Direct Evidence of Bitcoin Wash Trading *

Arash Aloosh† Jiasun Li‡

This version: April 2021

Abstract

We use the internal data of a major Bitcoin exchange leaked by hackers to detect wash trading – a type of market manipulation in which a single trader clears her own limit orders to "cook" transaction records. Our finding provides *direct* evidence for the widely-suspected "fake volume" allegation against cryptocurrency exchanges, which has so far only been backed by *indirect* inferences. Wash trades tend to follow low *past* transaction fee revenues to the exchange, significantly increase *subsequent* transaction fee revenues, and involve several known exchange insiders. The evidence is consistent with the hypothesis that the exchange itself commits wash trading to inflate apparent trading volume and boost fee revenues. We also use our direct evidence to evaluate the indirect inference techniques proposed in the literature.

Keywords: bitcoin; cryptocurrency; exchanges; forensics; market manipulation; regulation.

*We thank Patrick Augustin, Neil Gandal, Giacomo Sbrana, Christine Parlour, Leif Anders Thorsrud, Baolian Wang, and David Yermack, as well as attendants at NBER Summer Institute, DC Junior conference, NEOMA, Digital Finance Chair of Université Paris 2 & Telecom Paris, Nankai Business School, Central University of Finance and Economics, Renmin-HKBaptist-NTU Joint Digital-Economy Seminar Series and the Int'l Joint Conference on Theoretical Computer Science for helpful comments.

†arash.aloosh@neoma-bs.fr. NEOMA BS, 1 Rue du Marechal Juin, 76130 Mont-Saint-Aignan, France.

‡jli29@gmu.edu. George Mason University, 4400 University Drive, MSN 5F4, Fairfax, VA 22030, USA.

There have been suspicions that many cryptocurrency exchanges manipulate the market by permitting or even engaging in so-called wash-trading, in which fake trading records are created by having the same trader clearing his or her own standing limit order(s).[1] Since exchanges only display trades and quotes information rather than the trader identity behind each trade or order, market participants are unable to distinguish between a wash trade and a genuine trade. Wash trading therefore inflates the volumes of cryptocurrency exchanges and misleads market participants about the actual exchange liquidity conditions. While wash trading is explicitly banned in all regulated stock, derivative, or commodity exchanges, similar regulations do not apply to many cryptocurrency exchanges. Such suspicions are so widely believed that many industry data providers now offer adjusted "real" trading volume of cryptocurrency exchanges (See for example Blockchain Transparency Institute).

It has immediate policy implications to understand whether such suspicions are indeed valid in reality, and if so, why they arise and what impact they have on the cryptocurrency market. First, it helps us weigh in the debate on whether cryptocurrency exchanges should be subject to more strict regulations, given that some exchanges have complained about over- or mis-regulation (e.g. New York attorney general vs. Kraken),[2] leading to some exchanges leaving the US market for good (e.g. Poloniex following its acquisition by Justin Sun) while others opting for working closely with regulators (e.g. Gemini founded by the Winklevoss twins). Second, understanding manipulations on centralized cryptocurrency exchanges could

[1]Some anecdotes: On July 22, 2020, Coindesk reports that Coinsquare will settle with the Ontario Securities Commission (OSC) over allegations that executives had employees fake trades to inflate the platform's volumes. As part of the settlement agreement, Coinsquare admitted that around 840,000 illicit wash trades were conducted on the platform, amounting to a total value of around 590,000 bitcoin (BTC) (worth almost $5.5 billion at press time). On Aug 26, 2020, CoinDesk EU News Editor Daniel Palmer reports that Coinbit, South Korea's third-largest cryptocurrency exchange, appears to have been seized by police over allegations that it faked most of its trading volume. Exchange insiders and police said up to 99% of the platform's trading volume was "manipulated," or washed, using "ghost" accounts – totalling over 100 billion won ($84 million) in faked income. Seoul Newspaper, which broke the news, said it had seen the books and that 99% of recorded trades could not be associated with deposits or withdrawals.

