discord, and Slack.

public from its inception to date. We exclude social media accounts that are not clearly related to the project based on information related to the project's name, ticker, website address, and team members. 1,917 ICOs have some social media activity by the last day of their ICO; 1,282 (293, 902, 882) have some Twitter (Reddit, Medium, BitcoinTalk) activity by that day.

To construct our measure of social media presence, we need to aggregate various measures of activity.[19] We define the importance of each of these measures as the inverse of the frequency of appearance of the measure in our sample across all ICOs and dates. We use the number of Medium articles as the numeraire and compute a project's cumulative social media activity on a given day as the sum over all social media measures of the product of cumulative activity measure on that day multiplied by that measure's importance.

2.2.5 Github data

Most ICO-backed projects are in very early stages of development and their R&D output is typically not protected by patents. As a result, many of these projects rely on open source code development. To examine the time-series evolution of project development, we focus on code revisions ("commits") posted on the largest open source platform – Github. 876 projects have some commits on Github by the last ICO day. We distinguish between crucial and less important code revisions by separating commits posted in a project's main repository ("source commits") from those posted in other repositories ("feature commits"). In the empirical analysis, we mainly focus on the more critical code updates, i.e. source commits.

2.2.6 Blockchain transaction data

For 1,087 ERC-based tokens that are listed on exchanges, we collect information on blockchain transactions from `www.ethplorer.io` and `www.etherscan.io`. Each transaction contains information about the addresses of wallets sending and receiving the tokens, the amount of tokens transferred in each transaction,[20] and the transaction's hash and time stamp. After merging blockchain transaction data with the

[19] Twitter's activity measures include tweets, replies, retweets, and likes; Reddit measures are posts, thumbs, and comments; Medium measures include articles, claps, and comments; and BitcoinTalk measures are posts and merits.

[20] ERC protocol does not allow token transactions with non-integer values. As a result, information on token divisibility is reported using a variable called decimals. This variable represents the number of digits after the decimal place. For example, if a transaction reports a number of tokens transferred equaling 150,000,000,000 and the value of decimals equaling 8, we adjust

ICO data, we are left with 637 ICOs with wallet information available as of the first trading day.

We use data on wallet transactions to construct the distribution of wallets containing each token at various points in time. In doing so, we exclude wallets belonging to crypto exchanges, which aggregate holdings of multiple investors, and "genesis wallets", i.e. wallets belonging to ICO issuers that are used to transfer tokens to ICO investors, contributors, and miners.

3 Summary statistics

Table 5 presents summary statistics for key dimensions of ICOs and their outcomes. As evident from Panel A, the average ICO hardcap is $U.S. 70 million, while in more than 50% of the ICOs it is larger than $U.S. 20 million. The percentage of tokens issued to the public in an ICO averages 56% of the total tokens outstanding, and in about 10% of ICOs all tokens outstanding are offered to investors. There are 176 ICOs (4% of the sample) in which some money is raised in a presale to large/institutional/VC investors prior to the official ICO start. 30% of ICOs feature advanced investor registration ("whitelist"), whereas 49% of ICOs have a "know your customer" (kyc) requirement. The average (median) number of team members involved in an ICO is 11 (9).

Industry affiliations are available for 52% of projects in our sample. Among projects with industry information present, the most frequent sector is finance, representing 31% of ICOs, while the least frequent sector is blockchain (9%). Location information is available for 79% of ICOs. Almost a third of ICOs are performed in Western Europe, Canada, and Australia. About 15% of ICOs are U.S.-based. 31% of ICOs are performed in jurisdictions that have adopted crypto-friendly policies, such as Singapore, Hong Kong, Switzerland, Estonia, Malta, British Virgin Islands, and Gibraltar. White papers are available for 26% of attempted ICOs. A typical white paper has about 1,600 unique words filtered using natural language processing. The average amount of unique technology-related words over the total number of unique words is 29%.

45% of ICOs are able to raise an amount that exceeds $U.S.10,000. Conditional on raising money, the average (median) amount raised in an ICO is $U.S. 15 million ($U.S. 5 million). Ventures are able to

the reported value by subtracting 8 zeros, thus obtaining a value of 1,500 tokens transferred.

raise 44% of ICO hardcap on average, and only 26% of ICOs reach or exceed their hardcap. A typical ICO lasts around 47 days, while the largest ICO to date, EOS, lasted 342 days. The ICO of Brave, a free web browser, raised $US 30 million in 30 seconds om May 31 2017, making it the fastest ICO to date. Conditional on raising money, 39% of tokens end up being listed on an exchange. A token is traded on 5 different exchanges on average.

Finally, there is wide dispersion of data quality across ICOs, which ranges between 0.09 and 3.74 with a standard deviation of 0.73, on the order of magnitude of average and median quality measures. This variation highlights the need to focus the empirical analysis on a subset of ICOs characterized by relatively high-quality data.

Panel B presents summary statistics for the social media variables collected from four platforms: Twitter, Reddit, Medium, and BitcoinTalk, and aggregated across various measures of social media activity, as described above. We report cumulative activity for each platform at seven points in time: 90 days prior to the end of an ICO, the ICO start day, the ICO end day, 90 days following the end of the ICO, and -90 days (0 days, 90 days) relative to the first trading day for ICOs that are eventually listed on an exchange. There is wide variation in the extent of social media activity across ICOs. An interesting observation – which we explore further in the empirical analysis – is that social media activity is much higher in the 90 days prior to either the ICO end or the first post-ICO trading day than in the following 90 days, as evident from the differences in cumulative activity values at various points in time. For example the difference between mean log Twitter activity at ICO end and that 90 days prior to ICO end is $1.59 - 0.78 = 0.81$, implying that cumulative Twitter activity increases by a factor of 2.25 on average in the 90 days preceding ICO end, while the difference between mean log Twitter activity 90 days post-ICO and that on the ICO end day is $1.78 - 1.59 = 0.19$, implying a 21% change in cumulative Twitter activity in the first 90 post-ICO days.

Panel C presents the evolution of total, source, and feature commits around ICO end and around the first day of post-ICO trading. Similar to the evolution of social media activity, the growth in code production slows down substantially around the ICO end and around the first trading day. For example, the cumulative number of source commits grows by 25% on average (log growth is $0.78 - 0.56 = 0.22$) in the 90 days prior to ICO end, whereas the growth slows down to 12% over the following 90 days (log growth is $0.89 - 0.78 = 0.11$). This evidence is reminiscent of the decline in post-IPO operating performance documented in Jain and Kini (1994) and Chemmanur et al. (2008).

In Panel D, we report the evolution of the distribution of token holdings across crypto wallets at various dates. The average number of token holders increases from around 100 at the end of the first ICO day, to approximately 250 at ICO end, to over 800 by the time a token starts trading on an exchange, and to over 2,000 three months after the commencement of trading. The concentration of token holdings, measured by its Herfindahl index, is decreasing monotonically in time, as evident from the bottom part of the panel.

In Panel E, we present summary statistics of post-ICO returns of listed tokens over various horizons, as well as their liquidity and return volatility. We winsorize all returns at the top and bottom 5% to attenuate the influence of outliers. We calculate ICO "end-to-open" return using a token's opening price during the first day of trading on an exchange and the average ICO price computed as the ratio of the amount raised and the number of tokens in circulation. Mean (median) ICO end-to-open return – i.e. the adjustment of average ICO price from the ICO end day to its first trading day – is 269% (108%). These very large end-to-open returns are in line with results documented in other studies: Benedetti and Kostovetsky (2018) and Lee et al. (2018) report average ICO returns of 179% and 158%, respectively. As evident from the differences between mean and median returns, high mean end-to-open returns are driven by a few observations with extremely high returns, even after winsorization. These tend to be relatively small ICOs in terms of the amount raised.

Mean first-day return, computed as the difference between the closing and opening prices of the first day of trading, is 12%, and more than 50% of ICOs have positive first-day returns. These findings are in line with the results in Momtaz (2018a). The evidence on positive returns during both the pre-listing period and during the first trading day is reminiscent of the vast evidence on underpricing of IPOs (e.g., Beatty and Ritter (1986) and Ritter and Welch (2002)). Mean first-day return of ICOs is of the same order of magnitude as mean IPO first-day return. For example Loughran and Ritter (2004) report that average IPO first-day return tends to be about 15%, whereas it is 65% during the internet bubble of 1999-2000. However, in the case of ICOs, most of the price adjustment occurs not during the first day of trading but during the period between ICO end and first trading, as reported above. Once both the first-day return and end-to-open return are considered, mean "ICO underpricing" tends to be significantly higher than average IPO underpricing even in times of hottest IPO markets.

Subsequent longer-term post-ICO cumulative returns – measured 30, 90, 180, and 365 days after the first trading day – are on average positive. Mean post-ICO cumulative return ranges from 1% for the 30-day horizon to 60% for the 365-day horizon. However, similar to the case of end-to-open returns, mean long-term returns are driven by a few ICOs with very high cumulative returns. Median post-ICO cumulative return is negative for all horizons, ranging from -31% to -69%. In addition, 67% of 30-day cumulative returns are negative, and this fraction increases to 77% for 365-day cumulative returns. This result is consistent with the large evidence of long-term post-IPO underperformance (e.g., Ritter (1991) and Loughran and Ritter (1995)).[21]

In the last two parts of Panel E, we report summary statistics of liquidity and volatility of tokens traded on crypto exchanges. We follow Howell et al. (2018) and calculate a token's liquidity over a given period as the negative of the log value of the average Amihud (2002) illiquidity measure.[22] Average liquidity value over the various horizons of around 12 is consistent with Howell et al. (2018). The (exponent of the) average liquidity measure is one-to-two orders of magnitude lower than corresponding quantity for stocks: log of average Amihud (2002) liquidity measure for equities is roughly 15 and it is around 17 in the last 20 years (e.g., Amihud (2002) and Harris and Amato (2018)).

Tokens are very volatile. Average and median daily volatilities are 13%-19% over various horizons, translating into 250%-360% annualized volatility. These values are very high both compared to the typical volatility of stocks (averaging annualized 40%-50%, e.g., Andersen, Bollerslev, Diebold, and Ebens (2001)), as well as typical volatility of established crypto currencies, such as Bitcoin, with average annualized volatility on the order of 80%-100%, and Ether, with average annualized volatility of 120%-150%.

4 Empirical analysis

We begin our empirical analysis by examining determinants of various measures of ICO success: the amount raised in the ICO, whether the issued token begins trading on one of the crypto exchanges, and whether the listed token is not delisted or significantly declines in value in the year within listing. We

[21]More recent studies (e.g., Brav, Geczy, and Gompers (2000), Eckbo, Masulis, and Gompers (2000), and Lyandres, Sun, and Zhang (2008)) question the extent of long-term IPO underperformance.

[22]The average Amihud's illiquidity measure over T periods is given by $\frac{1}{T}\sum_{i=1}^{T}\frac{|\ln p_i - \ln p_{i-1}|}{p_i \times volume_i}$, where p_i is the token price on day i and *volume* is the token's trading volume on day i.

then examine longer-term effects of ICOs on project-related social media activity and open source code production. We proceed by examining ICO returns at various horizons: pre-trading, first-day of trading, and cumulative longer-term returns. Finally, we analyze traded tokens' liquidity and return volatility.

4.1 ICO success

As data limitations reduce the number of observations in multivariate regressions substantially, we begin by reporting univariate relations between various ICO characteristics and measures of ICO success. Table 6 presents mean values of five success measures – an indicator equaling one if at least a minimal amount ($U.S. 10,000) was raised, log amount raised, the ratio of amount raised to hardcap, an indicator equaling one if a token was eventually listed on an exchange, and an "disaster indicator" equaling one if a token is delisted within a year of listing or experiences cumulative return lower than -95% a year after listing – for subsamples of ICOs with high and low values of various ICO characteristics. In the case of characteristics taking the form of indicators, high (low) corresponds to the value of 1 (0). In the case of continuous variables, top (bottom) quartile of observations are classified as high (low).

ICOs preceded by a presale to institutional investors and to VCs are more likely to succeed along all dimensions except for avoiding a disastrous outcome post-listing. On the other hand, ICOs with higher proportion of tokens for sale tend to be less successful. Whitelist and kyc requirements are associated with higher likelihood and degree of ICO success, as are the size of the venture's team and the availability of white paper and its informativeness, as measured by the number of nlp words and by the ratio of technical words. Social media presence and the extent of coverage in various social media channels, as well as the presence of Github source repository and the number of commits in it at the time of ICO start are also positively associated with various measures of ICO success, including, in most cases, avoidance of post-listing disaster. In what follows, we examine the relations between ICO characteristics and measures of ICO success in a multivariate regression framework, and suggest interpretations for the observed relations.

In the first two columns of Table 7, we report results of logistic regressions in which the dependent variable is an indicator equaling one if a minimal amount of money was raised in an ICO. To facilitate the interpretation of the results, in all logistic regressions here and below, we report the marginal effects of each independent variable. In the first column (and all odd columns in this table), we estimate the regression using the whole sample of ICOs with available data on all explanatory variables. To mitigate potential issues related to data quality, discussed in Section 2, in the second column (and all even columns), we restrict attention to ICOs belonging to the top tercile of our data quality measure ("High Q"). Because of substantial time-series variation in average ICO characteristics (e.g., the proportion of technology-related words in white papers has been declining over time), and because our sample's end point is very recent, reducing our ability to observe listing of latest ICOs on crypto exchanges, all regressions include time (quarter) fixed effects.

