The Economics of Cryptocurrency Pump and Dump Schemes*

JT Hamrick[†2], Farhang Rouhi[‡3], Arghya Mukherjee[*2], Amir Feder[§1], Neil Gandal[‡1], Tyler Moore[*2], and Marie Vasek[†3]

[1]Berglas School of Economics, Tel Aviv University, Israel
[2]Tandy School of Computer Science, The University of Tulsa, USA
[3]Department of Computer Science, University of New Mexico, USA

Abstract

The surge of interest in cryptocurrencies has been accompanied by a proliferation of fraud. This paper examines a pervasive tactic long known to financial markets: pump and dump schemes. While the fundamentals of the ruse have not changed in the last century, the recent explosion of nearly 2,000 cryptocurrencies in a largely unregulated environment has greatly expanded the scope for abuse. The paper first quantifies the scope of cryptocurrency pump and dump on Discord and Telegram, two widely popular group messaging platforms with 130 million users and 200 million users respectively. Both platforms can handle large groups with thousands of users, and they are the most popular outlets for pump and dump schemes involving cryptocurrencies. We identified 3,767 different pump signals advertised on Telegram and another 1,051 different pump signals advertised on Discord during a six-month period in 2018. The schemes promoted more than 300 cryptocurrencies. These comprehensive data provide the first measure of the scope of pump and dump schemes across cryptocurrencies and suggest that this phenomenon is widespread and often quite profitable. This should raise concerns among regulators. We then examine which factors that affect the "success" of the pump, as measured by the percentage increase in price near the pump signal. We find that the coin's rank (market capitalization/volume) is the most important factor in determining the profitability of the pump: pumping obscure coins (with low volume) is much more profitable than pumping the dominant coins in the ecosystem.

*Initial draft: submitted on October 4, 2018 to the 12th Digital Economics Conference, Toulouse, France. This draft version: December 18, 2018.

[†]{jth563,arm3606,tyler-moore}@utulsa.edu
[‡]{frouhi@unm.edu, vasek@cs.unm.edu}
[§]{amirfeder@mail.tau.ac.il,gandal@post.tau.ac.il

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. 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 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 drives up the price. When outside buyers are attracted and begin making purchases, 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. We 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. This gave us a total of 316,244,976 price data points across all of the coins listed. The data collected are at the finest granularity presented by `coinmarketcap.com` at the time of collection, namely 5-minute intervals.

Pump data was then gathered by collecting messages posted to hundreds of dedicated Discord and Telegram channels using their APIs and manually labeling messages that signaled pumps. We then ensured that the extracted pump signals announced on Discord and Telegram consistently mapped to the trading data.

Through a comprehensive data collection process, we identified 1,051 pump and dump schemes from Discord data and 3,767 pump and dump schemes from Telegram data that took place during the (slightly less than) six month period from mid-January 2018 to early July 2018. These comprehensive data provide the first measure of the scope of pump and dump schemes involving cryptocurrencies and suggest that this phenomenon is widespread. The scope should raise "red flags" for regulators, especially as mainstream financial institutions begin investing in cryptocurrencies.

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 (12 percent) in just five minutes. 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.

We then examined what factors explained the ability to increase price. The most impor-

tant variable in explaining success of the pump is the ranking of the coin.[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 were much more profitable 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. We discuss the effect of other variables on the "success" of the pumps when we conduct the analysis.

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 by 2016, one bitcoin was worth approximately $200. Despite the bursting of the bubble, cryptocurrencies were on the map and massive entry (as well as non-trivial exit) has occurred in the industry during the last four 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 and Halaburda [3] analyzed how network effects affected competition in the cryptocurrency market during the price spike and subsequent fall in the price of Bitcoin. 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 Bil-

[1] Bitcoin has rank #1.

lion. As of February 2018, the total market capitalization was approximately $414 Billion.[2]

In February 2018, there were more than 300 cryptocurrencies with market capitalization between $1 Million and $100 Million. In January 2014, there were less than 30 coins with market capitalization between $1 million and $100 million. This has raised concerns of an increased potential for price manipulation.

