The Economics of Cryptocurrency Pump and Dump Schemes

JT Hamrick, Farhang Rouhi, Arghya Mukherjee,
Amir Feder, Neil Gandal, Tyler Moore and Marie Vasek

Working Paper No. 3-2019

The Foerder Institute for Economic Research
and
The Sackler Institute of Economic Studies

The Economics of Cryptocurrency Pump and Dump Schemes

JT Hamrick, Farhang Rouhi, Arghya Mukherjee, Amir Feder, Neil Gandal, Tyler Moore, and Marie Vasek*

Abstract

The surge of interest in cryptocurrencies has been accompanied by a proliferation of fraud. This paper examines pump and dump schemes. The recent explosion of nearly 2,000 cryptocurrencies in an unregulated environment has expanded the scope for abuse. We quantify the scope of cryptocurrency pump and dump on Discord and Telegram, two popular group-messaging platforms. We joined all relevant Telegram and Discord groups/channels and identified nearly 5,000 different pumps. Our findings provide the first measure of the scope of pumps and suggest that this phenomenon is widespread and prices often rise significantly. We also examine which factors affect the pump's "success."

1 Introduction

As mainstream finance invests in cryptocurrency assets and as some countries take steps toward legalizing bitcoin as a payment system, it is important to understand how susceptible cryptocurrency markets are to manipulation. This is especially true since cryptocurrency assets are no longer a niche market. The market capitalization of all cryptocurrencies exceeded $800 Billion at the end of 2017. Even after the huge fall in valuations, the market capitalization of these assets is currently around $140 Billion. This valuation is greater than the fifth largest U.S. commercial bank/commercial bank holding company in 2018, Morgan Stanley, which has a market capitalization of approximately $100 Billion.[1]

In this paper, we examine a particular type of price manipulation: the "pump and dump" scheme. These schemes inflate the price of an asset temporarily so a

*Hamrick: University of Tulsa, jth563@utulsa.edu. Rouhi: University of New Mexico, frouhi@unm.edu. Mukherjee: University of Tulsa, arm3606@utulsa.edu. Feder: Technion, amirfeder@mail.tau.ac.il. Gandal: Tel Aviv University, gandal@post.tau.ac.il. Moore: University of Tulsa, tyler-moore@utulsa.edu. Vasek: University of New Mexico, vasek@cs.unm.edu.

[1]The largest U.S.commercial bank/commercial bank holding company is JPMorgan Chase, with a market capitalization of slightly less than $400 Billion.

select few can sell at the artificially higher price. In the case of cryptocurrencies, at the beginning of a pump, a signal indicating the currency to buy is transmitted to insiders via a group messaging platform. Ideally, from the standpoint of the pumpers, the coordinated buying increases the trading activity and begins to drive up the price. When outside buyers are attracted and begin making purchases, the price rises further; then the pumpers sell the positions they acquired previously at lower prices.

The proliferation of cryptocurrencies and changes in technology have made it easier to conduct pump and dump schemes (pumps). Many of the nearly 2,000 cryptocurrencies available today are illiquid and are characterized by very low trading volumes on most days, with occasional volume and price spikes.

Our goal is to describe how the pumps work in the cryptocurrency realm, quantify the extent of the phenomenon, and to examine what factors (e.g., coin popularity, the number of exchanges on which it is traded) affect the "success" of a pump.

The data collection required for the analysis was substantial. Pump data was gathered by collecting messages posted to hundreds of dedicated Discord and Telegram channels using their APIs and manually labeling messages that signaled pumps. In order to obtain the data, we had to join hundreds of channels on Discord and Telegram and manually process their communication. We describe this process in detail in the body of the paper.

We then collected price data on nearly 2,000 coins across 220 cryptocurrency trading exchanges from `coinmarketcap.com`, the leading website of aggregated data on cryptocurrency trading during the six month period from January to July 2018. The price data is captured at the finest granularity presented by `coinmarketcap.com` at the time of collection, namely 5-minute intervals. Finally, we merged these two data sets in order to conduct the analysis.

Overall, we identified nearly 5,000 pump and dump schemes on Discord and Telegram that took place during the (slightly less than) six month period from mid-January 2018 to early July 2018. This provides the first measure of the scope of pump and dump schemes involving cryptocurrencies and indicates that the phenomenon is widespread. The cryptocurrency pumps also show us what financial markets might be like without regulation.

We then measured the success of the schemes, which we define to be the percentage increase in the price following a pump. Ten percent of the pumps on Telegram (Discord) increased the price by more than 18 percent (11 percent) in just five minutes. Recall that the January-July 2018 period was a period in which cryptocurrency prices were falling significantly; hence "moderate" percentage increases were an achievement for the pump.

Further, we examined what factors explained the ability to increase price. The most important variable in explaining success of the pump is the ranking of the coin.[2]

[2] Bitcoin has rank #1.

Coins with lower market capitalization typically have lower average trading volume. Lower average volume gives the pump scheme a greater likelihood of success. We found that pumps using obscure coins with low market capitalization on average led to greater price increases than pumping the dominant coins in the ecosystem: the median price increase was 3.5% (4.8%) for pumps on Discord (Telegram) using the top 75 coins; it was 23% (19%) on Discord (Telegram) for coins ranked over 500.

Additionally, we find that the number of exchanges the coin is traded on is negatively associated with success. This makes intuitive sense, because with fewer exchanges, pump schemes have better control over the total volume of the coin. We also find that the number of members that belong to different servers is positively associated with success. Information about member count is only available for Discord. This variable essentially measures the potential market for participating on pump schemes promoted on that server. This may be capturing a "herding" affect. In the analysis, we control for coin "volatility." We find that volatility is positively correlated with pump success, and furthermore that the other identified factors remain significant even after controlling for this volatility.

The road map for the paper is as follows. In the remainder of this section, we provide background information and review the literature. Section 2 provides a detailed description of the methodology and how we collected the data. In section 3 we describe the Discord and Telegram data, while 4 provides and discusses our results. Section 5 briefly concludes.

1.1 Background

History of the Cryptocurrency Market Bitcoin (BTC), the first cryptocurrency, was founded in 2009. While the market took off slowly, a massive spike in the price of bitcoin in late 2013 led to wider interest in what had been until then a niche industry. The value of Bitcoin increased from around $150 in mid 2013 to over $1,000 in late 2013. The fall was dramatic as well and bitcoin fell to $400 in a very short period of time. Despite the dramatic fall, cryptocurrencies were on the map and massive entry (as well as non-trivial exit) has occurred in the industry during the last five years.