The likelihood of raising at least a minimal amount in an ICO is higher for ICOs with kyc requirement and white paper availability: such ICOs are 20-24% more likely to raise a minimal amount, as follows from the combination of coefficients on kyc and white paper indicators. The probability of minimal success is also increasing in the the size of the venture's team, in cumulative social media activity (aggregated across Twitter, Reddit, Medium, and BitcoinTalk) at the time of ICO start, as well as in the number of source commits at ICO start. In particular, a one-standard-deviation increase in social media activity (1.85) is associated with 5%-6% increase in the likelihood of raising at least a minimal amount in the ICO; A one-standard-deviation increase in log number of source commits (2.12) is also associated with around 6% increase in the probability of minimal success. All this evidence suggests that a reduction in opaqueness surrounding an ICO raises the likelihood of securing at least minimal funding. This interpretation is similar to that of the positive relation between the amount of information and success of IPOs of equity. Theoretically, Benveniste and Spindt (1989) link IPO success to the ability of underwriters to elicit private information from informed investors and Welch (1992) relates IPO success to sequential learning by later (less informed) investors from earlier (more informed) ones. Empirically, Dunbar (1998) shows that IPO success is increasing in investment bank reputation, which is likely inversely related to the degree of information asymmetry between an issuing firm and IPO investors. Asian ICOs are more likely to reach at least minimal success, while U.S. ICOs are less likely to do so.

In columns 3 and 4 of Table 7, the dependent variable is the logarithm of the amount raised in the ICO, and the regression is estimated using OLS for the subsample of ICOs that achieved at least a minimal degree of success, as defined in columns 1 and 2. The amount raised in an ICO is increasing in hardcap – a 1% increase in hardcap is associated with around 0.7% increase in the amount raised. The amount raised is strongly increasing in the presale indicator – the ability to sell tokens to large/institutional investors prior to ICO is associated with over 50% increase in amount raised.[23] This result is consistent with de Jong et al. (2018) and Fisch (2019), who find a positive effect of presale on the total amount raised in an ICO, while it is in contrast with Lee et al. (2018) who find a (marginally) negative effect. Since the presale indicator is strongly associated with the presence of institutional /sophisticated investors, our interpretation of this result is that it is consistent with the asymmetric-information-based theories: having a presale increases the amount of information available to ICO investors, signals ICO credibility to retail investors, and reduces the information asymmetry surrounding the ICO.

Importantly, the amount raised is decreasing in the proportion of tokens for sale, suggesting that "skin in the game" (i.e. the proportion of tokens retained by the entrepreneurs) is an important determinant of ICO success. A one percentage point decrease in the proportion of tokens for sale is associated with 0.6% increase in the amount raised. This finding is consistent with Lee et al. (2018), who report a significantly negative impact of the percentage of tokens for sale on the probability of raising funds in an ICO and on the amount raised. The negative association between ICO success and entrepreneurs' skin in the game is reminiscent of a similar finding in the venture capital literature (e.g., Conti, Thursby, and Rothaermel (2013)).

The amount raised is increasing in the whitelist indicator, as well as in the kyc indicator, suggesting that the process of pre-ICO registration, which may be somewhat akin to the process of book building in IPOs, is conducive of ICO success. Our results are in contrast with Lee et al. (2018), who find that kyc indicator has an insignificant and negative impact on the amount raised and a significant and negative impact on the probability of raising funds. In addition, the amount raised is related to some of the measures of project transparency – the number of identifiable team members and the number of code revisions at the start of the ICO. Finally, Eastern European ICOs, as well as those in Western Europe, Canada, and Australia, tend to raise 30-50% less funds than elsewhere.[24]

[23]Note that we do not include the presale indicator in the logistic regressions in columns 1 and 2, as all 176 ICOs with successful presale ended up raising at least a minimal amount in the ICO, making it impossible to estimate the coefficient on the presale dummy.

[24]In unreported results, the coefficient on the blockchain indicator is positive and significant, suggesting that blockchain firms are likely to raise more money conditional on raising any. This result is supportive of the idea that while an ICO can be (and has been) used as means of raising funds for many types of projects, investors consider blockchain-based financing technology to be best suited for blockchain-related applications.

In columns 5 and 6, the dependent variable is the ratio of amount raised to hardcap. The results are quite similar to those for log amount raised, with one important exception. The ratio of amount raised to hardcap is decreasing in hardcap: a one-standard-deviation increase in log hardcap (1.38) is associated with about 10% reduction in the raised-to-hardcap ratio. This result is consistent with the theoretical argument that large offerings may send a negative signal to the market (e.g., Leland and Pyle (1977) and Miller and Rock (1985) for the case of IPOs) and/or with the idea that in the presence of downward-sloping demand, larger offerings reduce the likelihood of (relative) success (e.g., Scholes (1972)). It is also consistent with empirical evidence that various measures of IPO success are decreasing in offering size (e.g., Hanley (1993), Dunbar (1998), and Dunbar and Foerster (2008)).

In the columns 7-8 of Table 7, the dependent variable is an indicator equaling one if the tokens begin trading on an exchange following an ICO. Consistent with substantial fixed costs of listing,[25] larger ICOs (i.e. those with larger hardcap) are more likely to be listed. ICOs that are preceded by a successful presale are 14-17 percentage points more likely to be listed, consistent with the positive effects of presale on other measures of ICO success. Consistent with the evidence in the previous columns, the listing probability is significantly negatively associated with the percentage of tokens for sale – a one-standard-deviation increase in the proportion of tokens issued in ICO (0.24) is associated with about 2.5 percentage point reduction in the likelihood of exchange listing. This finding is consistent with the evidence in Amsden and Schweizer (2018). Similar to the amount-raised regressions, the probability of listing is increasing in measures of project transparency – kyc, availability of white paper, and social media presence and the number of open source code revisions at ICO start.

The last two columns of Table 7 show that the likelihood of disaster is negatively associated with two (inverse) measures of project opaqueness – the number of team members and the number of source commits at ICO start. For example, a one-standard-deviation increase in log number of team members (0.74) is associated with 7-10% reduction in the likelihood of a disastrous outcome. This result is reminiscent of evidence for IPOs: Busaba et al. (2001) and Dunbar and Foerster (2008) find that the likelihood of IPO failure (withdrawal) is decreasing in the presence of an early (likely informed) VC investor.

Table 8 zooms in on the subsample of ICOs with available white papers to examine the effects of

[25] See for example, https://www.ccn.com/heres-what-it-costs-to-get-your-ico-token-listed-on-an-exchange/ or https://www.businessinsider.com/cryptocurrency-exchanges-listing-tokens-cost-fees-ico-2018-3?r=UK&IR=T .

white paper contents on measures of ICO success. As the number of ICOs with a white paper and a 365-day return history that reach a disastrous end is very small, we do not include disaster-avoidance-based measure of success in this table. The informativeness of a white paper, as measured by the number of nlp words or the proportion of technical words, is positively associated with all four success measures. For example, increasing log number of nlp words by one standard deviation (0.49) is associated with around 2% increase in amount raised. A one percentage point increase in tech ratio is associated with 0.6-0.8% increase in the likelihood of raising a minimal amount and with 0.5% increase in the probability of listing conditional on raising a minimal amount. These findings complement Bourveau et al. (2018), who show a negative impact of the white paper's opacity on the amount raised.

4.2 Evolution of social media activity

Figure 3 plots the evolution of cumulative social media activity around ICO end for a median firm belonging to i) the full sample of ICOs with some social media activity (solid black line), ii) the subsample of ICOs that succeeded in raising at least a minimal amount (dashed-dotted red line), and iii) the subsample of ICOs that failed to do so (dashed blue line). As evident from the figure, up to one quarter prior to ICO end projects associated with both (ex-post) successful ICOs and those associated with (ex-post) failed ICOs exhibit a very similar sharp quarter-to-quarter growth in cumulative social media activity. Following ICO end, social media activity drops dramatically within both subsamples, more so for the subsample of failed ICOs, which reach near-zero social media activity one quarter after ICO end. For successful ICOs, social media activity drops from 75% two quarters prior to ICO end and over 100% during last pre-ICO-end quarter to 25% in the quarter following ICO end and even slower growth thereafter.

In this subsection, we examine determinants of post-ICO social media activity. The dependent variable is the log difference between cumulative social media activity in the 90 days following ICO end and cumulative social media activity at ICO end, i.e. growth in cumulative social media activity in the first post-ICO quarter. In the first column of Table 9, the dependent variable is based on aggregate social media activity, whereas columns 2-5 focus separately on Twitter, Reddit, Medium, and BitcoinTalk activity.

The main independent variable is a measure of ICO success. In Panel A, we focus on the "extensive margin" of success, i.e. on whether a venture succeeded in raising at least a minimal amount in its ICO. In

Panel B, we focus on the "intensive margin", i.e. the logarithm of the amount raised in a successful ICO. We control for the trend in the evolution of social media activity by including the log difference between cumulative social media activity at ICO end and that 90 days prior to ICO end, i.e. growth in social media activity in the last pre-ICO-end quarter. In addition, we control for the level of pre-ICO social media activity by including the level of cumulative activity 90 days prior to ICO end.

Consistent with Figure 3, post-ICO growth in cumulative social media activity is higher for ICOs that manage to raise at least a minimal amount, i.e. for ICOs with raised dummy equaling one. This result holds for overall social media activity, as well as for every social media channel except for Reddit, for which the number of ICOs with social media presence is by far the lowest. The economic effect of the ability to raise money in an ICO on the growth in cumulative social media activity is substantial: projects that raise funds in their ICO exhibit an approximately 20% higher growth in cumulative social media activity in the first post-ICO quarter than those that failed to raise funds in an ICO. These findings are echoed by the intensive margin results in Panel B. The amount raised in a successful ICO is significantly positively related to post-ICO growth in cumulative social media activity. A one-standard-deviation increase in log amount raised (2.22) is associated with about 15 percentage point increase in post-ICO growth in cumulative social media activity. This result holds for all four social media channels separately. The growth in social media activity tends to be negatively related to the level of activity 90 days prior to the ICO. This is the result of the scale effect – the larger the overall level of cumulative social media activity the slower the growth in it.

4.3 Evolution of commits

Similar to the case of social media activity, Figure 4, which describes the evolution of cumulative source commits around all, successful, and failed ICOs, shows that cumulative growth in commits slows down substantially around ICOs. For the full sample of ICOs, the quarter-to-quarter growth rate drops to 3% in the quarter following ICO end from 12% two quarter prior to ICO end. This drop is more dramatic for failed ICOs, which experience a drop from 6% to 0 for a median venture, than for successful ICOs, experiencing a drop from 22% to 5%.

In Table 10, we examine the relation between measures of ICO success and post-ICO projects' output, as measured by the growth in cumulative source commits in the first post-ICO quarter. The dependent variable is the log difference between cumulative commits in the 90 days following ICO end and cumulative commits at ICO end. In the first column of Table 9, the dependent variable is based on total commits, in the second column it is based on source commits, and in the third column it is based on feature commits. As in Table 9, in Panel A the main independent variable is an indicator equaling one for ventures able to raise at least a minimal amount in their ICOs, whereas in Panel B the main independent variable is log amount raised in the ICO.

The positive coefficients on the measures of ICO success in both panels suggest that raising money in an ICO is positively associated with venture's post-ICO output. The relations between both the extensive and intensive margins of ICO success on one hand and the growth in cumulative number of commits in the first post-ICO quarter are positive and marginally significant. The relations become strongly statistically significant when we focus on important code updates, i.e. source commits, in column 2. The economic significance of these relations is quite large as well. Projects that are able to raise at least a minimal amount in their ICOs have an approximately 15% higher growth in cumulative source commits in the first post-ICO quarter than projects that do not raise money in their ICOs. A one-standard-deviation increase in log amount raised in an ICO leads to 11% increase in source commits production in the quarter following ICO.

4.4 ICO returns

As reported in Panel E of Table 5, mean and median returns between ICO end day and first trading day are positive and very high. Mean return on the first day of a token's trading is also positive, although much smaller in magnitude. In this subsection, we examine factors associated with ICO returns. The first three columns of Table 11 report regressions in which the dependent variable is the return between ICO end day and the opening of the first day of trading (end-to-open return). In the first column, we examine the full sample of listed ICOs. In the second column, we focus on the subsample of ICOs that have information on cumulative social media activity, cumulative commits, and distribution of token holdings across crypto wallets on the first day of trading. In column 3, we further restrict our focus to the subset of high-data-quality observations. In all regressions we include time (quarter), industry sector, and geographical region fixed effects.

The first potential determinant of ICO returns is the contemporaneous return of the crypto market as a whole. Two of the largest determinants of crypto market returns are returns on Bitcoin and Ether, whose shares of the overall crypto market range from 85% and 5% respectively in the beginning of 2017 to 50% and 11% respectively in November 2018. Within the full sample of listed ICOs, end-to-open returns are significantly positively related to contemporaneous returns on Bitcoin during the same period. However, this association is not robust within the more restrictive sample, in which the loading on Bitcoin is insignificantly negative, whereas the loading on Ether is significantly positive. This result is likely driven by the fact that due to the restriction on wallet data availability in the subsample in the second column, all ICOs in that subsample are written on top of Ethereum protocol, creating an implicit link between the values of Ether and tokens.

End-to-open return is significantly negatively related to ICO size, as measured by the logarithm of the amount raised during ICO, in both samples. A one-standard-deviation increase in log amount raised (1.84) is associated with a 0.61 reduction in ICO log return for the full sample and a reduction of 0.64 for the subsample in column 2. This result supports the notion that larger ICOs are less obscure and investors in larger ICOs face lower degree of information asymmetry, which, in turn, is associated with lower underpricing. This result is consistent with the negative relation between issue size and underpricing of IPOs (e.g., Beatty and Ritter (1986), Megginson and Weiss (1991), and Michaely and Shaw (1994)). End-to-open return is decreasing in cumulative social media activity at the time of ICO end, consistent with the role of social media in mitigating information asymmetry between token issuers and investors. Surprisingly, end-to-open return is increasing in the number of commits at ICO end. The results are virtually the same when we restrict our sample to high-data-quality observations (column 3).