The Larger Picture Cryptocurrency manipulations tie in to a concern in 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."[3] The U.S. Securities and Exchange Commission (SEC) report also documents that OTC trading has increased significantly over time.[4]

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, 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. [5]

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, first released in 2015, has similar capabilities and 130 million users as of May 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, many of the pump groups do not hide their goals.

[2]In early December 2018, the total market capitalization was approximately $122 Billion.

[3]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).

[4]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 McCrank [11].

[5]An exception to the "hands off" policy took place in March 2018, when the SEC pressed charges against PlexCorps for an ICO that likely defrauded investors. Additionally, the SEC announced increased enforcement regarding violations of securities law.

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. [10] 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. [7] 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 buying a pennies worth.

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 by Griffin and Shams examined whether Tether, a digital cryptocurrency that is pegged to the USD, affected the price of Bitcoin and other cryptocurrency prices during the huge increase in cryptocurrency valuations in 2017 [5]. 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.

The New York State Office of the Attorney General [13] investigated cryptocurrency fraud at the cryptocurrency exchange level. They found that most trading platforms had insufficient controls to evade abusive behavior, such as pump and dump trading activity. They also unsurprisingly found that cryptocurrency exchanges overwhelmingly lacked consumer protections for the currency that they held. Given that it is common for currency exchanges to be hacked [12], this is quite worrisome.

Other researchers have studied financial fraud using cryptocurrencies. In two separate

studies, Vasek and Moore [15, 16] 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. Similarly to our work, their work depends on mining information from the large social ecosystem around the cryptocurrency fraud they investigated.

Our work is quite different from the 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.

Concurrent Work on Pump and Dumps Three other (essentially) concurrent papers also examine pump and dump schemes on cryptocurrencies, but with a different emphasis. Kamps and Kleinberg [6] use market data to identify suspected pump and dumps based on sudden price and volume spikes. They evaluate the accuracy of their predictions using a small sample of manually identified pump signals. Xu and Livshits [17] 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. [8] use a difference-in-difference model to show that pump and dumps lower the trading price of affected coins.[6]

Our work is distinguished from this other concurrent work in several important ways. First, we have collected many more pump signals from channels on Discord and Telegram and evaluate them all, without restricting ourselves to the successful pumps. 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 in order to assess what makes a pump and dump scheme successful.

2 Methodology

Here we describe the data we collected on pump signals from social media and public messaging sources.

2.1 Pump Signals Data from Discord and Telegram

We collected pump signals from Discord and Telegram. 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.

[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 [9]. This was followed by work by Shifflet and Vigna in a Wall Street Journal article published in August 2018 [14].

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.

With Discord, people join the servers. Individual channels 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. Telegram is a cloud-based service where individual channels are set up by individual operators and hosted on Telegram's infrastructure.

Next we join as many cryptocurrency related channels as possible. 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. After going through most of the channels, we discovered that there several patterns are repeated in the pump and dump channels. Based on these patterns, all of the channels could be organized in three 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 plans 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.

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 tried to announce the exchange 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. 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. These channels typically had lots of members: the idea is that when many people enter the market at the same time, a pump can happen. They also tried to convince people to buy the coin and participate.

Copied Pumps: The third category was channels that 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 tried to avoid these channels, because we wanted to collect our data from original resources. We primarily used these channels to ensure complete coverage, i.e., to find the pump sources and follow them. However, we have included a few of them in our analysis. The reason is that some of these channels had a large number of followers suggesting that they were capable of creating pumps on their own. Additionally, some of the sources that they used were very popular, but we did not have access to them since they were private channels.

Pump Signal Collection: We programmatically scraped Discord and Telegram channels about pump and dumps using their respective APIs. We seeded our collection with URLs from a bitcointalk page on Discord pump groups: `https://bitcointalk.org/index.php?topic=2887116.0`. We complemented this by inspecting all groups with over 4,000 users from an Android app that tracks the popularity of pump and dump groups (`https://padl.mine.nu/`). Then we filtered the data based on some specific keywords. These keywords were chosen specifically for each channel based on the specific patterns that each channel followed. We 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. Any additional channels that we learned about on a particular channel we parsed were then added to our list of channels to examine. Our data was collected in August 2018 and our data spans the time period from mid-January to mid-July 2018.