While Bitcoin dominated the market through most of the 2009-2016 period, in 2013, a few other cryptocurrencies competed with Bitcoin. These coins began appreciating much more quickly than Bitcoin during the price rise. Gandal et al. analyzed how network effects affected competition in the cryptocurrency market during the price spike and subsequent fall in the price of Bitcoin [3]. Their analysis suggests that there were strong network effects and winner-take-all dynamics following the fall in the price of Bitcoin in early 2014. From July 2014 to February 2016, Bitcoin's value was essentially constant against the USD, while the other currencies depreciated dramatically against the USD. Litecoin, the number two coin in the market, declined by 70% in value, while other "main" coins declined by more than 90% in value. In early 2016, Bitcoin accounted for 94% of the total market capitalization, while Litecoin (the number two cryptocurrency) accounted for 2%. Despite its shortcomings, Bitcoin had emerged at that point as the clear winner and beneficiary of network effects.

In 2017, things changed dramatically. Bitcoin began rising again and by early 2017, the value of bitcoin was again more than $1,000. It had taken more than three years for the value of bitcoin to return to the 2013 peak level, but that was only the beginning. Eventually, in December 2017, Bitcoin reached a peak of more than $19,000 before plummeting over the next few months to $6,000.

The market capitalization of cryptocurrency grew stunningly in the past few years. In February 2014, the market capitalization of all cryptocurrencies was approximately $14 Billion. In January 2018, near Bitcoin's peak, the total market capitalization reached $825 Billion. As of February 2019, total market capitalization is approximately $132 billion.

In February 2018, there were 715 cryptocurrencies with market capitalization between $1 million and $100 million.[3] January 2014, there were less than 30 coins with market capitalization between $1 million and $100 million. This sharp four year rise in high-valued coins raises concerns of an increased potential for price manipulation.

The Larger Picture Cryptocurrency manipulations tie in to a concern over trading in unregulated financial exchanges. The potential for manipulation in the Over-the-Counter (OTC) markets is a significant concern for financial regulators. OTC trading is conducted directly between two parties, without going through a stock exchange. In a recent white paper, the SEC noted that "OTC stocks are also frequent targets of market manipulation by fraudsters."[4] The U.S. Securities and Exchange Commission (SEC) report also documents that OTC trading has increased significantly over time.[5]

Pump and dump schemes were outlawed in the 1930s. Nevertheless, the practice has continued. In the early 1990s the brokerage Stratton Oakmont artificially increased the price of "penny" stocks it owned by creating a "hype" around the stock. Once the price rose, the firm sold its shares in the relevant holding. The founder of Stratton Oakmont, Jordan Belfort, was convicted for securities fraud.

The U.S. SEC actively prosecutes pump and dump cases using publicly traded stocks. Such schemes involving cryptocurrencies are not any different. However, regulators have yet to prosecute pump and dumps involving cryptocurrencies. With the exception of insuring that taxes are paid on cryptocurrency profits and individual

[3] As of February 2019, there are 751 such coins.

[4] Outcomes of Investing in OTC Stocks, by Joshua White, December 16, 2016, U.S. Securities and Exchange Commission Division of Economic and Risk Analysis (DERA).

[5] In 2008 around 16 percent of U.S. stock trades were of the OTC type. By 2014, OTC trades accounted for 40 percent of all stock trades in the US. Like cryptocurrency trading, OTC trades are not transparent and not regulated, and there is concern that such trading is more harmful than high-frequency trading via regulated exchanges – See [12].

state-based regulation, US regulatory policy towards cryptocurrencies and initial coin offerings (ICOs) has been generally been "hands-off." One problem in moving forward in the regulatory sphere is that – unlike stocks, commodities, or fiat currency – cryptocurrencies do not have a regulatory agency in charge of all cryptocurrency policy.

Technologies like Telegram and Discord allow people to easily coordinate such schemes. Telegram is a cloud-based instant messaging service and uses Voice over Internet Protocol (VoIP). Users can send messages and exchange photos, videos, stickers, audio and files of any type. Messages can be sent to other users individually or to groups of up to 100,000 members. As of March 2018, Telegram had 200 million active users. Discord has similar capabilities and had 150 million users as of August 2018.

Discord and Telegram are primary sources for cryptocurrency pumps and have been used for pump and dump schemes on a large scale. Perhaps because of the regulatory vacuum, some of the pump groups do not hide their goals.

1.2 Literature Review

The academic literature on price manipulation and pump and dump schemes involving stocks includes Aggarwal and Wu [1]. They examined SEC litigation against market manipulators in OTC markets. They find stocks with low volume are subject to manipulation. They find that stock prices, volume, and volatility increase during the pump and dump scheme, but end quickly once it is over. They write that while manipulative activities have declined on main exchanges, it is still a serious issue in the over-the-counter (OTC) market in the United States.

Massoud et al. [11] studied OTC companies that hire promoters to engage in secret stock promotions to increase their stock price and trading volume. They find that the "promotions," or informal pump and dump schemes, coincide with trading by insiders. Brüggemann et al. [2] show that OTC stocks have lower levels of liquidity than a matched sample of similar NASDAQ listed stocks.

Cryptocurrency Price Manipulation Krafft et al. [8] created bots that executed penny trades in 217 different cryptocurrency markets. While their intent was not to incite bubble-type behavior, their bots created large price swings in the individual currencies after very small purchases.

Gandal et al. [4] identify and analyze the impact of suspicious trading activity on the Mt. Gox Bitcoin currency exchange, in which approximately 600,000 bitcoins (BTC) valued at $188 million were fraudulently acquired. They find that the USD-BTC exchange rate rose by an average of four percent on days when suspicious trades took place, compared to a slight decline on days without suspicious activity. They conclude that the suspicious trading activity by the Mt. Gox exchange itself likely caused the unprecedented spike in the USD-BTC exchange rate in late 2013, when the rate jumped from around $150 to more than $1,000 in two months.