In columns 4 and 5 of Table 11, the dependent variable is the return on the first day of a token's trading on an exchange, measured as the percent difference between the first-trading-day closing and opening prices. The main finding is that unlike the end-to-open return, the first-day return is not significantly associated with most explanatory variables. This is consistent with most of the ICO price revision (underpricing) occurring in the weeks between ICO end and the start of trading, and not during the first trading day. This is also evident from observing the R squared, which range between 7% and 14% in the first-day return regressions, compared with 21%-37% in the regressions of end-to-open returns.

One exception is the negative relation between the first-day return and the end-to-open return, suggesting that ICO investors overreact to information revealed between ICO end and the token's listing on an exchange. This finding stands in contrast with the partial adjustment to offer price revision in IPOs: IPO first-day return tends to be positively associated with IPO offer price revision during the underwriting process (e.g., Hanley (1993) and Bradley and Jordan (2002)). Another exception is the negative relation between first-day return and the amount raised during the ICO, which is qualitatively similar to the relation between end-to-open return and amount raised. Also in this case, the results are virtually the same when we restrict our sample to high-data-quality observations (column 6).

4.5 Post-ICO longer-term cumulative returns

In this subsection, we examine determinants of post-ICO cumulative returns over horizons, ranging from 30 days to 365 days, complementing Benedetti and Kostovetsky (2018), Lee et al. (2018), and Momtaz (2018b). The results of estimating longer-term post-ICO return regressions are reported in Table 12, which has four sets of estimates for various return horizons. For each return horizon, the first column reports results for the full sample of traded tokens, while the second column considers the subsample of ICOs with information on social media activity, code revisions, and distribution of wallet holdings. As in the ICO return regressions, we include time (quarter) dummies, industry sector dummies, and geographical region dummies in all regressions. Most ICOs with longer-term return observations belong to the highest-data-quality tercile, eliminating the need to report separate results for the "High Q" subsample.

Post-ICO longer-term cumulative returns tend to be related to contemporaneous Bitcoin and Ether returns. At relatively short horizons (30 and 90 days), loadings on Ether are significant and trump those on Bitcoin. This result is reversed at longer horizons – loadings on Bitcoin are significantly larger than those on Ether for 180-day and 365-day returns. In unreported tests examining the distribution of loadings on Bitcoin and Ether obtained in time-series regressions of individual daily token returns, we find that a typical loading on Ether is around 0.7 and that on Bitcoin is around 0.4 when both are included in the regression for virtually all samples and investment horizons. In another set of unreported tests, we find that factors that have explanatory power in the equity market (such as the Fama and French (2015) five factors) have virtually zero power in explaining token returns.

In both samples and for all return windows, the coefficients on both the end-to-open return and on the first-trading-day return are significantly negative. As an example, using the results in columns 1 and 5, a one-standard-deviation increase in the first-trading-day log return (around 400%) reduces the 30-day log return by 44 percentage points and the 180-day log return by 50 percentage points. This reversal result has an analog in the IPO literature – Ritter (1984), Ritter (1991), and Ofek and Richardson (2003) among others document a negative relation between IPO underpricing and long-term post-IPO returns. A possible reason proposed to explain this phenomenon in the IPO setting is the combination of winner's curse and short-sale constraints during post-IPO trading (e.g., Ljungqvist et al. (2006)). This explanation seems even more plausible in the ICO setting, as tokens are impossible to short, and the degree of valuation uncertainty, exacerbating the winner's curse, is significantly larger for ICO-financed projects than for a typical stock.

Post-ICO cumulative returns are decreasing in ICO size, as measured by the amount raised during ICO. Using the results in columns 1 and 5, we find that a one-standard-deviation increase in log amount raised reduces the 30-day log return by 15 percentage points and the 180-day log return by 11 percentage points. This finding is in contrast with the positive relation between IPO proceeds and long-term post-IPO returns (e.g., Brav and Gompers (1997), Carter et al. (1998), and Teoh et al. (1998)). A possible reason for the contrasting finding is that the positive relation in the case os IPOs may be driven by more reputable banks underwriting larger and more successful IPOs (e.g., Chemmanur and Fulghieri (1994)). In the ICO setting, on the contrary, in the absence of underwriter certification, raising a large amount in an ICO may be a sign of ICO investors overpaying for a token, leading to correction once it starts trading.

Tokens that are traded on multiple exchanges enjoy higher long-term returns, consistent with both the exchanges being willing to trade tokens of successful projects and with projects that executed a successful ICO being more willing to pay listing fees to multiple exchanges. A token listed on an additional exchange has a 30-day (180-day) log return that is 6% (14%) higher, ceteris paribus. Finally, longer-term post-ICO returns tend to be positively associated with contemporaneous changes in the number of crypto wallets holding a token.

4.6 Liquidity and return volatility

We proceed to examine the liquidity of traded tokens, defined as minus the log of Amihud illiquidity measure (in Table 13) and the volatility of their returns, defined as daily return standard deviation (in Table 14). Both tables have four sections, corresponding to various return horizons. Each section has two columns, similar to Table 12.

ICOs with higher end-to-open returns have higher liquidity at all horizons and within all samples. This result is consistent with the empirical evidence of a positive relation between IPO underpricing and subsequent liquidity, documented by Hahn, Ligon, and Rhodes (2013). The economic effect is also large: A 1% increase in log end-to-open return increases the 30-day (180-day) liquidity measure by 0.6% (0.7%). On the other hand, first-day return does not tend to be significantly related to token liquidity at horizons longer than 30 days.

Not surprisingly, and consistent with Howell et al. (2018), tokens issued in larger ICOs are more liquid, as are tokens that trade on multiple exchanges. A 1% increase in log amount raised is associated with about 1% increase in liquidity over various horizons. This finding is consistent with IPO evidence: Hahn et al. (2013) report that IPO size is positively associated with post-ICO share liquidity.

Token liquidity is positively associated with contemporaneous social media activity over all horizons, the relation being significant in most cases – a result in line with Bourveau et al. (2018). This result is parallel to equity-based evidence in Blankespoor et al. (2014), who report that companies with an active Twitter account tend to have lower bid-ask spreads.

However, augmenting the liquidity regressions by contemporaneous change in the number of wallets holding the token tends to reduce the coefficient on contemporaneous social media activity and makes it insignificant over most horizons. This suggests that social media activity related to ICO-funded projects increases investors' interest in tokens, which manifests itself in wider investor participation and higher liquidity. Token liquidity is indeed positively and significantly associated with the number of wallets holding the token and with contemporaneous change in it in all samples and over all horizons. For example, a 1% increase in log number of wallets on the first trading day increases the 30-day (180-day) liquidity measure by 0.27-0.55%. This result is consistent with Naes et al. (2011), who show a positive association between investor participation and stock market liquidity.

We now turn to examining the determinants of token return volatility. Token volatility is positively associated with the volatility of ether, at least at shorter horizons. ICO end-to-open return has a negative and significant effect on post-ICO volatility. This result is at odds with the theory of Beatty and Ritter (1986) in IPO setting, who link larger ex-ante uncertainty in IPO valuation, manifesting in higher underpricing, to higher IPO return and a higher post-IPO volatility. Large ICOs, as measured by the amount raised, have lower token return volatility. This finding is reminiscent of the negative relation between firm size and equity return volatility. Tokens of projects in more advanced stages of development (i.e. those with more source commits at the start of trading) tend to have less volatile returns.

5 Conclusions

In this paper, we extend the empirical ICO literature by providing one of the most comprehensive empirical analyses of Initial Coin Offerings. ICO data is generally of low quality, therefore a significant portion of our paper deals with ways to characterize data quality. We propose a data quality measure both at the level of data source-variable pair and at the level of a particular ICO. This measure allows us to focus our empirical investigation on subsamples of ICOs characterized by the highest data quality, thus mitigating concerns about wrong inference due to poor data quality.

We proceed to examine determinants of ICO success, post-ICOs token returns, and longer-term returns, volatility and liquidity, as well as the evolution of project-related social media activity and projects' output around ICOs. In our empirical analysis, we draw parallels with the vast IPO literature. Most of our results for ICOs – such as regarding determinants of ICO success, ICO underpricing and longer-term post-ICO cumulative returns – are broadly consistent with empirical regularities known to characterize IPOs and with theories that were developed to explain regularities in the IPO market. However, some results for ICOs are different from the corresponding findings for IPOs. We argue that some of the discrepancies may be due to differences in institutional settings between the ICO and IPO markets.

Overall, our results contribute to the debate about whether tokens issued in an ICO should be considered securities, a crucial question for designing optimal ICO regulation. Our evidence that tokens tend to behave similar to equities supports the view that they should be treated as securities.

References

Adhami, S., G. Giudici, and S. Martinazzi (2018). Why do businesses go crypto? an empirical analysis of initial coin offerings. *Journal of Economics and Business Forthcoming*.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets 5*(1), 31–56.

Amsden, R. and D. Schweizer (2018). Are blockchain crowdsales the new "gold rush"? success determinants of initial coin offerings. *McGill University Working Paper*.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics 61*(1), 43–76.

Beatty, R. P. and J. R. Ritter (1986). Investment banking, reputation, and the underpricing of initial public offerings. *Journal of Financial Economics 15*, 213–232.

Benedetti, H. and L. Kostovetsky (2018). Digital tulips? returns to investors in initial coin offerings. *Boston College Working Paper*.

Benveniste, L. M. and P. A. Spindt (1989). How investment bankers determine the offer price and allocation of new issues. *Journal of Financial Economics 24*(2), 343–361.

Blankespoor, E., G. S. Miller, and H. D. White (2014). The role of dissemination in market liquidity: Evidence from firms' use of twitter. *Accounting Review 89*(1), 79–112.

Boreiko, D. and N. Sahdev (2018). To ico or not to ico – empirical analysis of initial coin offerings and token sales. *Massachusetts Institute of Technology Working Paper*.

Bourveau, T., E. T. De George, A. Ellahie, and D. Macciocchi (2018). Initial coin offerings: Early evidence on the role of disclosure in the unregulated crypto market. *Columbia University Working Paper*.

Bradley, D. J. and B. D. Jordan (2002). Partial adjustment to public information and ipo underpricing. *Journal of Financial and Quantitative Analysis 37*(4), 595–616.

Brav, A., C. Geczy, and P. A. Gompers (2000). Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics 56*(2), 209–249.

Brav, A. and P. A. Gompers (1997). Myth or reality? the long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies. *Journal of Finance 52*, 1791–1821.

Busaba, W. Y., L. M. Benveniste, and R.-J. Guo (2001). The option to withdraw ipos during the premarket: empirical analysis. *Journal of Financial Economics 60*, 73–102.

Carter, R. B., F. H. Dark, and A. K. Singh (1998). Underwriter reputation, initial returns, and the long-run performance of ipo stocks. *Journal of Finance 53*, 285–311.

Catalini, C. and J. Gans (2018). Initial coin offerings and the value of crypto tokens. *Massachusetts Institute of Technology Working Paper*.

Chemmanur, T. J. and P. Fulghieri (1994). Investment bank reputation, information production, and financial intermediation. *Journal of Finance 49*, 57–79.

Chemmanur, T. J., S. He, and D. K. Nandy (2008). The going public decision and the product market. *Review of Financial Studies 23*, 1855–1908.

Chod, J. and E. Lyandres (2018). A theory of icos: Diversification, agency, and information asymmetry. *Boston University Working Paper*.

Cong, L., Y. Li, and N. Wang (2018). Tokenomics: Dynamic adoption and valuation. *University of Chicago Working Paper*.

Conti, A., M. Thursby, and F. T. Rothaermel (2013). Show me the right stuff: Signals for high-tech startups. *Journal of Economics and Management Strategy 22*(2), 341–364.

Davvydiuk, T., D. Gupta, and S. Rosen (2018). De-crypto-ing signals in initial coin offerings: Evidence of rational token retention. *Carnegie Mellon University Working Paper*.

de Jong, A., P. Roosenboom, and T. van der Kolk (2018). What determines success in initial coin offerings? *Monash University Working Paper*.

Dunbar, C. G. (1998). The choice between firm-commitment and best-efforts offering methods in ipos: The effect of unsuccessful offers. *Journal of Financial Intermediation 7*, 60–90.

Dunbar, C. G. and S. R. Foerster (2008). Second time lucky? withdrawn ipos that return to the market. *Journal of Financial Economics 87*, 610–635.

Eckbo, E. B., R. W. Masulis, and P. A. Gompers (2000). Seasoned public offerings: Resolution of the 'new issues puzzle'. *Journal of Financial Economics 56*(2), 251–291.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

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

Florysiak, D. and A. Schandlbauer (2018). The information content of ico white papers. *University of Southern Denmark Working Paper*.

Hahn, T., J. A. Ligon, and H. Rhodes (2013). Liquidity and initial public offering underpricing. *Journal of Banking and Finance 37*(12), 4973–4988.

Hanley, K. W. (1993). The underpricing of initial public offerings and the partial adjustment phenomenon. *Journal of Financial Economics 34*, 231–250.

Harris, L. and A. Amato (2018). Illiquidity and stock returns: Cross-section and time-series effects: A replication. *Critical Finance Review Forthcoming*.

Howell, S. T., M. Neissner, and D. Yermack (2018). Initial coin offerings: Financing growth with cryptocurrency token sales. *New York University Working Paper*.

Hu, A., C. Parlour, and U. Rajan (2018). Cryptocurrencies: Stylized facts on a new investible instrument. *University of Michigan Working Paper*.

Huang, W., M. Meoli, and S. Vismara (2018). The geography of initial coin offerings. *University of Bergamo Working Paper*.

Jain, B. A. and O. Kini (1994). The post-issue operating performance of ipo firms. *Journal of Finance 49*, 1699–1726.