[2]Summarized in Underwood (2018) and Kraken's Blog. As a side note, Kraken's CEO Jesse Powell contributed to rebuilding the Mt.Gox exchange following its 2011 hack, an event to be discussed in more details later in the paper.

help us make more fair comparisons between centralized cryptocurrency exchanges and the emerging applications of decentralized exchanges, given the various types of manipulations identified among those applications.[3] Finally, in its numerous rejections against various bitcoin ETF proposals, the SEC has referred to potential market fragility due to thin volume disguised by seemingly abundant "fake" volumes. A deeper understanding of the issue may further enlighten the regulatory decisions.

That said, so far all studies into potential wash trading and fake volume on exchanges have been exclusively based on indirect inferences. For example, using price impact analysis from large order execution experiments, Ribes (2018) document potential fake volume on exchanges. In an influential presentation to the SEC, Bitwise (2019) claims that 95% of cryptocurrency exchange volumes are fake, even though the accuracy of this estimate has been subsequently questioned. In the academic literature, Cong, Li, Tang and Yang (2020) examine 3 regulated and 26 unregulated crypto exchanges by comparing statistical patterns of public exchange data. Similar indirect inference techniques have also been adopted by Amiram, Lyandres and Rabetti (2020) to relate wash trading and exchange competition.

Despite these ample *indirect* estimates, without *direct* evidence from individual-level transaction data, it is impossible to verify the various *indirect* inference techniques. Since all the techniques proposed in the literature are based on statistical irregularities, they cannot directly point fingers at wash trading, as the captured anomalous transactions may be motivated by various strategies other than wash trading (e.g. stealth trading à la Alexander and Peterson (2007)). Indeed, Tibeaudou (2019) has presented counter-arguments against some indirect inference techniques for identifying wash trading. Furthermore, even if indirect inferences do accurately quantify the magnitude of wash trading, they still cannot indisputably answer many follow-up questions that the community may be interested in: For example, who are the perpetrators of wash trading, and specifically are they associated with

[3] See, for example, Daian, Goldfeder, Kell, Li, Zhao, Bentov, Breidenbach and Juels (2020).

the exchanges themselves? What are the motivations behind wash trading, and specifically are they motivated by exchange insiders' desires to inflate apparent trading volumes and thus boost fee (commission) revenues, assuming that liquidity begets liquidity? Separately, given that most of the fake volume allegation surfaced *circa* 2018, one may also ask whether wash trading is a recent phenomenon, or is it actually an "original sin" of cryptocurrency exchanges that has accompanied their births?

In this paper, we fill these gaps by providing direct evidences on wash trading. Our analysis uses the internal transaction logs from Mt.Gox, the largest bitcoin exchange by reported volume during our sample period. The internal data include complete transaction records with timestamps, price levels, sizes, and most importantly, trader identities (IDs) for both the buyer and seller of each trade. The data was first leaked by hackers to the bitcoin community upon the exchange's collapse in 2014, and have subsequently been studied by both academics (Gandal, Hamrick, Moore and Oberman (2018)) and practitioners (Nilsson (2014) or more commonly known as the Willy report) to investigate potential market price manipulations that led to the 2014 Bitcoin bubble, which occurs at a different time from our analysis.[4] Section 1 provides more institutional details and summary statistics for the data.

Section 2 presents our direct evidence of wash trading on the major bitcoin exchange Mt.Gox: Among 16 million buy/sell records, representing 8 million unique trades, we find more than 115 thousand trades with both sides (buy and sell) having the same trader ID. These trades first emerged in June 26$^{\text{th}}$, 2011, immediately after the exchange restored itself following a one-week halt of service due to a hacker attack, and has been prevalent ever since until May 20$^{\text{th}}$, 2013, when the exchange was investigated by regulators for a separate litigation. Within the June 26$^{\text{th}}$, 2011 - May 20$^{\text{th}}$, 2013 period, wash trades constitute more than 2% of the total number of trades. These wash trades involve 2,887 unique trader IDs, which we define as *wash trader IDs*, and transactions among the 2,887 wash trader IDs

[4]See also Investopedia and the WSJ article about the Willy report.

account for roughly 33% of all trades during the June 26$^{\text{th}}$, 2011 - May 20$^{\text{th}}$, 2013 period.