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 that were for coins to "hodl" - 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.

It stands to reason that pumps will be more successful if they can attract more participants. Many channels in Telegram and Discord try to achieve this by collaborating with other channels in other platforms. Approximately, 5 percent of our Discord data overlaps with the Telegram data due to such collaborations.

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. As a result, most of these groups try to start by copying pump signals originally published by other channels. This will guarantee a certain volume of trades, while it contributes to the success of such pumps. Since most of Discord channels are smaller than Telegram channels, and Discord is less popular for pump and dumps, this trend leads to approximately 10 percent of our Discord data having overlaps with the Telegram data. These copied pump announcements can be found in less popular channels and at the beginning of large channels in Discord.[7]

2.2 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 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.

2.3 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.

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

[7] Our results remain qualitatively unchanged if we eliminate the overlapping observations.

(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 1,051 of the Discord pump signals and 3,767 of the Telegram pump signals to more than 300 cryptocurrencies.

2.4 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. 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 the coin price relative to BTC as our measure of pump success. While it may often be the case that the price rise lasts longer or shorter than 5 minutes, we have found that the 5 minute percentage price increase is a robust indicator, particularly given how volatile the coin price data can be.

3 Data Analysis

In this section, we discuss the Discord and Telegram 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 span the nearly six month period from mid-January to early July 2018. In the full data set, 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 1,051 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) 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 important to note that bitcoin is not immune from the pump and dump phenomenon.[8]

3.2 Independent Variables

We have the following independent variables.

- 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. When we include this variable in the analysis, we are left with 1,033 observations for the Discord regressions.

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 3.51%. 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. However, 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.

[8] Alternatively, we could include the bitcoin pumps in the analysis by converting bitcoin to USD. Given the relatively small number of bitcoin pumps, our results are qualitatively unchanged if we include pumps of bitcoin.

[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.

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% (10.0%), 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.

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 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. 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 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 the higher explanatory power, in the sense that it has a higher adjusted R-squared. This is true both for Discord and Telegram. Hence, the log/log regression is the preferred regression.

  In the case of Discord, the log/log regression has an adjusted R-squared of 0.34 versus 0.24 for the linear/linear regression. In the case of Telegram, the log/log regression has an adjusted R-squared of 0.24 versus 0.12 for the linear/linear regression.

- The ranking of the coin is positively associated with success.[10] 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 obtains 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. This makes intuitive sense, because with fewer exchanges, pump schemes have better control over the total volume of the coin. This effect is statistically significant and obtains for both Discord and Telegram in the preferred regression.

[10] Recall that higher rank means more obscure.

- 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.

4.2 Explaining the Number of Exchanges Used in the Pump

In the case of Discord, we also have data on "Pump-Exchanges," i.e., the number of exchanges in which the pump took place for more than 50 percent of the 1,051 pumps. We scraped this data from the pump signal, counting any exchanges directly mentioned in the signal message. This is clearly an endogenous variable. Here, we briefly examine what factors explain the number of pump exchanges employed.

While most pumps occur on a single exchange, more than 18 percent of the pumps occurred on more than one exchange. Correlations among the number of pump exchanges and the independent variables are shown in Table 8. The distribution of the number of pump exchanges is shown in Table 9.

Here employ two dependent variables: (i) the number of exchanges used in the pump and (ii) a 0/1 binary variable where the variable takes on the value one if the pump occurred on more than one exchange and zero otherwise. We find the following:

- Pumps on coins with smaller market capitalization (higher rank) take place on fewer exchanges - and this effect is statistically significant. This is the only variable that is statistically significant and it is significant in both regressions. The results are shown in Table 10.

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 and conduct pump and dump schemes.

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 variety of ranking, pumping obscure coins gave the pump scheme a greater likelihood of success.

Our results have implications for regulatory policy. The U.S. SEC recently (July 26, 2018) rejected a proposal to include Bitcoin in a managed Exchange Traded Fund (ETF).[11] The rejection of the proposal was in part taken because concerns of possible price manipulation.