Kaal, W. A. (2018). Initial coin offerings: The top 25 jurisdictions and their comparative regulatory responses. *University of St. Thomas Working Paper*.

Lee, J., T. Li, and D. Shin (2018). The wisdom of crowds and information cascades in fintech: Evidence from initial coin offerings. *University of Florida Working Paper*.

Leland, H. E. and D. H. Pyle (1977). Informational asymetries, financial structure, and financial intermediation. *Journal of Finance 32*(2), 371–387.

Li, J. and W. Mann (2018). Initial coin offerings and platform building. *UCLA Working Paper*.

Ljungqvist, A., V. Nanda, and R. Singh (2006). Hot markets, investor sentiment, and ipo pricing. *Journal of Business 79*(4), 1667–1702.

Loughran, T. and B. McDonald (2013). Ipo first-day returns, offer price revisions, volatility, and form s-1 language. *Journal of Financial Economics 109*(2), 307–326.

Loughran, T. and J. R. Ritter (1995). The new issues puzzle. *Journal of Finance 50*(1), 23–51.

Loughran, T. and J. R. Ritter (2004). Why has ipo underpricing changed over time? *Financial Management 1*(1), 5–37.

Lyandres, E., L. Sun, and L. Zhang (2008). The new issues puzzle: Testing the investment-based explanation. *Review of Financial Studies 21*(6), 2825–2855.

Malinova, K. and A. Park (2018). Tokenomics: When tokens beat equity. *University of Toronto Working Paper*.

Megginson, W. L. and K. A. Weiss (1991). Venture capitalist certification in initial public offerings. *Journal of Finance 46*, 879–903.

Michaely, R. and W. H. Shaw (1994). The pricing of initial public offerings: Tests of adverse-selection and signaling theories. *Review of Financial Studies 7*, 279–319.

Miller, M. H. and K. Rock (1985). Dividend policy under asymmetric information. *Journal of Finance 40*(4), 1031–1051.

Momtaz, P. P. (2018a). Initial coin offerings. *UCLA Working Paper*.

Momtaz, P. P. (2018b). Putting numbers on the coins: The pricing and performance of initial coin offerings. *UCLA Working Paper*.

Naes, R., J. A. Skjeltor, and B. A. Odegaard (2011). Stock market liquidity and the business cycle. *Journal of Finance 66*(1), 139–176.

Ofek, E. and M. Richardson (2003). Dotcom mania: The rise and fall of internet stock prices. *Journal of Finance 58*, 1113–1137.

Ritter, H. and I. Welch (2002). A review of ipo activity, pricing, and allocations. *Journal of Finance 57*, 1795–1828.

Ritter, J. R. (1984). The ”hot issue” market of 1980. *Journal of Business 57*(2), 215–240.

Ritter, J. R. (1991). The long-run performance of initial public offerings. *Journal of Finance 46*, 3–27.

Scholes, M. S. (1972). The market for securities: Substitution versus price pressure and the effects of information on share prices. *Journal of Business 45*(2), 179–211.

Sockin, M. and W. Xiong (2018). A model of cryptocurrencies. *Princeton University Working Paper*.

Teoh, S. H., I. Welch, and T. J. Wong (1998). Earnings management and the long-run market performance of initial public offerings. *Journal of Finance 53*, 1935–1974.

Welch, I. (1992). Sequential sales, learning, and cascades. *Journal of Finance 47*(2), 695–732.

Figure 1. ICOs over time. This figure reports monthly values of the number of ICOs that raise at least $U.S. 10,000 (left axis) and the total amount raised across all ICOs each month ($ billions, right axis). Monthly observations go from August 2016 to November 2018. The observations reported for the month of August 2016 group all 33 ICOs up to August 2016.

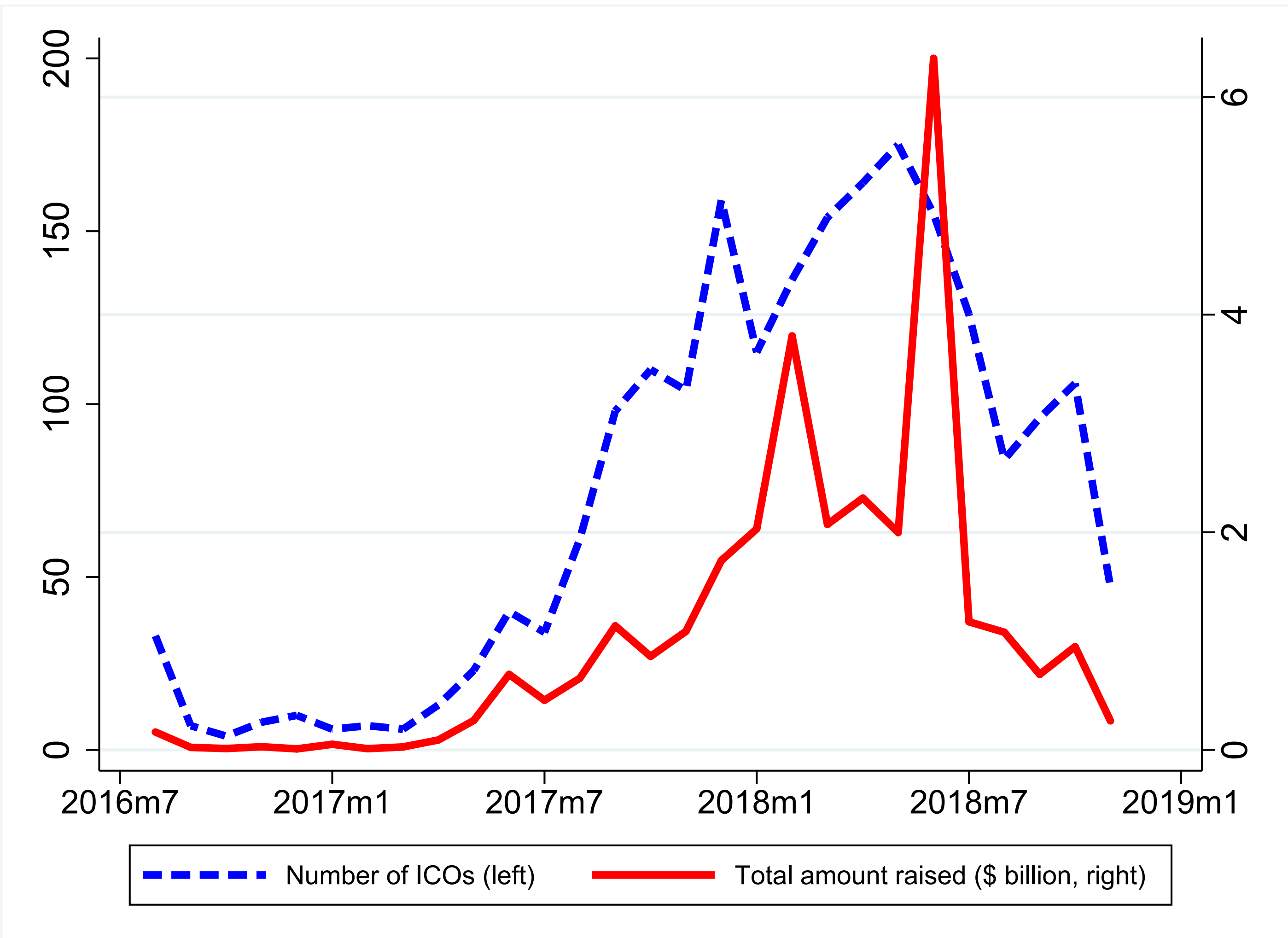

Figure 2. ICOs quality. This figure reports the overall ICO data quality in our dataset as a function of the total number of sources and average consistency across sources for each ICO. Darker points refer to low ICO data quality, while lighter points refer to high ICO data quality.

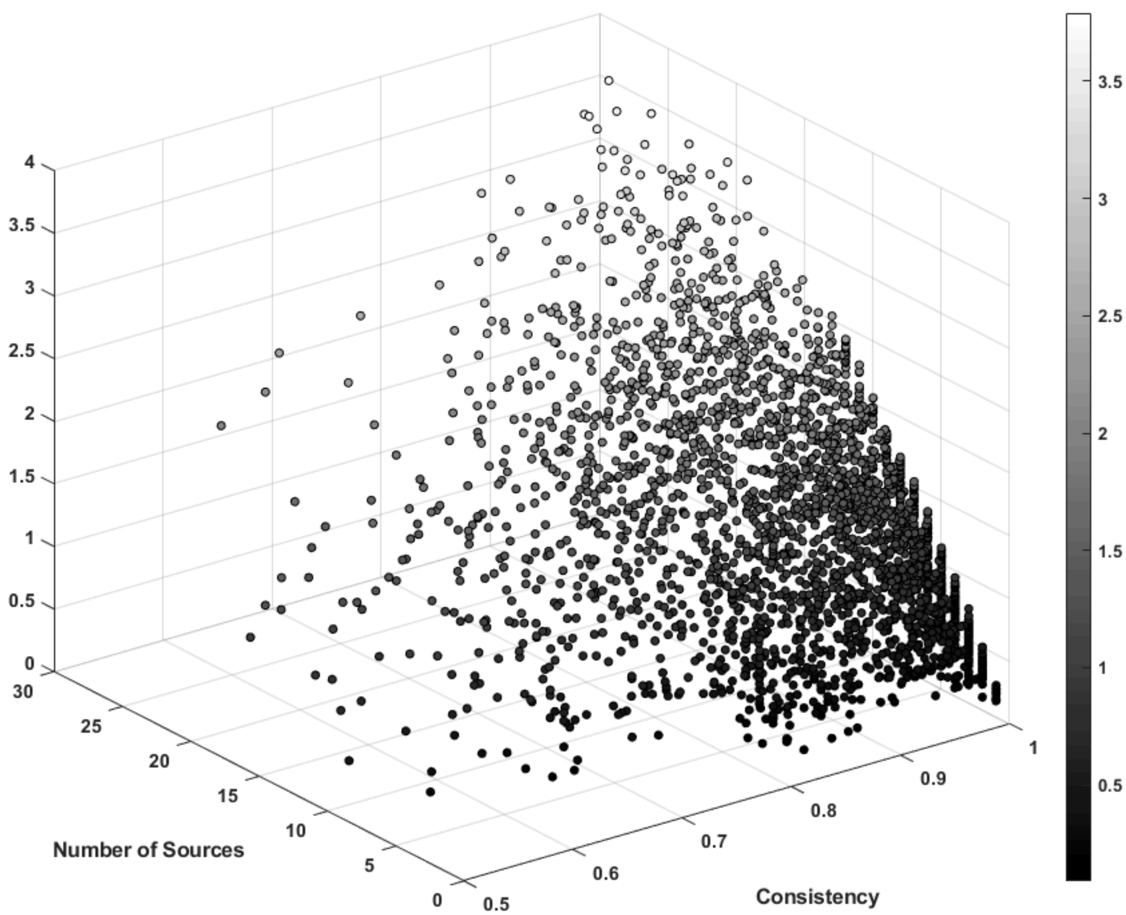

Figure 3. Evolution of social media activity. This figure reports the growth in cumulative social media activity by quarter, starting from 4 quarters before ICO end and ending 4 quarters after ICO end. Growth in quarter t is defined as the percentage difference between the cumulative social media activity at the end of quarter t and that at the end of quarter $t-1$. Every quarter (t) we compute changes for firms with positive cumulative social media activity at ends of both quarters t and $t-1$ and report the median change. Solid black curve refers to all ICOs. Dashed-dotted red line refers to ICOs that managed to raise at least \$U.S. 10,000 in their ICOs. Dashed blue line refers to ICOs that did not manage to raise at least \$U.S. 10,000 in their ICOs.