A June 2018 working paper examined whether Tether, a digital cryptocurrency that is pegged to USD, affected the price of Bitcoin and other cryptocurrency prices during the huge increase in cryptocurrency valuations in 2017 [6]. Since they do not have data on which accounts initiated trades, they use algorithms to analyze blockchain data. They find that purchases with Tether occur following falls in Bitcoin prices and that the Tether purchases led to subsequent price rises in Bitcoin (and other cryptocurrency) prices. In particular, they find that short periods with especially heavy Tether trading volume are associated with "50 percent of the meteoric rise in Bitcoin and 64 percent of other top cryptocurrencies." They conclude that these purchases cannot be explained by investor demand, but that they are consistent with the hypothesis that Tether was used to provide price support and manipulate cryptocurrency prices.

Other researchers have studied financial fraud using cryptocurrencies. In two separate studies, Vasek and Moore [16, 17] researched online Ponzi schemes using cryptocurrencies. They measured millions of dollars reaped in by Ponzi scheme runners. Furthermore, they found that the most successful scams depend on large contributions from a very small number of victims. They then investigated Ponzi schemes advertised on the Bitcoin forum and the ecosystem that perpetuates them. Similar to our work, they mine information from the large social ecosystem around the cryptocurrency fraud they investigated.

Our work is quite different from the existing research on price manipulations; to the best of our knowledge, this is the first study to assess the scope of pump and dump schemes involving cryptocurrencies. We are also the first to examine which factors affect the "success" of pumps, where success means a large percentage increase in price.

Four other (essentially) concurrent papers also examine pump and dump schemes on cryptocurrencies, but with a different emphasis. Kamps and Kleinberg [7] use market data to identify suspected pump and dumps based on sudden price and volume spikes (and the following sharp decreases). They evaluate the accuracy of their predictions using a small sample of manually identified pump signals. Employing a similar approach with a different dataset, Mirtaheri et al. [13] use data collected from Twitter on cryptocurrencies cross-referenced with pump signal data from Telegram and market data. They note that a lot of the tweets are automated and attempt to predict pumps using only the Twitter traffic. Xu and Livshits [18] use data on just over 200 pump signals to build a model to predict which coins will be pumped. Their model distinguishes between highly successful pumps and all other trading activity on the exchange. Li et al. [9] use a difference-in-difference model to show that pump and dumps lower the trading price of affected coins.[6]

Our work is different from these papers in several important ways. First, we

[6] There have been media articles about the pump and dump phenomenon as well. Mac reported on pump and dump schemes in a Buzzfeed article published in January 2018 [10]. This was followed by work by Shifflet and Vigna [15] in a Wall Street Journal article published in August 2018.

have collected many more pump signals from channels on Discord and Telegram and evaluate them all, without restricting ourselves to the successful pumps. Our goal was to the extent possible to reach all pump and dump schemes on Discord and Telegram. Second, we investigate reported pumps for all coins with public trading data, not only those taking place at selected exchanges. This enables us to incorporate ecosystem-wide explanatory variables such as the number of exchanges on which a coin is traded on, the rank of the coin, etc., in order to assess what makes a pump and dump scheme successful.

2 Methodology

In this section, we discuss the methodology we used to collect the data on pump signals from social media and public messaging sources, as well as how we gathered pricing data and measured pump success.

2.1 Pump Signals Data from Discord and Telegram

Collecting Pump Signals Our objective was to collect as many pump signals as possible from all channels in these platforms. These platforms are the main outlets for pump and dump schemes.

A pump signal is an announcement to encourage people to buy a cryptocurrency and then take advantage of the price manipulation created by the surge in purchasing. The first step in collecting this data was to become familiar with the Discord and Telegram applications.

We programmatically scraped Discord and Telegram channels about pump and dumps using their respective APIs. We started our collection with URLs from a bitcointalk page on Discord pump groups: `https://bitcointalk.org/index.php?topic=2887116.0`. We then inspected all groups with over 4,000 users from an Android app that tracks the popularity of pump and dump groups (`https://padl.mine.nu/`). Afterward, we filtered the data based on keywords chosen specifically for each channel based on their posting patterns. This required a huge effort because communications/language on the channels were not uniform. We then manually inspected the filtered data and verified whether the post actually described an attempted pump or not, recording those that appeared to be pump signals. We parsed any additional channels that we learned about from a particular channel, adding them to our database. We are confident that we managed to get to most of the relevant channels. Our data was collected in August 2018 and our data spans the time period from mid-January to mid-July 2018.

With Discord, people join the servers. Individual channels/groups are associated with servers. The main purpose of the channels is to organize data, and any member of a server has access to all channels in that server. Thus, in the case of Discord, we were able to collect data on the number of members that belong to a specific server. It is not specific to a particular pump, since servers contain many channels; it essentially measures the potential market for participating on pump schemes promoted on channels on that server. Telegram is a cloud-based service where individual channels are set up by individual operators and hosted on Telegram's infrastructure. Hence, there is no analogous variable to number of members that belong to a specific server in the Telegram data.

We joined as many cryptocurrency-related channels as possible on both Discord and Telegram. The main challenge is that the only way to join many channels is by invitation. Another challenge is to make sure that an announcement is actually a pump signal. We systematically ignored a few types of posts. We did not consider posts about users predicting the future prices of the coins. We also ignored signals coins to "hodl" coins, which is a cryptocurrency meme for holding on to coins for a long period of time. Since "hodl"-ing is antithetical to the short term pump and dumps, we ignored these. We ignored channels with very few members. From the conversations between members of these unpopular channels, it became clear that even the few members do not actually participate in the pumps.

After painstakingly going through the channels, we discovered that several patterns are repeated in the pump and dump channels. Based on these patterns, we learned how the pumps worked and were able to characterize the channels into three broad categories.

Obvious Pumps The first category was the most straightforward to identify. These channels used the words "pump" and "dump" everywhere, including in the name of their channels. They usually had only a few pump announcements, and they posted pump signals infrequently. They usually posted the first announcement between 24 to 48 hours before the pump. Then, they posted many other announcements about timing and the cryptocurrency exchange where the pump would occur. When the time of pump came, they posted the name of the coin. They usually posted the pump results a few hours afterward, along with the date of the next pump. In many cases, they posted the name of the coin in an image. In other cases, they gave many coin names with a check mark in front of the actual coin being pumped. This was likely done to combat automated scraping of the coin name.