[11] 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.

Regulators should be very concerned that price manipulation via pump and dump schemes is so widespread.

Acknowledgements We gratefully acknowledge support from the following research grants: US-Israel Binational Science Foundation grant No. 2016622 and US National Science Foundation Award No. 1714291.

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, May 2018.

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

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

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

[8] 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.

[9] 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`.

[10] 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.

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

[12] Tyler Moore and Nicolas Christin. Beware the middleman: Empirical analysis of Bitcoin-exchange risk. In *Financial Cryptography and Data Security*, pages 25–33. Springer, 2013.

[13] 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`.

[14] 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/`.

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

[16] Marie Vasek and Tyler Moore. Analyzing the Bitcoin Ponzi scheme ecosystem. In *Fifth Workshop on Bitcoin and Blockchain Research*, Lecture Notes in Computer Science. Springer, 2018.

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

6 Appendix

Table 1: Descriptive Statistics: Discord

Variable	Obs	Mean	Std. Dev.	Min	Max
Max Price inc.	1,033	7.41	17.89	0.64	221.90
Exchanges	1,033	21.18	26.16	1	182
Pair Count	1,033	24.56	88.45	1	759
Rank	1,033	264	323	2	1,869
Server Member Count	1,033	5,458	9,325	141	84,823
Pump Exchanges	546	1.21	0.49	1	4

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 %
≤ 75	52	69.33	348	3.51	56	74.67	1,000	4.81
76-200	58	46.40	259	5.22	62	49.60	736	6.46
201-500	75	25	286	5.32	84	28	948	8.10
> 500	80	5.84	158	23.23	176	11.46	1,083	18.74

Table 4: Correlations Among Variables: Discord, N=1,033

Variable	% Price inc.	Exchanges	Pair Count	Rank	Server Member Count
% Price inc.	1				
Exchanges	−0.15	1			
Pair Count	−0.058	0.70	1		
Rank	0.48	−0.42	−0.18	1	
Server Member Count	0.043	−0.021	0.0080	0.010	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 (1) % Price Increase linear/linear	Dependent Variable (2) % Price Increase log/log
Exchanges	-0.015***	-0.15***
	(0.0053)	(0.0059))
Pair Count	0.0028*	0.059
	(0.0017)	(0.050)
Rank	0.0029***	0.0013***
	(0.00041)	(0.00018)
Server Member Count	0.000034***	0.17***
	(0.000011)	(0.023)
Observations	1,033	1,033
Adjusted R-squared	0.24	0.34

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

*** p<0.01, ** p<0.05, * p<0.1

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

Independent Variables	Dependent Variable (1) % price change linear/linear	Dependent Variable (2) % price change log/log
Exchanges	0.0024	-0.28***
	(0.017)	(0.054)
Pair Count	-0.0011	0.034
	(0.0025)	(0.043)
Rank	0.018***	0.15***
	(0.0024)	(0.041)
Observations	3,767	3,767
Adjusted R-squared	0.12	0.24

Standard errors in parentheses: clustered standard errors at the level of the coin
*** p<0.01, ** p<0.05, * p<0.1

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

Table 9: Pump Exchanges Distribution

Percentiles	# of Exchanges	Distribution	
1%	1		
5%	1		
10%	1	Obs	546
25%	1		
50%	1	Mean	1.20
		Std. Dev.	0.48
75%	1		
90%	2	Variance	0.23
95%	2	Skewness	2.74
99%	3	Kurtosis	11.34

Table 10: What Affects Number of Exchanges used in the pump: Discord Data

Independent Variables	Dependent Variable (1) # of Pump Exchanges OLS	Dependent Variable (2) "More than one" Logit
Exchanges	0.0066***	0.0084
	(0.0022)	(0.0069)
Pair Count	-0.00086*	-0.0020
	(0.00046)	(0.0015)
Rank	-0.00021***	-0.0064***
	(0.000055)	(0.0016)
Observations	546	546
	0.12 (adjusted R^2)	0.16 (Pseudo) R-squared

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

*** p<0.01, ** p<0.05, * p<0.1