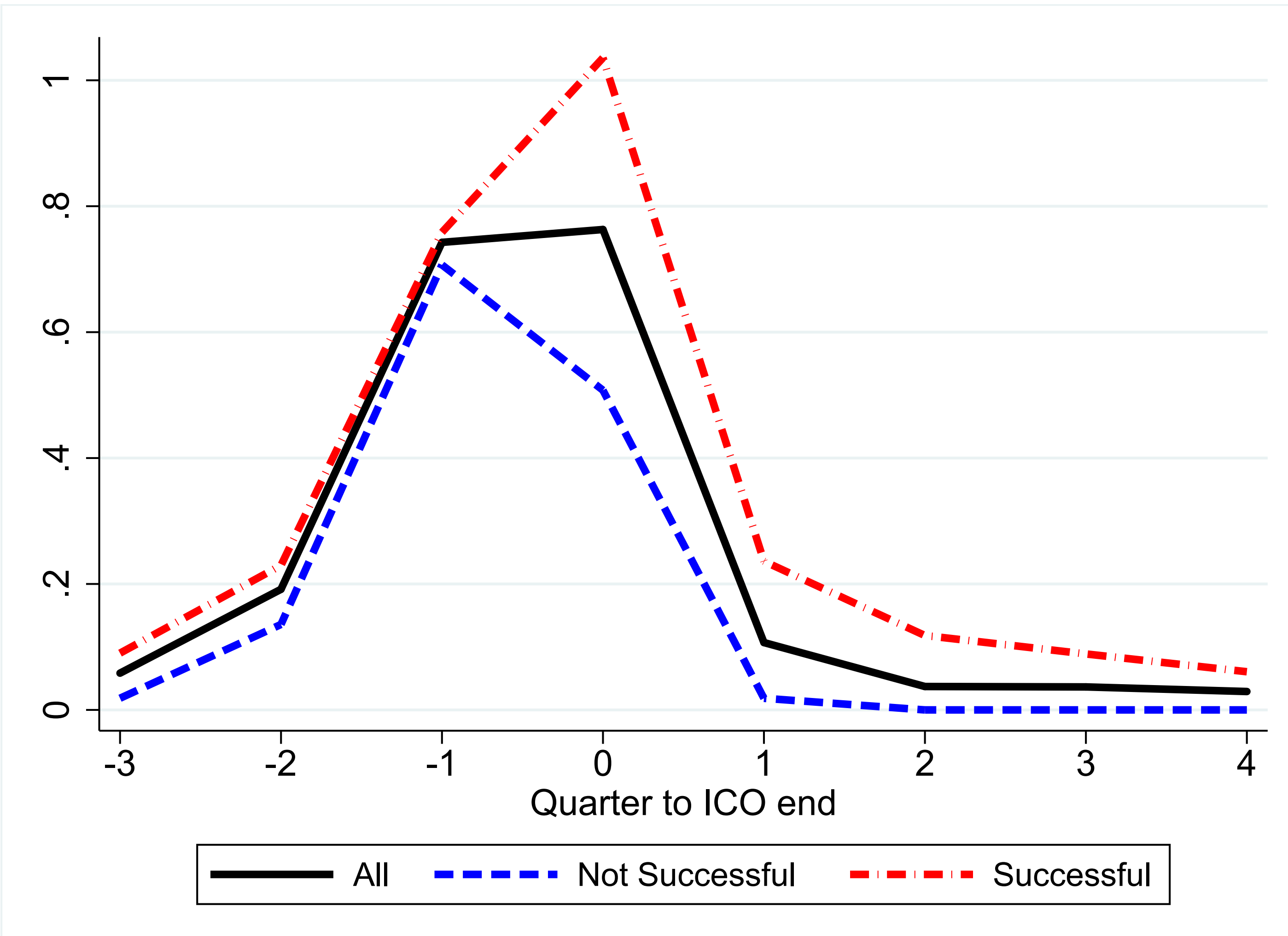

Figure 4. Evolution of commits. This figure reports the growth in cumulative source commits by quarter, starting from 4 quarters before ICO end and ending 4 quarters after ICO end. Growth in quarter t is defined as the percentage difference between the cumulative source commits at the end of quarter t and that at the end of quarter $t-1$. Every quarter (t) we compute changes for firms with positive cumulative commits at ends of both quarters t and $t-1$ and report the median change. Solid black curve refers to all ICOs. Dashed-dotted red line refers to ICOs that managed to raise at least \$U.S. 10,000 in their ICOs. Dashed blue line refers to ICOs that did not manage to raise at least \$U.S. 10,000 in their ICOs.

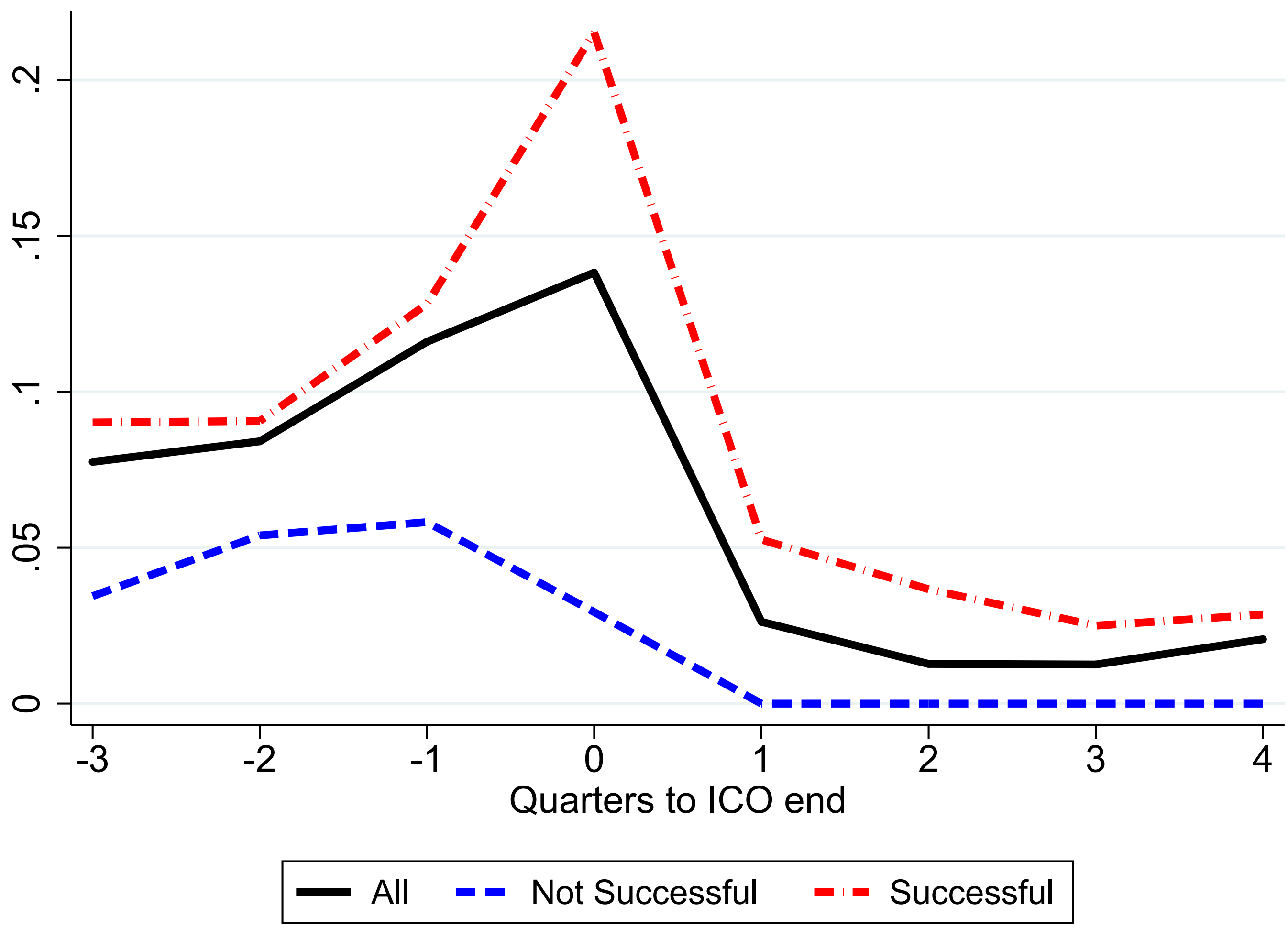

Table 1. Variable definitions. This table lists the variables used in the empirical analysis. For each variable we report the name and a brief description.

Name	Type	Description
amount raised	continuous	total amount raised prior to ICO (in all currencies, converted to $U.S.)
hardcap	continuous	maximum amount allowed to be raised (in all currencies, converted to $U.S.)
token supply	continuous	total amount of tokens that can be issued according to the smart contract
tokens for sale	continuous	total amount for tokens released for the crowd sale
% for sale	continuous	ratio of the total supply of tokens released in a crowd sale over the total supply of tokens that may be issued (token supply)
presale	indicator	whether the project has information on the amount raised in a presale
whitelist	indicator	whether the project offers whitelist for early investors
kyc	indicator	whether the project complies with "know your customer" requirement
team size	integer	total number of team members verified using LinkedIn, Twitter or Facebook accounts
white paper	indicator	whether the project has a white paper associated with it
industry	indicator	whether the project has information about industry
finance	indicator	whether the project is in the finance sector
other software	indicator	whether the project is in the other software sector
business services	indicator	whether the project is in the business services sector
entertainment	indicator	whether the project is in the entertainment sector
blockchain	indicator	whether the project is in the blockchain sector
location	indicator	whether the project has information about location
West. Europe, Can., Austr.	indicator	whether the project is located in Western Europe, Canada, or Australia
Eastern Europe	indicator	whether the project is located in Eastern Europe
Asia	indicator	whether the project is located in Asia
USA	indicator	whether the project is in the USA
rest of the world	indicator	whether the project is in located in a different location from the ones listed above
#nlp words	integer	words count after natural language treatment
tech ratio	continuous	amount of unique tech words over the total unique words in the white paper
raised dummy	indicator	whether the project raised more than 5% of hardcap or more than $10,000 if hardcap is missing
raised-to-hardcap	continuous	ratio of the amount raised in the ICO to the hardcap
ICO length	integer	number of days between ICO end date and ICO start date
listing	indicator	whether the project is listed on at least one exchange
# exchanges	integer	number of exchanges over on which the token/coin is traded
disaster	continuous	whether the token is delisted within one year of listing or has return lower than -95% one year after listing
ICO quality	continuous	ICO data quality measure

Name	Type	Description
Twitter	continuous	a measure of cumulative Twitter activity by day t, computed as # tweets $\times$ I(tweets) + # replies $\times$ I(replies) + # retweets $\times$ I(retweets) + # likes $\times$I(likes), where # (V) corresponds to the cumulative number of V by day t, and I(V) corresponds to the inverse of occurrence of V in the full sample of all ICOs and all dates relative to the most infrequent variable, which is articles on Medium. I(V) takes the following values: I(tweets)= 0.146, I(replies)=0.015, I(retweets)=0.0026, and I(likes)=0.0018
Reddit	continuous	a measure of cumulative Reddit activity by day t, computed as # posts $\times$ I(posts) + # thumbs $\times$ I(thumbs) + # comments $\times$ I(comments), where # (V) corresponds to the cumulative number of V by day t, and I(V) corresponds to the inverse of occurrence of V in the full sample of all ICOs and all dates relative to the most infrequent variable, which is articles on Medium. I(V) takes the following values: I(posts)= 0.323, I(thumbs)=0.028, and I(comments)=0.061
Medium	continuous	a measure of cumulative Medium activity by day t, computed as # articles $\times$I(articles) + # claps $\times$ I(claps) + # comments $\times$ I(comments), where # (V) corresponds to the cumulative number of V by day t, and I(V) corresponds to the inverse of occurrence of V in the full sample of all ICOs and all dates relative to the most infrequent variable, which is articles on Medium. I(V) takes the following values: I(articles)= 1, I(claps)=0.0048, and I(comments)=0.092
BitcoinTalk	continuous	a measure of cumulative BitcoinTalk activity by day t, computed as # posts $\times$I(posts) + # merits $\times$ I(merits), where # (V) corresponds to the cumulative number of V by day t, and I(V) corresponds to the inverse of occurrence of V in the full sample of all ICOs and all dates relative to the most infrequent variable, which is articles on BitcoinTalk. I(V) takes the following values: I(posts)=0.017, I(merits)=0.00006
social media	continuous	Twitter+Reddit+Medium+BitcoinTalk
social media growth (contemp.)	continuous	change in log social media corresponding to the time period over which the dependent variable is measured.
total commits	continuous	total commits in Github main and other repositories by day t
source commits	continuous	total commits in Github main repository by day t
feature commits	continuous	total commits in Github repositories other than the main one by day t
total (source, feature) commits growth	continuous	change in log total (source, feature) commits corresponding to the time period over which the dependent variable is measured
Github feature commits	continuous	total commits in Github repositories other than the main one by day t
# wallets	continuous	number of cryptographic wallets containing at least one token at time t
wallets growth (contemp.)	continuous	change in log # wallets corresponding to the time period over which the dependent variable is measured
ICO end-to-open return	continuous	log difference between the token value at opening during the first trading day and the token value at the end of ICO
ICO first day return	continuous	log difference between the token value at closing and the token price at opening during the first trading day
t-day return (log)	continuous	log difference between the token value at closing after t days and the token value at closing in the first trading day
t-day mean liquidity	continuous	we first compute the average Amihud illiquidity measure during the first t days after the first trading day; then we take the negative of the log value
t-day return volatility	continuous	daily return volatility during the first t days after the first trading day

Table 2. Summary of data sources. This table provides a summary of the data sources used in the paper. In Panel A we list each source and the number of projects covered by it. An ICO is considered to be covered by a source if information is available on the value of at least one of the four variables: amount raised, hardcap, token supply, and tokens for sale. We also provide a description of the type of data and main variables that we extract from each source. In Panel B, we report the distribution of projects across the 10 ICO data sources. 1 means a project that is covered by just one source, 2 means a project that is covered by two sources, and so on.

Panel A: ICO data sources

Source	Type	Projects	Main variables
www.etherscan.io	ico	814	listed exchanges, circulating supply.
www.coindesk.com	ico	841	raised, project name, ICO end date, cumulative funding.
www.coingecko.com	ico	2,326	raised, hardcap, token supply, location.
www.cryptocompare.com	ico	821	raised, hardcap, location, tokens for sale.
www.icobench.com	ico	2,723	raised, hardcap, location, ratings.
www.icodrops.com	ico	488	raised, hardcap, location, token supply.
www.icorating.com	ico	2,965	raised, hardcap, location, ratings,
www.icomarks.io	ico	2,467	raised, hardcap, location, token supply.
www.icodata.io	ico	1,810	raised, hardcap, location, token supply.
www.foundico.com	ico	1,903	raised, hardcap, location, industry, rating.
www.coinmarketcap.com	price	3,401	open, high, low, close, volume, market cap.
www.etherscan.io	transactions	814	wallet address, transaction value.
www.github.com	source code	1,437	commits, source commits, feature commits.
www.twitter.com	social media	3,621	tweets, replies, retweets, likes.
www.reddit.com	social media	1,923	posts, thumbs, comments.
www.medium.com	social media	2,372	articles, claps, comments.
www.bitcointalk.org	social media	3,685	posts, merits.
white papers	white paper contents	1,610	nlp word count, tech ratio.