These channels usually had premium membership as well. The premium membership was based on how many people a person had recruited to the channel. Users could also buy premium membership plans. Based on the type of plans, premium members would receive the pump signals a certain amount of time before others. A concept that these groups frequently used is known as a "collaboration pump." A collaboration pump means that many different channels post the same coin at the same time to increase the trading volume. These pump signals were not posted by channel owners. They were posted by bots. Since these signals were repeated many times in many channels, we tracked them to avoid having more than one copy of the same signal. "Obvious pumps" accounted for 27 Discord groups (with 139 corresponding pump signals) and 12 Telegram channels (with 565 corresponding pump signals).

Target Pumps The second category was not as brazen as the first category. These channels usually avoided the words "pump" and "dump". The main concern that was reflected in their chatrooms was that members were not sure if pump and dump was legal, so they avoided using the name. They had many more pump signals than the first category. They posted the name of the coin and the current price, without any previous announcement. They usually announced the exchange for the pump as well. They also gave target prices asking participants to sell at any of these prices, but not lower than them. In many cases, they also gave some news about the coin. These channels typically did not a have premium membership option. However, some of the these channels required payment for membership. These channels usually posted the same signals a certain amount of time before the freely available channels.

The main challenge for us in this category was to make sure the announcements were actually pump announcements. These channels typically had lots of members. The first indicator was that they had another channel where they talked about the pump results and thanked members for participation. The other indicator was that they did not use any technical analysis or technical indicators to analyze the market. Further, they tried to convince people to buy the coin and participate. "Target pumps" accounted for 24 Discord groups (with 543 pumps) and 13 Telegram channels (with 4,159 pumps).

Copied Pumps The third category was copied the signals from other sources. Although they usually posted the signals hours after the pump, they included the actual time that a pump was published. They also included the source of that pump. We preferred not to use these signals, because we wanted to collect our data from primary sources. We used these channels to ensure complete coverage, i.e., to find the pump sources and follow them. We included them in the analysis when we could not get access to the source channels. This was an important source since copied pumps accounted for 4 Discord groups with 514 associated pump signals not found elsewhere. There are no Telegram pumps in this category because of the complete overlap between these Telegram groups and other signals already collected.

Although most of the copied pumps were from other Discord channels, approximately five percent of our Discord data overlaps with our data from Telegram. We include the copied pumps for completeness, but our results are qualitatively unchanged if we remove the small number of these copied pumps from the Discord analysis.

Summary of Pump Signals In the case of Telegram, 88 percent of the signals were target pumps and 12 percent were obvious pumps. In the case of Discord, 42 percent of the signals were target pumps, 40 percent of the signals were copied pumps, and 11 percent were obvious pumps. Thus more than 10 percent of the pumps were blatant and unconcerned about any regulatory actions.

2.2 Observations on Pump Organizer Behaviors

In the process of collecting the data, we learned quite a bit about how the organizers of cryptocurrency pumps and dumps operate and which exchanges the are used in the schemes. Here we provide a brief summary of what we found.

Pump Promotion When channels begin operating, they usually have a small number of members. Thus, they cannot schedule their own pumps. Some of these channels wait until they reach a certain number of members, and then start publishing pump signals. However, this can be a long wait for them, because not many people tend to join such inactive groups. Thus, most of these groups try to start by copying pump signals originally published by other channels. Some of them later move to organizing their own pumps and we categorize them based on the time in which we join the group and the amount of pump signal overlap.

Pump group organizers have two basic strategies to promote coins: "news" and "science." News strategies use rumors purported as news, almost always later found out to be false information. These rumors often were ill-defined; a common message before a pumping a coin is "heard good news is coming soon." Sometimes this is made marginally more specific, like from the user `ADA_All_The_Way` on the Pure Investments channel: "Heard some big news coming in April for Ada." Others referenced rumors from other platforms (for example, a message from the Crypto Cartel Original channel reads: "ETC undervalued. `https://twitter.com/eth_classic/status/950546415394029568/photo/1`").

Further, some pump organizers used more "scientific" measures. They would post graphs of the price of a currency and strategize about when would be optimal to pump. They also would also use algorithmic methods to spot dips in price for coins, and advocating buying after the price hits a target low.

Cryptocurrency Exchange Selection Participants usually try to focus their trading on one exchange to maximize the effect of the pump. Often the name of the exchange is published hours before the actual pump. Features that influence exchange selection include popularity, coin availability, chatroom ability, registration information required, and accessibility.

Many of these channels use a voting system to identify the exchange preferred by users. Choosing a popular channel is also helpful for promoting pumps on social media platforms to attract outside traders. The more popular the exchange, the more of these outsiders will end up trading in that exchange. The only exception is the smaller pump and dump channels. These groups know that they cannot induce as many purchases, which makes influencing prices more difficult. In these cases, the channels might decide to target a smaller and less popular exchange. Even in these cases, the channels employ a voting strategy, letting their member choose among smaller exchanges.

An additional consideration is where the pumped coin is traded. Not all exchanges support less frequently traded coins. Exchanges that support many low volume coins include Binance, HitBTC and Cryptopia. These exchanges are frequently chosen by pumpers.

Another important factor is the availability and popularity of the chatroom of each exchange. Pumpers promote their activities on the exchange's chatroom to attract outsiders. Many exchanges provide such chatrooms on their websites, while others rely on social media platforms. The more popular the chatroom, the more outsiders are exposed to the pump.

A final consideration is accessibility. Many cryptocurrency exchanges limit the countries in which they operate, and they often require extensive information in order to register, trade, or withdraw funds. Discord is based in North America and most of its users are Americans. This could explain why Poloniex, which requires a US social security number for registration, is a preferred trading platform for pumps on Discord but is only rarely used by pumps promoted on Telegram. On the other hand, Telegram users tend to be more international. So, BitMEX, which is restricted in North America, is a relatively popular choice among pump groups using Telegram, but not Discord.

2.3 Pricing Data on Cryptocurrencies

We collected price data on nearly 2,000 coins and tokens (henceforth referred to as coins) across 220 exchanges as reported to `coinmarketcap.com`, the leading website of aggregated data on cryptocurrency trading. We collected all price data for each of the coins listed on `coinmarketcap.com` from mid-January through early July 2018. This gave us a total of 316,244,976 collective volume and price data points across all of the coins listed. The data points collected are at the finest granularity presented by `coinmarketcap.com` at the time of collection, a 5-minute interval.