Panel B: Distribution of projects across ICO data sources

Number of Matches	1	2	3	4	5	6	7	8	9	10
Number of Projects	1,642	989	827	704	498	353	289	131	57	5
Percent of Projects	30%	18%	15%	13%	9%	6%	5%	2%	1%	0%

Table 3. Distribution of ICOs across data sources. For each of the 10 ICO data sources, this table reports the number of observations for four key variables of the ICO process: total amount raised, maximum amount allowed to be raised (hardcap), total supply of tokens that may be issued (token supply), and total supply of tokens released for crowd sale (tokens for sale). For each of the four variables, we also report the mean deviation from the consensus (average) value at the source level. For variable j in source i, the deviation from the average value is measured as the absolute value of the difference between the variable value $x_{j,i}$ and the average value across all sources reporting values for this variable, $\bar{x}_i$, divided by the sum of the two values, namely $\left|\frac{x_{j,i}-\bar{x}_i}{x_{j,i}+\bar{x}_i}\right|$.

Source	Amount raised		Hardcap		Token supply		Tokens for sale	
	Obs	Mean Deviation	Obs	Mean Deviation	Obs	Mean Deviation	Obs	Mean Deviation
www.etherscan.io	242	0.068	0	NA	811	0.128	0	NA
www.coindesk.com	813	0.085	0	NA	0	NA	0	NA
www.coingecko.com	1,033	0.079	1,440	0.103	1,332	0.087	948	0.141
www.cryptocompare.com	525	0.132	562	0.129	627	0.147	605	0.181
www.icobench.com	1,340	0.093	1,623	0.089	1,704	0.063	2,112	0.088
www.icodrops.com	447	0.074	415	0.083	393	0.112	365	0.126
www.icorating.com	1,012	0.115	1,438	0.106	0	NA	1,800	0.147
www.icomarks.io	0	NA	811	0.060	1,609	0.053	1,897	0.081
www.icodata.io	1803	0.322	1803	0.314	373	0.313	0	NA
www.foundico.com	0	NA	676	0.130	0	NA	1,600	0.162

Table 4. Bancor data quality. This table reports the calculation of our data quality measure for the ICO of Bancor. Panel A reports available data on amount raised, hardcap, token supply, and tokens for sale across the 10 ICO data sources. For each variable, we report the average value of the variable across sources and consistency, defined as 1 minus the mean deviation from the average value. The latter value is measured for each variable as the absolute value of the difference between the variable value x_i and the average value $\bar{x}$, divided by the sum of the two values, namely $\left|\frac{x_i - \bar{x}}{x_i + \bar{x}}\right|$. Panel B reports for each available data source the quality of that source for that variable, computed as the inverse of that data source's mean deviation, reported in Table 3, divided by the largest value across data sources. For each variable we report the sum of the source quality values across all four variables above. The adjusted quality value for each variable is given by the product of the sum of source quality values and corresponding consistency value, reported in Panel A. The ICO data quality is the simple average of the adjusted quality values.

Panel A: Available data

Source	Amount raised	Hardcap	Token supply	Tokens for Sale
www.etherscan.io			77,566,371	
www.coindesk.com	153,000,000			
www.coingecko.com			75,783,855	
www.cryptocompare.com	153,000,000	36,000,000	79,320,000	39,660,000
www.icobench.com	153,000,000			
www.icodrops.com	153,000,000			
www.icorating.com	153,000,000			
www.icomarks.io				
www.icodata.io	153,000,000	0	56,889,807	
www.foundico.com				
Average	153,000,000	36,000,000	72,390,008	39,660,000
Consistency	1.000	0.333	0.944	1.000

Panel B: Source quality (Bancor example)

Source	Amount raised	Hardcap	Token supply	Tokens for sale
www.etherscan.io			0.414	
www.coindesk.com	0.800			
www.coingecko.com			0.609	
www.cryptocompare.com	0.515	0.465	0.361	0.448
www.icobench.com	0.731			
www.icodrops.com	0.919			
www.icorating.com	0.591			
www.icomarks.io				
www.icodata.io	0.211	0.191	0.169	
www.foundico.com				
Sum of source qualities	3.768	0.656	1.553	0.448
Adjusted quality	3.768	0.219	1.466	0.448
ICO data quality	1.475			

Table 5. Summary statistics. This table reports the mean, standard deviation, minimum value, median value, maximum value, and number of observations for the variables used in the empirical analysis. Variables are described in Table 1.

Panel A: ICO Variables

	Mean	Std. Dev.	Min	Median	Max	Obs.
ICO characteristics						
hardcap	69.75	989.41	0.00	20.00	48,093.76	3,207
hardcap (log)	2.81	1.38	-6.32	3.00	10.78	3,207
% for sale	0.56	0.24	0.00	0.57	1.00	2,815
presale	0.04	0.20	0.00	0.00	1.00	4,411
whitelist	0.30	0.46	0.00	0.00	1.00	4,411
kyc	0.49	0.50	0.00	0.00	1.00	4,411
team size	10.67	8.14	1.00	9.00	74.00	3,230
team size (log)	2.21	0.74	0.69	2.30	4.32	3,230
Industry						
industry	0.52	0.50	0.00	1.00	1.00	4,411
finance	0.31	0.46	0.00	0.00	1.00	2,310
business services	0.22	0.41	0.00	0.00	1.00	2,310
entertainment	0.22	0.42	0.00	0.00	1.00	2,310
other software	0.15	0.36	0.00	0.00	1.00	2,310
blockchain	0.09	0.29	0.00	0.00	1.00	2,310
Location						
location	0.79	0.41	0.00	1.00	1.00	4,411
West. Europe, Can., Austr.	0.31	0.46	0.00	0.00	1.00	3,469
Eastern Europe	0.20	0.40	0.00	0.00	1.00	3,469
Asia	0.19	0.39	0.00	0.00	1.00	3,469
USA	0.15	0.36	0.00	0.00	1.00	3,469
rest of the world	0.15	0.35	0.00	0.00	1.00	3,469
crypto friendly	0.31	0.46	0.00	0.00	1.00	3,469
White paper characteristics						
white paper	0.26	0.44	0.00	0.00	1.00	4,411
# nlp words	1,672.79	751.96	93.00	1,561.50	5,062.00	1,136
# nlp words (log)	7.31	0.49	4.54	7.35	8.53	1,136
tech ratio	0.29	0.06	0.04	0.28	0.65	1,136
ICO Outcomes						
raised dummy	0.45	0.50	0.00	0.00	1.00	4,411
amount raised	14.70	103.96	0.00	4.57	4,197.96	2,040
amount raised (log)	1.08	2.22	-13.82	1.52	8.34	2,040
raised-to-hardcap	0.44	0.39	0.00	0.30	1.00	1,711
ICO length	47.42	39.67	0.00	31.00	575.00	4,411
listing	0.39	0.49	0.00	0.00	1.00	2,040
# exchanges	5.39	8.31	1.00	3.00	100.00	764
ICO data quality						
ICO data quality	1.02	0.73	0.09	0.89	3.79	4,411

Panel B: Social media

	Mean	Std. Dev.	Min	Median	Max	Obs.
Total (log)						
90 days before ICO end (log)	1.40	1.76	0.00	0.29	11.20	4,406
at ICO start (log)	2.23	1.85	0.00	2.14	11.20	4,408
at ICO end (log)	2.67	1.91	0.00	2.80	11.20	4,408
90 days after ICO end (log)	2.89	1.97	0.00	3.07	11.20	3,832
90 days before first trade (log)	2.15	2.01	0.00	2.00	11.20	876
at first trade (log)	3.60	1.79	0.00	3.91	11.20	878
90 after first trade (log)	4.09	1.77	0.00	4.46	11.20	805
Twitter (log)						
90 days before ICO end (log)	0.78	1.45	0.00	0.00	7.77	4,406
at ICO start (log)	1.26	1.71	0.00	0.00	7.89	4,408
at ICO end (log)	1.59	1.86	0.00	0.76	8.07	4,408
90 days after ICO end (log)	1.78	1.94	0.00	1.21	9.76	3,832
90 days before first trade (log)	1.19	1.80	0.00	0.00	7.61	876
at first trade (log)	1.99	2.16	0.00	0.96	7.85	878
90 after first trade (log)	2.36	2.27	0.00	1.93	9.82	805
Reddit (log)						
90 days before ICO end (log)	0.22	0.80	0.00	0.00	6.49	4,406
at ICO start (log)	0.35	1.00	0.00	0.00	6.54	4,408
at ICO end (log)	0.48	1.21	0.00	0.00	6.92	4,408
90 days after ICO end (log)	0.56	1.31	0.00	0.00	7.05	3,832
90 days before first trade (log)	0.11	0.54	0.00	0.00	5.86	876
at first trade (log)	0.31	0.95	0.00	0.00	6.44	878
90 after first trade (log)	0.52	1.28	0.00	0.00	6.54	805
Medium (log)						
90 days before ICO end (log)	0.68	1.29	0.00	0.00	11.20	4,406
at ICO start (log)	1.06	1.53	0.00	0.00	11.20	4,408
at ICO end (log)	1.27	1.68	0.00	0.00	11.20	4,408
90 days after ICO end (log)	1.34	1.74	0.00	0.00	11.20	3,832
90 days before first trade (log)	1.01	1.62	0.00	0.00	11.20	876
at first trade (log)	1.62	1.94	0.00	0.00	11.20	878
90 after first trade (log)	1.88	2.09	0.00	0.69	11.20	805
BitcoinTalk (log)						
90 days before ICO end (log)	0.34	0.79	0.00	0.00	4.86	4,406
at ICO start (log)	0.69	1.08	0.00	0.00	5.35	4,408
at ICO end (log)	0.88	1.25	0.00	0.00	5.54	4,408
90 days after ICO end (log)	1.02	1.38	0.00	0.00	6.69	3,832
90 days before first trade (log)	0.88	1.28	0.00	0.00	6.09	876
at first trade (log)	1.88	1.51	0.00	1.88	6.21	878
90 after first trade (log)	2.18	1.62	0.00	2.23	6.72	805

Panel C: Commits

	Mean	Std. Dev.	Min	Median	Max	Obs.
Total commits (log)						
90 days before ICO end (log)	0.99	2.35	0.00	0.00	13.19	4,406
at ICO start (log)	1.20	2.47	0.00	0.00	13.18	4,408
at ICO end (log)	1.26	2.51	0.00	0.00	13.23	4,408
90 days after ICO end (log)	1.39	2.65	0.00	0.00	12.68	3,832
90 days before first trade (log)	3.00	3.54	0.00	1.10	12.67	876
at first trade (log)	3.63	3.51	0.00	3.04	12.68	878
90 after first trade (log)	3.84	3.60	0.00	3.37	12.68	805
Source commits (log)						
90 days before ICO end (log)	0.56	1.72	0.00	0.00	11.62	4,406
at ICO start (log)	0.77	2.12	0.00	0.00	13.18	4,408
at ICO end (log)	0.78	1.96	0.00	0.00	11.67	4,408
90 days after ICO end (log)	0.89	2.14	0.00	0.00	11.71	3,832
90 days before first trade (log)	1.97	2.92	0.00	0.00	11.65	876
at first trade (log)	2.59	3.11	0.00	0.00	11.70	878
90 after first trade (log)	2.85	3.28	0.00	0.69	11.75	805
Feature commits (log)						
90 days before ICO end (log)	0.66	2.04	0.00	0.00	13.19	4,406
at ICO start (log)	0.73	1.91	0.00	0.00	11.66	4,408
at ICO end (log)	0.79	2.15	0.00	0.00	13.23	4,408
90 days after ICO end (log)	0.89	2.28	0.00	0.00	12.64	3,832
90 days before first trade (log)	2.08	3.34	0.00	0.00	12.63	876
at first trade (log)	2.39	3.40	0.00	0.00	12.63	878
90 after first trade (log)	2.55	3.50	0.00	0.00	12.64	805

Panel D: Wallets

	Mean	Std. Dev.	Min	Median	Max	Obs.
			Total Wallets (log)			
at ICO start (log)	4.63	2.29	0.69	4.78	10.35	348
at ICO end (log)	5.51	2.38	0.69	5.93	11.34	450
at first trade (log)	6.73	1.67	1.61	6.80	11.83	637
30 days after first trade (log)	7.28	1.46	1.10	7.33	11.99	636
90 days after first trade (log)	7.66	1.40	1.10	7.71	11.99	590
180 days after first trade (log)	7.93	1.28	3.53	8.01	11.39	463
365 days after first trade (log)	7.88	1.33	3.91	7.93	11.39	215
			Wallet Concentration			
at ICO start (log)	0.48	0.36	0.00	0.44	1.00	348
at ICO end (log)	0.44	0.35	0.00	0.36	1.00	450
at first trade (log)	0.33	0.27	0.00	0.26	1.00	637
30 days after first trade (log)	0.29	0.25	0.00	0.22	1.00	636
90 days after first trade (log)	0.26	0.23	0.00	0.20	1.00	590
180 days after first trade (log)	0.25	0.23	0.00	0.19	1.00	463
365 days after first trade (log)	0.24	0.22	0.01	0.17	0.87	215

Panel E: Returns

	Mean	Std. Dev.	Min	Median	Max	Obs.
ICO end-to-open returns						
ICO return (%)	269.18	405.91	4.23	108.98	1,625.06	801
ICO first-day returns						
First day return (%)	11.85	25.86	-20.86	2.45	86.15	878
Longer-term cumulative post-ICO returns						
30-day return (%)	0.75	89.73	-81.48	-30.51	258.38	835
90-day return (%)	9.09	140.55	-93.43	-46.38	448.70	771
180-day return (%)	32.04	230.04	-97.55	-68.90	796.31	615
365-day return (%)	60.30	295.98	-98.78	-66.12	1,095.33	302
Liquidity						
30-day mean liquidity	12.77	3.39	1.32	13.14	21.43	856
90-day mean liquidity	12.45	3.62	2.38	13.05	21.36	790
180-day mean liquidity	12.42	3.78	2.36	13.10	21.22	623
365-day mean liquidity	11.95	4.31	1.92	12.80	20.38	305
Return volatility						
30-day average vol (%)	18.94	11.37	6.56	15.53	50.72	857
90-day average vol (%)	17.48	10.08	7.61	14.25	46.72	790
180-day average vol (%)	16.28	8.69	7.80	13.32	41.71	623
365-day average vol (%)	17.21	9.36	9.06	13.69	45.90	305

Table 6. ICO success, unconditional analysis. This table reports univariate analysis of determinants of ICO success. We use five definitions of ICO success: (i) dummy variable that takes the value of one if a project raised more than $10,000; (ii) total amount raised; (iii) total amount raised as a fraction of hardcap; (iv) dummy variable that takes value of one if the token is traded on an exchange; (v) dummy variable that takes the value of one if a traded token is delisted from all exchanges within one year or has a one-year return lower than -95%. For dummy variables, the column *Low* (*High*) reports the average value of the success variable when the dummy variable takes value of zero (one). For continuous variables, the column Low (High) reports the average value of the success variable for observations in the top (bottom) quartile. The column Diff reports the difference between the High and Low values, whereas the column tstat reports the corresponding t statistic.