We realize there are limitations to this method of data collection. For instance, `coinmarketcap.com` does not list every coin or token available for purchase or trade. Further, this data is slightly different than what one would be able to collect from an exchange API. Since the website is collecting data from so many sources, it reports a volume weighted average of all of the prices reported at each exchange to calculate the price it reports. On the plus side, this approach is more comprehensive in the number of exchanges and coins covered.

Every internet service experiences outages planned or otherwise; the services we are interested in are no exception to the rule. During the initial data exploration phase, gaps in the data were discovered. To make sure these gaps were recorded in the data and not a result of our collection efforts, we programmatically check the data for proper intervals. If a gap exists in the data that spans a time period equal to or greater than 7.5 minutes, we report that data point as missing. We chose 7.5 minutes because of the 5 minute average interval in the data collected. After iterating through the timeline of each of the coins, we create an hour long window surrounding the missing data points and query `coinmarketcap.com` for that data. If the gap persists after the additional data collection, we surmise it is because of an outage either due to the exchange or `coinmarketcap.com`. In total we are missing approximately 3,806,474 volume and price records across all of the coins, or approximately 1% of the data.

Matching Discord/Telegram Information with Trading Data For the purpose of our study, it was essential to ensure a consistent mapping between what is announced in the pump signal to what is associated with the trading data. In particular, pump signals are by no means consistent when it comes to the coin names used in the messages. Some users refer only to the coin ticker such as DOGE, which is the ticker for Dogecoin. This can be a bad idea as several cryptocurrencies employ identical tickers (being decentralized, there is no equivalent to NYSE or NASDAQ to enforce the uniqueness of ticker symbols). Others use the full coin or token name, but that can be problematic because many coins have similar names. For instance, the cryptocurrency IOTA has the ticker MIOTA; the coin name is similar to the ticker for IoTex, which is IOTX. Still others use some combination of the ticker and full or partial name. For example, "Bitcoin (BCD)" refers to Bitcoin Diamond and not Bitcoin as the ticker for Bitcoin is BTC and not BCD.

We normalized reports to the name used by `coinmarketcap.com`. To do this, we created a name map that contains several variations of the actual cryptocurrency name based on our observations. We then removed special characters from the names reported in Discord and performed a case insensitive comparison to the map we created. If a match was found, we replaced the pump name with a clean version that matches the name elsewhere in our data. Some of the names required manual replacement since cryptocurrencies have the ability to rebrand. In this way, we were able to map 1,034 of the Discord pump signals and 3,767 of the Telegram pump signals to more than 300 cryptocurrencies.[7]

Identifying Pump Timing and Success Throughout the processes of aggregating, combining, and cleaning the data, it became increasingly apparent that we could not reliably use the time of a pump signal to mark the beginning of a period of anomalous trading activity.[8]

[7] We have more total pumps than that, but approximately 5% do not have complete data and cannot be used in the analysis.

[8] This may be because "insiders," i.e., those running the pump, strategically purchase before the agreed upon time. This is consistent with the other work in this area. [7] noticed that pumps sometimes occurred exactly when a signal was put out and other times occurred afterwards. [9]

Hence, instead of taking the pump signal time as given, we treat it as the starting point to identify associated spikes in trading activity. We inspect 48 hours before and after the time of the reported signal to find the maximum percentage jump between two consecutive price data points (typically spaced 5 minutes apart).

In the data analysis described in the next section, we use this maximum 5-minute percentage increase in this 96 hour period in the coin's price relative to BTC as our measure of pump success. In the analysis, we will control for "pre-pump" coin volatility, as we explain below.

3 Data Description

In this section, we describe the data available for the study. Our goal is to examine what factors explain the success of the pump and dump scheme, where success means that the pump increased the price significantly.

The Discord and Telegram data spans the nearly six month period from mid-January to early July 2018. In the full dataset, a small number of observations were duplicates in the sense that they involved the same coin, took place on the same day and at roughly the same time (within an hour) on the same exchanges. We eliminated the duplicates, but the results are qualitatively unchanged if we include them.

Once we eliminate the duplicate observations and a few observations for which we did not have complete data, we are left with 1,034 observations with complete data on Discord and 3,767 observations with complete data on Telegram. This gives a sense of the scope of the pump and dump phenomenon on these platforms.

3.1 Dependent Variable

We employ the maximum % price increase (as described above) in the 48 hours preceding and following the pump as the dependent variable. We denote this variable as % Price Increase.

Most of the cryptocurrencies cannot be directly traded with USD, but they can be traded with bitcoin. Hence, we use coin prices in bitcoin. Because of this, we cannot include the pumps using bitcoin itself. There were 6 pumps of bitcoin on Discord and 76 pumps of bitcoin on Telegram. While these pumps account for only 1.7% of all pumps, it is interesting to note that even bitcoin is not immune from the pump and dump phenomenon.

3.2 Independent Variables

We have the following independent variables.

collected more pump signal information than [7] and observed the same effect. [18] collected *hourly* market data, and found that the markets move as much as 72 hours before an announced pump.

- Exchanges: the number of exchanges on which the coin can be traded. We measured this variable twice: once at the end of 2017 and once in September 2018. The correlations are above 0.99 and the results are unchanged regardless which date we choose. The 2018 variable has more observations, so we use that one.

- Rank: the rank of the coin in terms of market capitalization. Bitcoin is #1. Coins with higher rank have lower market capitalization.

- Pair Count: the number of other coins that the coin can be traded with.[9]

- Server-Member-Count (Discord Only): the number of members that belong to a server (which is not specific to a particular pump). This variable essentially measures the potential market for participating on pump schemes promoted on that server.

These variables are clearly exogenous to the pump. Descriptive statistics for all variables used in the analysis appear in Table 1 and Table 2.

Table 3 groups coins by popularity. In Table 3, the entry # of coins represents the number of coins in that category that were pumped. Thus in the first row, 52 of the top 75 coins were pumped on Discord; this represents 69% of these coins. In the same row, there were 348 such pumps involving coins in the top 75, and the average percent increase from these pumps was 3.5%. Table 3 shows that while many of the pumps involve coins with light trading and low market capitalization (similar to penny stocks), pumps are not limited to obscure coins. Coins with greater market caps experience smaller spikes in prices: the median price increase is between 3.5-4.8% for the top 75 coins, compared to 19–23% for coins ranked over 500. See Table 3 for the full breakdown.

The pumping of more "mainstream" coins may be because it is not always easy to pump obscure coins that are traded on a small number of exchanges. Additionally, there is less volatility in mainstream coins, and some "investors" (pumpers) may have preferred a relatively lower risk level.

Overall, in the case of Discord data, the median (mean) percentage price increase was 3.5% (7.4%), while the 75th percentile of the distribution was 6.3%. In the case of Telegram data, the median (mean) percentage price increase was 5.1% (9.8%), while the 75th percentile of the distribution was 9.2%. Recall that the January-July 2018 period was a period in which cryptocurrency prices and trading volume were falling significantly; hence "moderate" percentage increases were an achievement for the pump.

[9]Similar to exchanges, we measured this variable twice, once at the end of 2017 and once in September 2018. The correlations are above 0.99 and the results are unchanged regardless which date we choose. The 2018 variable has more observations, so we use that one.

From the above discussion, it is not surprising that the coin rank is the independent variable that is most highly correlated with the percent price increase of the pump, both on Discord (0.48) and Telegram (0.35.) The correlations among the variables are shown in Table 4 and Table 5. As Table 4 and Table 5 show, the correlations are similar across the Discord and Telegram platforms.

4 Analysis and Results: Discord and Telegram Pumps

4.1 Explaining Success in Increasing Price

In the regressions in Table 6 and in Table 7, we use the percentage price increase as the dependent variable. We first run two regressions with different functional forms: (1) a linear/linear ordinary least squares (OLS) regression when using the percent price increase as dependent variable and the explanatory variables in levels (2) a log/log OLS regression using the natural logarithm of the variables, both the dependent variable and the independent variables.[10] We employ clustered standard errors at the level of the coin, since many of the coins appear more than once in the data set. The regressions appear in the first two columns of Table 6 and Table 7. The results are as follows:

- The log/log regression where we use the natural logarithm of the variables (both the dependent variable and the independent variables) has much higher explanatory power, in the sense that it has a much higher adjusted R-squared. This is true both for Discord and Telegram. This is not surprising given that all variables (except for Rank) are highly skewed.

  In the case of Discord, the log/log regression has an adjusted R-squared of 0.29 versus 0.21 for the linear/linear regression. In the case of Telegram, the log/log regression has an adjusted R-squared of 0.28 versus 0.12 for the linear/linear regression. Hence, the log/log regression is the preferred regression.

- The ranking of the coin is positively associated with success for both Discord and Telegram. This effect is highly significant in all regressions.[11] Coins with lower market capitalization typically have lower average volume. Lower average volume gives the pump scheme a greater likelihood of success. This effect is statistically significant and is obtained regardless of functional form for both Discord and Telegram.

- The number of exchanges on which the coin can be traded is negatively associated with success in the preferred regression; the effect is statistically significant and is obtained for both Discord and Telegram in the preferred regression.

[10]Since rank is not a skewed variable, we use the level of rank in this regression, but nothing changes qualitatively if we would use the log of the rank.

[11]Recall that higher rank means more obscure.

This makes intuitive sense, because with fewer exchanges, pump schemes have better control over the total volume of the coin.

- In the case of Discord, the variable server member count is positively associated with success; the effect is statistically significant in the case of the log/log regression, which is the preferred regression.
- The number of other coins that the coin can be traded with is not statistically associated with success.

These results are consistent with liquidity effects and herding behavior that have been identified for financial assets:

- Higher rank implies lower liquidity and higher volatility.
- More exchanges imply higher liquidity and lower volatility.
- More members on the chat server imply the possibility of more herding behavior and higher volatility.

Controlling for volatility It is possible that more volatile coins are selected to be pumped in order to improve the chances to profit from them. If this were true, some or all of the identified price rise could be attributed to the coin's inherent volatility. In other words, pumpers might behave like surfers who wait for the best waves to ride.

To explore this possibility, we have devised a measure of volatility and re-run the regressions. In order to try to control for coin volatility, we calculated the following variable, denoted "Volatility," which is defined to be the maximum five-minute percentage increase in the 96 hours preceding the pump.[12] We believe that this is a reasonable proxy for the volatility of the coin. Of course it is not ideal, because some coins were certainly pumped before we began collecting the data. In such a case, this variable would be endogenous. Nevertheless, these pumps are a relatively recent phenomenon, so the exercise seems reasonable.

Not surprisingly, there is a positive correlation between Volatility and the percentage price increase associated with the pump. In the case of Discord it is 0.42. In the case of Telegram it is 0.21. When we add this variable to the right hand side of the regression, we find the following:

- The volatility of the coin is indeed associated with a higher percentage price increase from the pump.
- Controlling for volatility, we find that our results are qualitatively unchanged.

[12]Since some coins are pumped more than once, we calculate the variable in the 96 hours preceding the first pump in our data. For a small number of observations, we do not have the Volatility measure.

- Higher rank is associated with a higher percentage price increase from the pump.
- More exchanges are associated with a lower percentage price increase from the pump.
- (For Discord) More members on the chat server is associated with a higher percentage price increase from the pump.

These results (using the same functional form as the regressions in column 2) are shown in the third regressions of Table 6 and Table 7.

What happens after the pump is over An interesting question is what happens after the pump is over. To address this issue, we calculate two additional variables.

- **Starting price**: this is the starting price associated with the maximum five minute percentage increase in price. It can be interpreted as the "pre-pump" price.
- **End price**: This is the minimum price in the 48 hours after pump.
- We then calculate the following variable: $\frac{\text{End price}-\text{Starting price}}{\text{Starting price}}$. This is the percentage change in price from the pre-pump period to the post-pump period.

We find the following: The median percentage change in price from the pre-pump period to the post pump period is -41% for Discord data and -38% for Telegram data. Overall, more than 60% of the coins have a lower "post-pump" price than the "pre-pump" price. Even though prices were generally falling during this period, a 40% fall in prices in 48 hours is large.[13]

Trading volume data We do not have corresponding volume data, since volume data on `coinmarketcap.com` is reported continuously over the preceding 24 hour period and it is not clear how often volume information is updated.[14]

But we did calculate the following volume variable: "Per-change volume after," which equals the maximum (24 hour) volume in the 24 hours following a pump signal less the minimum (24 hour) volume in the 24 hours following a pump signal divided by the minimum (24 hour) volume in the 24 hours following a pump signal. We find the following.