	Raised dummy (N=4,411; mean=0.45)				Amount raised (N=2,040; mean=14.90)				Raised-to-hardcap (N=1,711; mean=0.44)				Listing dummy (N=2,040; mean=0.39)				Disaster dummy (N=346; mean=0.23)			
	Low	High	Diff	tstat	Low	High	Diff	tstat	Low	High	Diff	tstat	Low	High	Diff	tstat	Low	High	Diff	tstat
presale (dummy)	0.43	1.00	0.57	15.80	14.74	16.55	1.82	11.36	0.40	0.76	0.37	12.47	0.35	0.81	0.46	12.65	0.23	0.25	0.02	0.24
whitelist (dummy)	0.43	0.51	0.08	4.95	14.60	15.50	0.90	9.07	0.40	0.50	0.10	5.45	0.39	0.40	0.01	0.33	0.23	0.26	0.03	0.34
kyc (dummy)	0.35	0.55	0.20	13.42	14.22	15.37	1.15	12.27	0.34	0.49	0.14	7.30	0.25	0.49	0.25	11.56	0.29	0.21	-0.08	-1.64
team size	0.43	0.71	0.28	12.21	14.53	15.59	1.06	8.38	0.42	0.48	0.05	1.98	0.38	0.46	0.08	2.65	0.34	0.14	-0.20	-2.88
white paper (dummy)	0.37	0.70	0.33	20.42	14.66	15.28	0.62	6.43	0.39	0.50	0.11	5.79	0.26	0.60	0.34	16.25	0.30	0.18	-0.12	-2.71
# nlp words	0.57	0.76	0.19	4.86	14.29	16.04	1.75	9.51	0.41	0.58	0.17	3.73	0.44	0.70	0.26	5.32	0.24	0.16	-0.09	-0.93
tech ratio	0.61	0.80	0.20	5.01	14.86	15.59	0.72	3.72	0.38	0.58	0.20	4.73	0.44	0.73	0.29	5.97	0.40	0.09	-0.31	-3.83
social media at ICO start (dummy)	0.30	0.50	0.20	11.56	14.65	14.95	0.30	2.31	0.41	0.44	0.03	1.03	0.28	0.41	0.14	4.62	0.31	0.22	-0.09	-1.57
cumulative social media at ICO start	0.35	0.63	0.28	11.86	14.57	15.36	0.79	5.39	0.43	0.48	0.04	1.43	0.35	0.44	0.09	2.60	0.21	0.19	-0.02	-0.29
commits at ICO start (dummy)	0.36	0.72	0.36	22.10	14.59	15.38	0.79	8.28	0.38	0.51	0.13	7.01	0.26	0.60	0.34	16.46	0.34	0.15	-0.19	-4.23
cumulative source commits at ICO start	0.58	0.84	0.26	7.18	14.85	15.81	0.96	5.49	0.36	0.64	0.28	7.07	0.37	0.74	0.37	8.03	0.28	0.09	-0.19	-2.21

Table 7. ICO success, multivariate regressions. This table reports multivariate regressions of determinants of ICO success. We use five definitions of ICO success: We use five definitions of ICO success: (i) dummy variable that takes the value of one if a project raised more than \$10,000; (ii) total amount raised; (iii) total amount raised as a fraction of hardcap; (iv) dummy variable that takes value of one if the token is traded on an exchange; (v) dummy variable that takes the value of one if a traded token is delisted from all exchanges within one year or has a one-year return lower than -95%. For each variable, we run the analysis both on all the available observations and on observations belonging to ICOs with high data quality. An ICO is defined to have high quality data if it belongs to the top tercile of the ICO data quality variable distribution. We use logit regressions when the success variable is the raised dummy or the listing dummy. In this case, the reported coefficient is the marginal effect of the independent variable and the reported R-squared is the pseudo R-squared. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	Raised dummy		Amount raised (log)		Raised-to-hardcap		Listing dummy		Disaster dummy	
	All	High Q	All	High Q	All	High Q	All	High Q	All	High Q
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
hardcap (log)	0.002	0.002	0.705***	0.680***	-0.065***	-0.064***	0.010	0.022*	-0.041	-0.021
	(0.262)	(0.227)	(17.712)	(15.229)	(-7.355)	(-6.066)	(1.046)	(1.864)	(-1.483)	(-0.589)
presale			0.808***	0.750***	0.270***	0.267***	0.168***	0.139***	-0.017	0.027
			(5.389)	(5.130)	(8.181)	(7.753)	(4.648)	(3.751)	(-0.168)	(0.298)
% for sale	0.010	-0.031	-0.588***	-0.616***	-0.158***	-0.164***	-0.092**	-0.094*	-0.101	0.032
	(0.235)	(-0.597)	(-3.060)	(-2.890)	(-3.737)	(-3.272)	(-1.981)	(-1.717)	(-0.803)	(0.225)
whitelist	0.001	-0.011	0.162	0.235**	0.039*	0.046*	-0.010	-0.007	0.175*	0.113
	(0.061)	(-0.481)	(1.539)	(2.106)	(1.676)	(1.770)	(-0.403)	(-0.263)	(1.790)	(1.228)
kyc	0.116***	0.146***	0.551***	0.635***	0.089***	0.095***	0.161***	0.189***	0.149**	-0.000
	(5.199)	(5.552)	(5.010)	(5.126)	(3.660)	(3.268)	(6.262)	(6.358)	(2.081)	(-0.001)
team size (log)	0.108***	0.059***	0.165**	0.179**	0.014	0.007	0.054***	0.051***	-0.142***	-0.104**
	(7.617)	(3.376)	(2.344)	(2.368)	(0.929)	(0.418)	(3.073)	(2.613)	(-3.406)	(-2.250)
white paper	0.084***	0.090***	0.150	0.139	0.040*	0.048**	0.153***	0.162***	-0.036	-0.033
	(3.891)	(3.524)	(1.574)	(1.386)	(1.890)	(2.047)	(7.708)	(7.520)	(-0.603)	(-0.541)
social media at ICO start (log)	0.032***	0.025***	0.024	0.018	0.002	0.001	0.017***	0.016**	-0.035*	-0.033
	(5.782)	(3.735)	(0.885)	(0.633)	(0.282)	(0.194)	(2.650)	(2.239)	(-1.929)	(-1.595)
source commits at ICO start (log)	0.039***	0.029***	0.102***	0.074***	0.024***	0.019***	0.040***	0.040***	-0.030***	-0.031**
	(6.545)	(4.077)	(5.225)	(3.730)	(5.706)	(4.188)	(8.468)	(7.919)	(-2.585)	(-2.321)
West. Europe, Can., Austr.	-0.051*	-0.021	-0.406***	-0.474***	-0.091***	-0.119***	-0.006	-0.003	0.053	0.022
	(-1.683)	(-0.583)	(-2.846)	(-3.109)	(-2.896)	(-3.309)	(-0.175)	(-0.069)	(0.616)	(0.271)
Eastern Europe	-0.031	-0.016	-0.512***	-0.628***	-0.113***	-0.163***	-0.109***	-0.111***	0.001	0.004
	(-0.954)	(-0.416)	(-3.333)	(-3.841)	(-3.328)	(-4.232)	(-2.891)	(-2.683)	(0.011)	(0.043)
Asia	0.099***	0.100**	0.003	-0.070	0.005	-0.032	0.056	0.044	-0.087	-0.039
	(2.941)	(2.416)	(0.019)	(-0.433)	(0.150)	(-0.826)	(1.552)	(1.053)	(-0.914)	(-0.391)
USA	-0.088**	-0.074*	-0.121	-0.163	-0.001	-0.032	-0.008	-0.010	-0.047	-0.164
	(-2.401)	(-1.713)	(-0.707)	(-0.878)	(-0.019)	(-0.721)	(-0.193)	(-0.201)	(-0.501)	(-1.291)
rest of the world	-0.017	-0.030	-0.050	-0.122	-0.009	-0.048	-0.025	-0.032	-0.178	-0.079
	(-0.493)	(-0.746)	(-0.305)	(-0.703)	(-0.255)	(-1.171)	(-0.652)	(-0.742)	(-1.326)	(-0.676)
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,068	1,219	1,268	980	1,268	980	1,243	965	161	128
R-squared	0.230	0.228	0.361	0.387	0.247	0.276	0.374	0.408	0.282	0.264

Table 8. ICO success, white paper available. This table reports multivariate regressions of determinants of ICO success conditional on having a white paper. We use four definitions of ICO success: (i) dummy variable that takes the value of one if a project raised more than $10,000; (ii) total amount raised; (iii) total amount raised as a fraction of hardcap; (iv) dummy variable that takes the value of one if the token is traded on an exchange. For each variable, we run the analysis both on all the available observations and on observations belonging to ICOs with high data quality. An ICO is defined to have high quality data if it belongs to the top tercile of the ICO data quality variable distribution. We use logistic regressions when the success variable is the raised dummy or the listing dummy. In this case, the reported coefficient is the marginal effect of the independent variable and the reported R-squared is the pseudo R-squared. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	Raised dummy		Amount raised		Raised-to-hardcap		Listing dummy	
	All	High Q	All	High Q	All	High Q	All	High Q
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# nlp words (log)	0.054*	0.035	0.578***	0.552***	0.103***	0.122***	0.152***	0.126***
	(1.683)	(1.044)	(3.693)	(3.405)	(2.730)	(2.984)	(3.684)	(2.941)
tech ratio	0.825***	0.567**	2.933**	3.814***	0.821***	1.033***	0.553	0.456
	(2.957)	(2.170)	(2.374)	(3.083)	(2.765)	(3.320)	(1.575)	(1.229)
hardcap (log)	0.008	0.006	0.634***	0.592***	-0.080***	-0.090***	0.013	0.014
	(0.618)	(0.395)	(10.149)	(8.758)	(-5.357)	(-5.277)	(0.859)	(0.834)
presale			0.906***	0.757***	0.311***	0.276***	0.228***	0.175***
			(5.159)	(4.455)	(7.358)	(6.462)	(3.892)	(3.085)
% for sale	-0.045	-0.171***	-0.398	-0.524*	-0.116*	-0.154**	0.018	0.031
	(-0.747)	(-2.687)	(-1.485)	(-1.874)	(-1.805)	(-2.194)	(0.249)	(0.386)
whitelist	-0.014	-0.046	0.136	0.350**	0.021	0.056	-0.003	0.030
	(-0.433)	(-1.633)	(0.893)	(2.234)	(0.575)	(1.431)	(-0.074)	(0.712)
kyc	0.123***	0.100***	0.761***	0.963***	0.142***	0.162***	0.185***	0.177***
	(3.924)	(3.219)	(4.747)	(5.432)	(3.693)	(3.630)	(4.906)	(4.138)
team size (log)	0.069***	0.041**	0.071	-0.006	0.000	-0.014	0.064**	0.065**
	(3.316)	(2.049)	(0.724)	(-0.057)	(0.008)	(-0.557)	(2.378)	(2.334)
social media at ICO start (log)	0.022***	0.008	-0.008	-0.017	0.001	-0.003	0.013	0.009
	(2.780)	(1.050)	(-0.226)	(-0.439)	(0.136)	(-0.322)	(1.302)	(0.821)
source commits at ICO start (log)	0.045***	0.056**	0.083***	0.069***	0.020***	0.017***	0.049***	0.053***
	(4.519)	(2.525)	(3.317)	(2.771)	(3.350)	(2.719)	(6.170)	(6.133)
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	669	446	528	439	528	439	517	430
R-squared	0.296	0.424	0.419	0.457	0.353	0.395	0.368	0.401

Table 9. Social media. This table reports regressions of determinants of the change in cumulative social media activity around ICO end date. The dependent variable is the log difference between cumulative social media activity 90 days after ICO end and cumulative social media at ICO end. We report the results for the change in the cumulative measure of social media activity (Column 1), and for Twitter, Reddit, Medium, and BitcoinTalk activity separately in columns 2-5. In Panel A, we use raised dummy as a measure of ICO success, while in Panel B we use total amount raised. Lagged social media growth is the log difference between cumulative social media activity at ICO end and cumulative social media activity 90 days before ICO end. Lagged social media level is (log) value of cumulative social media activity 90 days before ICO end. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	All	Twitter	Reddit	Medium	Bitcointalk
	(1)	(2)	(3)	(4)	(5)
Panel A: Extensive margin					
raised dummy	0.204***	0.152***	0.011	0.146***	0.055***
	(8.925)	(5.294)	(0.178)	(6.055)	(3.977)
lagged social media growth	-0.010	-0.008	0.026	0.046***	0.047***
	(-0.933)	(-0.579)	(0.806)	(3.123)	(6.000)
lagged social media level	-0.069***	-0.074***	-0.033	-0.034***	-0.012**
	(-9.515)	(-8.561)	(-1.422)	(-3.400)	(-2.056)
Time dummy	Yes	Yes	Yes	Yes	Yes
Observations	1,917	1,212	293	902	882
R-squared	0.112	0.087	0.121	0.101	0.282
Panel B: Intensive margin					
amount raised (log)	0.069***	0.047***	0.054***	0.049***	0.024***
	(9.536)	(5.240)	(2.989)	(6.154)	(6.061)
lagged social media growth	-0.074***	-0.041**	0.029	0.026	0.041***
	(-4.913)	(-2.313)	(0.789)	(1.240)	(4.232)
lagged social media level	-0.116***	-0.095***	-0.010	-0.056***	-0.020**
	(-10.969)	(-8.011)	(-0.378)	(-3.959)	(-2.577)
Time dummy	Yes	Yes	Yes	Yes	Yes
Observations	1,082	647	159	542	618
R-squared	0.181	0.135	0.301	0.133	0.327