[13]We ran regressions using the percentage change in price from the pre-pump period to the post-pump period as the dependent variable, and the right-hand-side variables as the independent variables. In these regressions, the adjusted R-squared was virtually zero.

[14]Since we do not have delineated trading volume, we cannot quantify profitability from the pumps. Even if did have trading volume by time, it still would be impossible to measure profitability. This is the "pumpers" act as individuals and others can trade as well. The only way to measure profitability would be to have access to trading activity over time at the individual level; labeled trading data is not available.

- On both Discord and Telegram, there is approximately a 30 percent correlation between (i) the maximum five-minute percentage change in price and (ii) the "Per-change volume after."

- Since the price signal occurs before the changes in volume, we could run a regression with volume change as the dependent variable and put the maximum five-minute percentage change in price as a right-hand-variable along with the other independent variables used in the price regressions. In such a case, only that variable is significant and the adjusted R-squared is relatively large. Since we are not exactly sure of the timing, we do not want to push this too much, but it does suggest the following: The pump organizers buy first, increasing the price. Then the "herd" jumps in, where the herd is comprised of other people who received the pump signal and outsiders (some of whom may be using trading algorithms.) During this period, the original "pumpers" are likely selling there shares as well.

- The two points above suggest that the maximum five-minute percentage change in price is a good proxy for success.

The number of exchanges mentioned by the pump Finally, we briefly summarize data regarding the number of exchanges mentioned by the pump. We observed this for 546 pumps on Discord, around half of the total. We scraped this data from the pump signal, counting any exchanges directly mentioned in the signal message.

While most pumps mention a single exchange, more than 18 percent of the pumps mention more than one exchange. Correlations among the number of pump exchanges and the independent variables are shown in Table 8. Not surprisingly, the number of exchanges used in the pump is negatively correlated with the rank of the coin (-0.25) and positively correlated with the number of exchanges the coins are traded on (0.29.) These numbers give us additional confidence that we are indeed picking up actual pumps.

5 Brief Conclusions

In this paper we examined the phenomenon of pump and dump schemes for cryptocurrencies. The proliferation of cryptocurrencies and changes in technology have made it relatively easy (and virtually costless) for individuals to coordinate their activities.

In terms of scope, we found that this pump and dump phenomenon is widespread on both Discord and Telegram. We also found out that the most important variable in explaining success of the pump is the ranking of the coin. While there are attempts to pump coins spanning a wide range of popularity, pumping obscure coins gave the pump scheme the potential for greater success at the expense of increased risk, i.e., volatility. In some sense, the choice between using lower or higher ranking is similar to conservative and risky investment strategies: the benefit of investing in assets with low expected returns is that the volatility is low. The key difference, of course, is that deliberately pumping cryptocurrencies for profit is unethical.

Our results have implications for regulatory policy. In July 2018, the U.S. SEC rejected a proposal to include Bitcoin in a managed Exchange Traded Fund (ETF).[15] They rejected it in part due to concerns over possible price manipulation. But in general, U.S. regulatory policy towards cryptocurrencies can be characterized as hands-off. U.S. regulatory policy is inhibited in part because overlapping agencies have authority for regulating different aspects of the cryptocurrency ecosystem. The Internal Revenue Service (IRS), Financial Crimes Enforcement Network (FinCEN), the Commodity Futures Trading Commission (CFTC), and the Securities and Exchange Commission (SEC) are all involved in regulation related to the issuance, sale, and exchange of cryptocurrencies. A recent paper notes that depending on the regulatory agency, according to U.S. Law, cryptocurrencies can be money, property, a commodity, and a security [5]. This causes confusion and creates a regulatory vacuum.[16]

While federal regulators have not been pursuing pump-and-dump schemes, state attorneys general have been active in investigating forms of price manipulation. The New York State Office of the Attorney General investigated cryptocurrency fraud at the cryptocurrency exchange level [14]. They found that while most trading platforms acknowledged that market manipulation and fraud were issues, they lacked controls to evade abusive behavior, such as pump and dump trading activity. One currency exchange, Kraken, did not submit to their formal inquiry, but rather submitted a statement admitting that they did not believe market manipulation to be an issue.

But state regulation is not enough. Federal regulators should be very concerned that price manipulation via pump and dump schemes is so widespread. The scope of the phenomenon should raise red flags, especially as mainstream financial institutions begin investing in cryptocurrencies. These schemes illustrate why we need clear and consistent regulatory guidance at the federal level.

[15] An ETF is an investment fund traded on stock markets. ETFs typically hold assets like stocks, bonds, and commodities. Unlike a mutual fund, an ETF is traded like a stock and prices change continuously throughout the day. ETFs have typically been index funds, for example, there are several S&P 500 index ETFs that track the S&P 500 index.

[16] Nevertheless, market manipulation such as cryptocurrency-related pump and dump schemes could be viewed as illegal in the United States under the Securities Exchange Act of 1934 Rule 10-b5 which makes interstate commerce using manipulation or deceptive devices illegal. This is not a legal opinion, of course; it is simply an example of an existing law that could apply. Such clarification would be helpful since there is widespread belief spread by pump organizers that these schemes might be legal under US law.

Acknowledgements

We gratefully acknowledge support from the following research grants: US-Israel Binational Science Foundation grant No. 2016622, US National Science Foundation Award No. 1714291, and a Blavatnik Interdisciplinary Cyber Research Center Research award. We thank Christophe Bisière and participants at the 12th Digital Economics Conference in Toulouse for excellent comments and suggestions that significantly improved the paper.

References

[1] Rajesh K Aggarwal and Guojun Wu. Stock market manipulations. *The Journal of Business*, 79(4):1915–1953, 2006.

[2] Ulf Brüggemann, Aditya Kaul, Christian Leuz, and Ingrid M Werner. The twilight zone: OTC regulatory regimes and market quality. *The Review of Financial Studies*, 31(3):898–942, 2017.

[3] Neil Gandal and Hanna Halaburda. Can we predict the winner in a market with network effects? Competition in cryptocurrency market. *Games*, 7(3):16, 2016.

[4] Neil Gandal, JT Hamrick, Tyler Moore, and Tali Obermann. Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics*, 95:86–96, 5 2018.

[5] Carol Goforth. U.S. Law - crypto is money, property, a commodity, and a security, all at the same time. *Journal of Financial Transformation*, forthcoming, 2019.

[6] John M Griffin and Amin Shams. Is bitcoin really un-tethered? Available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3195066`, 2018.

[7] Josh Kamps and Bennett Kleinberg. To the moon: defining and detecting cryptocurrency pump-and-dumps. *Crime Science*, 7(1):18, 2018.

[8] Peter M Krafft, Nicolás Della Penna, and Alex Sandy Pentland. An experimental study of cryptocurrency market dynamics. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 605, 2018.

[9] Tao Li, Donghwa Shin, and Baolian Wang. Cryptocurrency pump-and-dump schemes. Available at: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3267041.`, 2018.

[10] Ryan Mac. Bitcoin scammers are using this app to fleece people. *Buzzfeed News*, January 2018. Available at: `https://www.buzzfeednews.com/article/ryanmac/cryptocurrency-scammers-are-running-wild-on-telegram`.

[11] Nadia Massoud, Saif Ullah, and Barry Scholnick. Does it help firms to secretly pay for stock promoters? *Journal of Financial Stability*, 26:45–61, 2016.

[12] John McCrank. Dark markets may be more harmful than high-frequency trading. *Reuters Business News*, 4 2014. `http://www.reuters.com/article/us-dark-markets-analysis-idUSBREA3508V20140406`.

[13] Mehrnoosh Mirtaheri, Sami Abu-El-Haija, Fred Morstatter, Greg Ver Steeg, and Aram Galstyan. Identifying and analyzing cryptocurrency manipulations in social media. Available at `https://arxiv.org/abs/1902.03110`, 2019.

[14] New York State Office of the Attorney General. Virtual currencies integrity initiative report, September 2018. Available at: `https://ag.ny.gov/sites/default/files/vmii_report.pdf`.

[15] Shane Shifflett and Paul Vigna. Some traders are talking up cryptocurrencies, then dumping them, costing others millions. *The Wall Street Journal*, August 2018. Available at: `https://www.wsj.com/graphics/cryptocurrency-schemes-generate-big-coin/`.

[16] Marie Vasek and Tyler Moore. There's no free lunch, even using Bitcoin: Tracking the popularity and profits of virtual currency scams. *Financial Cryptography and Data Security*, pages 44–61, 1 2015.

[17] Marie Vasek and Tyler Moore. Analyzing the Bitcoin Ponzi scheme ecosystem. *Fifth Workshop on Bitcoin and Blockchain Research*, 2018.

[18] Jiahua Xu and Benjamin Livshits. The anatomy of a cryptocurrency pump-and-dump scheme. Available at: `https://arxiv.org/abs/1811.10109.`, 2018.

A Appendix

Table 1: Descriptive Statistics: Discord, N=1,034

Variable	Obs	Mean	Std. Dev.	Min	Max
Max % Price inc.	1,034	6.96	16.78	0.64	221.90
Exchanges	1,034	21.48	26.72	1	182
Pair Count	1,034	24.07	89.90	1	759
Rank	1,034	256.98	308.17	2	1,863
Server Member Count	1,034	5,616.05	9,741.03	141	84,823

Table 2: Descriptive Statistics: Telegram, N=3,767

Variable	Obs	Mean	Std. Dev.	Min	Max
Max % Price inc.	3,767	9.96	21.29	0.42	341.99
Exchanges	3,767	18.51	24.45	1	182
Pair Count	3,767	19.11	72.20	1	759
Rank	3,767	394	433	2	2,036

Table 3: Median Price Increases by Coin Rankings.

	Discord				Telegram			
	Coins		Signals	Price	Coins		Signals	Price
Rank	#	%	#	Inc %	#	%	#	Inc %
$\leq$ 75	52	69.33	342	3.51	56	74.67	1,000	4.81
76-200	58	46.40	257	5.22	62	49.60	736	6.46
201-500	75	25.00	285	5.32	84	28.00	948	8.10
$>$ 500	73	5.33	150	23.23	176	11.46	1,083	18.74

Table 4: Correlations Among Variables: Discord, N=1,034

Variable	% Price inc.	Exchanges	Pair Count	Rank	Server Member Count
% Price inc.	1				
Exchanges	−0.14	1			
Pair Count	−0.056	0.71	1		
Rank	0.45	−0.43	−0.18	1	
Server Member Count	−0.011	−0.0090	0.017	0.0015	1

Table 5: Correlations Among Variables: Telegram, N=3,767

Variable	% Price inc.	Exchanges	Pair Count	Rank
% Price inc.	1			
Exchanges	−0.16	1		
Pair Count	−0.067	0.69	1	
Rank	0.35	−0.45	−0.19	1

Table 6: Examining What Affects Success of Pump and Dump Schemes: Discord Data

Independent Variables	Dependent Variable % Price Increase linear/linear	Dependent Variable % Price Increase log/log	Dependent Variable % Price Increase log/log
Exchanges	0.54	-0.14	-0.11
	(0.0025)	(0.060)	(0.056)
Pair Count	-0.0055	0.037	0.049
	(0.0045)	(0.051)	(0.045)
Rank	0.026	0.0012	0.0011
	(0.0053)	(0.00018)	(0.00021)
Server Member Count	-0.000016	0.10	0.087
	(0.000041)	(0.020)	(0.020)
Volatility			0.30
			(0.053)
Observations	1,034	1,034	1,019
Adjusted R^2	0.21	0.29	0.33

Standard errors in parentheses: clustered standard errors at the level of the coin

Table 7: Examining What Affects Success of Pump and Dump Schemes: Telegram Data

Independent Variables	Dependent Variable % Price Change linear/linear	Dependent Variable % Price Change log/log	Dependent Variable % Price Change log/log
Exchanges	0.0024	-0.22	-0.021
	(0.017)	(0.044)	(0.045)
Pair Count	-0.0011	-0.016	-0.00056
	(0.0025)	(0.040)	(0.040)
Rank	0.018	0.00069	0.00061
	(0.0024)	(0.000090)	(0.000091)
Volatility			0.15
			(0.035)
Observations	3,767	3,767	3,740
Adjusted R^2	0.12	0.28	0.29

Standard errors in parentheses: clustered standard errors at the level of the coin

Table 8: Correlation Table: Pump Exchange and Independent Variables, N=546

Variable	Pump Exchanges	Exchanges	Pair Count	Rank75
Pump Exchanges	1			
Exchanges	0.29	1		
Pair Count	0.13	0.73	1	
Rank	−0.25	−0.40	−0.18	1