Table 10. Commits. This table reports regressions of determinants of the change in cumulative total commits around ICO end date. The dependent variable is log difference between cumulative Github commits 90 days after the ICO end and cumulative Github commits at ICO end. We report the results for the change in the cumulative measure of total commits (Column 1), and for source commits and feature commits separately in columns 2 and 3. In Panel A, we use raised dummy as a measure of ICO success, while in Panel B we use total amount raised. Lagged commits growth is log difference between cumulative commits at ICO end and cumulative commits 90 days before ICO end. Lagged commits level is (log) value of cumulative commits 90 days before ICO end. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	All	Source	Feature
	(1)	(2)	(3)
Panel A: Extensive Margin			
raised dummy	0.065*	0.155**	0.046
	(1.947)	(2.380)	(1.144)
lagged commits growth	0.150***	0.114***	0.149***
	(7.645)	(4.084)	(4.944)
lagged commits level	0.002	-0.019*	0.000
Time dummy	Yes	Yes	Yes
Observations	767	480	492
R-squared	0.105	0.101	0.078
Panel B: Intensive Margin			
amount raised (log)	0.017*	0.048***	0.009
	(1.833)	(2.669)	(0.867)
lagged commits growth	0.146***	0.092***	0.142***
	(6.397)	(2.850)	(4.068)
lagged commits level	-0.000	-0.033**	-0.003
	(-0.020)	(-2.544)	(-0.418)
Time dummy	Yes	Yes	Yes
Observations	587	395	383
R-squared	0.098	0.099	0.070

Table 11. ICO returns. This table reports regressions of determinants of ICO end-to-open return and of ICO first-trading-day return. ICO end-to-open return is the log difference between the token value at opening during the first trading day and the token value at the end of ICO. The latter quantity is calculated by dividing the amount raised by the circulating supply of tokens 7 days after the beginning of trading. ICO first-day return is the the log difference between the closing and opening prices of the first day of trading. For each dependent variable, we report the results for three different samples. The first contains ICOs with information on amount raised. The second contains information on additional explanatory variables. The third contains information on additional explanatory variables and is restricted to observations belonging to ICOs with high data quality. An ICO is defined to have high-quality data if it belongs to the top tercile of the ICO data quality variable distribution.We winsorize the ICO end-to-open return and first-day return at the top and bottom 1%. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	ICO end-to-open return (log)			ICO first-day return (log)		
	(1)	(2)	(3)	(4)	(5)	(6)
ICO end-to-open return (log)				-0.020***	-0.025**	-0.028**
				(-3.502)	(-2.532)	(-2.560)
return bct	0.443**	-0.290	-0.485	0.025	-0.109*	-0.115
	(2.500)	(-0.807)	(-1.079)	(0.907)	(-1.794)	(-1.512)
return eth	-0.046	1.172***	1.163***	0.001	0.046	0.039
	(-1.041)	(5.155)	(4.563)	(0.163)	(1.152)	(0.879)
raised (log)	-0.330***	-0.354***	-0.253***	-0.023***	-0.032***	-0.033***
	(-8.681)	(-6.483)	(-3.530)	(-3.727)	(-3.195)	(-2.649)
% for sale		-0.238	0.380		0.026	0.021
		(-0.742)	(1.090)		(0.485)	(0.361)
presale		0.256	0.144		0.015	0.015
		(1.260)	(0.709)		(0.428)	(0.430)
social media at ICO end		-0.114**	-0.114**		0.003	0.009
		(-2.443)	(-2.336)		(0.346)	(1.053)
source commits at ICO end		0.089***	0.072**		0.005	0.007
		(3.216)	(2.509)		(1.053)	(1.518)
# wallets (log) at ICO end		-0.031	-0.028		-0.001	-0.004
		(-0.915)	(-0.759)		(-0.186)	(-0.618)
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummy	Yes	Yes	Yes	Yes	Yes	Yes
Region dummy	Yes	Yes	Yes	Yes	Yes	Yes
Observations	790	321	260	790	321	260
R-squared	0.210	0.368	0.339	0.072	0.145	0.143

Table 12. Long-term returns. This table reports regressions of determinants of longer-term post-ICO cumulative returns over different horizons: 30 days, 90 days, 180 days, and 365 days. In columns 1 and 2 the independent variable is the log difference between the token value at closing on the 30th trading day and the token value at closing on the first trading day. Measures at other horizons are defined in a similar fashion. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	30-day		90-day		180-day		365-day	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
return btc (log)	0.186	-0.278**	0.194**	-0.034	0.476***	0.331***	0.172**	0.648***
	(1.398)	(-2.058)	(2.398)	(-0.406)	(6.820)	(3.911)	(2.306)	(4.371)
return ether (log)	0.551***	0.727***	0.307***	0.495***	0.053***	0.088***	-0.002	-0.082***
	(6.064)	(7.338)	(7.118)	(9.136)	(3.354)	(4.219)	(-0.183)	(-2.695)
ICO end-to-open return (log)	-0.110***	-0.198***	-0.138***	-0.264***	-0.125***	-0.301***	-0.206**	-0.330***
	(-4.612)	(-6.817)	(-4.319)	(-6.875)	(-2.867)	(-5.413)	(-2.538)	(-2.681)
ICO first-day return (log)	-0.419***	-0.756***	-0.613***	-0.767***	-0.584***	-0.898***	-0.435	-1.358***
	(-3.338)	(-5.282)	(-3.826)	(-4.270)	(-2.835)	(-3.803)	(-1.218)	(-2.825)
raised (log)	-0.112***	-0.219***	-0.111***	-0.309***	-0.074*	-0.374***	-0.039	-0.359***
	(-4.581)	(-6.916)	(-3.432)	(-7.601)	(-1.795)	(-6.786)	(-0.579)	(-3.426)
# exchanges		0.062***		0.118***		0.145***		0.125***
		(5.372)		(8.171)		(8.159)		(3.455)
social media 1^{st} trading		-0.007		-0.015		0.013		-0.041
		(-0.315)		(-0.516)		(0.343)		(-0.527)
Δ social media (contemp.)		0.074		-0.053		-0.071		0.002
		(0.761)		(-0.811)		(-1.229)		(0.019)
source commits 1^{st} trading		0.005		-0.008		0.016		0.036
		(0.405)		(-0.466)		(0.740)		(0.808)
Δ source commits (contemp.)		-0.118		-0.018		0.004		0.030
		(-1.071)		(-0.255)		(0.063)		(0.277)
wallets 1^{st} trading		0.022		0.021		0.019		0.065
		(0.711)		(0.492)		(0.331)		(0.508)
Δ wallets (contemp.)		0.151**		0.086		0.212***		0.037
		(2.447)		(1.367)		(2.988)		(0.255)
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	760	506	712	474	582	382	289	164
R-squared	0.154	0.304	0.248	0.452	0.395	0.576	0.302	0.479

Table 13. Liquidity. This table reports regressions of determinants of average token liquidity measured over different horizons: 30 days, 90 days, 180 days, and 365 days. In columns 1 and 2, the independent variable is the negative log value of the average Amihud (2002)'s illiquidity measure given by $\frac{1}{30}\sum_{i=1}^{30}\frac{\left|\ln p_i - \ln p_{i-1}\right|}{p_i \times volume_i}$, where p denotes token price. Measures at different horizons are defined in a similar fashion. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	30-day		90-day		180-day		365-day	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ICO end-to-open return (log)	0.588***	0.643***	0.672***	0.633***	0.699***	0.640***	0.808***	0.791***
	(9.435)	(8.611)	(9.014)	(7.133)	(7.844)	(5.864)	(5.625)	(3.601)
ICO first-day return (log)	0.878**	0.684*	0.498	0.073	0.443	0.007	0.293	-1.361
	(2.554)	(1.728)	(1.275)	(0.160)	(1.005)	(0.013)	(0.482)	(-1.597)
raised (log)	0.877***	0.732***	0.994***	0.810***	1.084***	0.833***	1.143***	0.990***
	(12.459)	(8.587)	(11.967)	(8.117)	(11.332)	(7.116)	(8.760)	(5.324)
# exchanges	0.057***	0.128***	0.060***	0.148***	0.073***	0.198***	0.065***	0.176***
	(4.685)	(3.989)	(4.462)	(4.055)	(5.097)	(4.970)	(3.939)	(2.722)
social media 1st trading	-0.045	-0.014	-0.011	0.013	-0.056	-0.032	0.048	0.067
	(-0.787)	(-0.213)	(-0.166)	(0.178)	(-0.761)	(-0.378)	(0.431)	(0.486)
Δ social media (contemp.)	0.785***	0.412	0.297*	0.140	0.136	0.098	0.317**	0.529***
	(3.144)	(1.537)	(1.952)	(0.846)	(1.143)	(0.742)	(2.382)	(3.291)
source commits 1st trading	0.050	0.010	0.054	0.027	0.082**	0.043	0.093	0.019
	(1.589)	(0.283)	(1.500)	(0.664)	(2.004)	(0.909)	(1.521)	(0.238)
Δ source commits (contemp.)	-0.449	-0.345	-0.047	0.005	-0.113	-0.014	-0.162	-0.167
	(-1.425)	(-1.106)	(-0.306)	(0.028)	(-1.005)	(-0.100)	(-1.205)	(-0.861)
wallets 1st trading		0.275***		0.402***		0.544***		0.354
		(3.219)		(3.813)		(4.328)		(1.587)
Δ wallets (contemp.)		0.795***		0.884***		1.041***		0.844***
		(4.872)		(5.737)		(6.510)		(3.312)
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	688	517	649	485	532	387	263	164
R-squared	0.409	0.437	0.376	0.419	0.394	0.493	0.486	0.620

Table 14. Return volatility. This table reports regressions of determinants of daily return volatility, measured as daily standard deviation over different horizons: 30 days, 90 days, 180 days, and 365 days. We report t-statistics in parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

	30-day		90-day		180-day		365-day	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
cumulative return (log) (contemp.)	0.016***	0.014***	-0.001	0.002	-0.004*	-0.001	-0.006**	-0.005
	(4.051)	(3.136)	(-0.347)	(0.466)	(-1.843)	(-0.498)	(-2.051)	(-1.387)
volatility btc	0.417	0.552	0.582	1.084**	-0.054	1.193*	-0.678	-2.119
	(1.345)	(1.561)	(1.427)	(2.110)	(-0.103)	(1.690)	(-0.604)	(-0.998)
volatility ether	1.090***	1.011***	1.058**	0.650	0.734	-0.958	-0.025	3.554
	(3.930)	(3.008)	(2.275)	(1.044)	(1.075)	(-0.901)	(-0.028)	(1.506)
ICO end-to-open return (log)	-0.009***	-0.014***	-0.011***	-0.016***	-0.012***	-0.017***	-0.014***	-0.022***
	(-3.308)	(-4.235)	(-4.274)	(-5.072)	(-5.071)	(-5.228)	(-3.523)	(-4.140)
ICO first-day return (log)	-0.003	0.009	-0.011	0.013	-0.032***	-0.014	-0.054***	-0.015
	(-0.248)	(0.625)	(-0.987)	(0.943)	(-3.021)	(-1.077)	(-3.499)	(-0.748)
raised (log)	-0.022***	-0.026***	-0.023***	-0.027***	-0.024***	-0.025***	-0.023***	-0.030***
	(-7.909)	(-7.627)	(-9.108)	(-8.468)	(-10.274)	(-7.975)	(-6.885)	(-6.671)
# exchanges	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.004**
	(0.632)	(0.293)	(1.079)	(1.311)	(1.347)	(1.154)	(0.573)	(2.285)
social media 1st trading	-0.001	-0.000	-0.001	-0.000	0.000	0.000	-0.001	0.000
	(-0.612)	(-0.164)	(-0.578)	(-0.195)	(0.037)	(0.085)	(-0.420)	(0.080)
Δ social media (contemp.)	0.000	-0.004	-0.003	0.000	-0.004	-0.005	-0.002	-0.009**
	(0.018)	(-0.437)	(-0.733)	(0.046)	(-1.452)	(-1.478)	(-0.651)	(-2.387)
source commits 1st trading	-0.004***	-0.004***	-0.003***	-0.003**	-0.004***	-0.003***	-0.002	-0.000
	(-3.341)	(-2.848)	(-3.061)	(-2.471)	(-3.840)	(-2.617)	(-1.205)	(-0.171)
Δ source commits (contemp.)	0.000	-0.004	-0.003	-0.003	0.001	-0.000	0.003	0.003
	(0.037)	(-0.312)	(-0.629)	(-0.617)	(0.452)	(-0.124)	(0.948)	(0.557)
wallets 1st trading		0.003		0.002		-0.003		-0.005
		(0.891)		(0.507)		(-1.090)		(-0.999)
Δ growth (contemp.)		0.020***		0.008*		-0.002		-0.013**
		(3.218)		(1.789)		(-0.574)		(-2.166)
Time dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	669	506	631	474	520	382	256	164
R-squared	0.337	0.379	0.328	0.394	0.406	0.450	0.294	0.423