Section 3 proceeds with analyzing the consequence of, as well as the market conditions associated with wash trading. Inspired by the fake volume allegations in the cryptocurrency community, we use a vector autoregression (VAR) framework to characterize the joint dynamics between the intensity of wash trading and trading fee (commission) revenues to the exchange. We find that wash trading tend to occur following low fee revenues, and a higher wash trading intensity significantly increases subsequent fee revenues to Mt.Gox. It appears that the motivation for wash trading is to boost fee revenues, presumably by inflating volumes to create an illusion of liquid market conditions and thus attract more orders. Because the exchange itself should have the strongest incentives to boost fee revenues, to better support the fee-boosting hypothesis, we look for further direct evidence of exchange insider involvement in wash trading activites.

Section 4 dives into the 2,887 wash trader IDs to shed light on the perpetrators committing wash trading, and in particular, whether the exchange insiders (Mt.Gox's owner or affiliated individuals/entities) engage in wash trading. We use multiple sources to identify trader IDs that are likely controlled by exchange insiders, and investigate their involvement in wash trading. Consistent with our hypothesis, many known insiders, including for example, a trader ID nicknamed MagicalTux (a pseudo-name of Mark Karpelés, owner and CEO of Mt.Gox), as well as trader ID #1, the first registered trader on Mt.Gox, who is likely closely connected to the exchange, indeed engage in wash trading.

We also use the identified exchange insiders to address a robustness concern: Because a single person or entity could potentially register multiple trader IDs on Mt.Gox (using different email addresses), our direct evidence may underestimate the true extent of wash trading in reality. Therefore, we also augment our wash trading sample with transactions that have both sides (buy and sell) being identified insiders and repeat our exercises. These transactions are what market participates typically refer to as painting the tape, a closely related concept to wash trading, in which a group of associated traders clear orders among themselves (instead of a single trader clearing her own orders) to "cook" fake transaction records. We report results with various identifications of exchange insiders.

Finally, Section 5 uses our direct evidence to evaluate Benford's law, an indirect inference techniques for anomalous transactions proposed in the literature, and assesses its merit and limitations.

Related literature Besides those papers mentioned in the introduction on indirect inferences of anomalous transactions in the cryptocurrency market, this paper is also related to an emerging literature documenting various manipulations in the cryptocurrency market. For example, using the same data as ours but focusing on a different time period, Gandal, Hamrick, Moore and Oberman (2018) and the Willy report relate the late-2013-early-2014 bitcoin bubble/crash to price manipulations on the Mt.Gox exchange by a trading bot commonly known as Willy, who started operation in September 2013. This manipulation is different from wash trading identified in this paper, as the former is done by one-sided buys with funds apparently created out of thin air, while the latter involves buying and selling simultaneously between oneself. Wash trading in our sample also occurs prior to the time when Willy starts operation.

Griffin and Shams (2020) document another price manipulation, which relates the 2017 bitcoin bubble to Tether issuance from a single large bitcoin address. Specifically, this single address, allegedly belonging to Bitfinex, appears to "mint" Tether out of thin air and purchases Bitcoin mostly around major bitcoin price support levels so as to create an illusion of strong buying pressure and to push bitcoin price higher. This price manipulation also involves one-sided buying, rather than simultaneous buying and selling as in wash trading studied in this paper.

A third type of price manipulation concerns so-called pump-and-dump schemes. For example, Li, Shin and Wang (2019) and Xu and Livshits (2019) provide direct evidence of pump-and-dump schemes in the cryptocurrency market using communication records on Telegram. Hamrick, Rouhi, Mukherjee, Feder, Gandal, Moore and Vasek (2018) further covers communication records on Discord.

In a broader sense, our findings also contribute to an emerging literature of the bitcoin and cryptocurrency trading market. For example, Makarov and Schoar (2020), Choi, Lehar and Stauffer (2020), and Yu and Zhang (2020) document large and recurrent arbitrage opportunities across exchanges and especially across borders. Liu, Tsyvinski and Wu (2019), Liu and Tsyvinski (forthcoming), and Li and Yi (2019) study the factor structures in cryptocurrency returns. Shams (2020) and Benetton and Compiani (2020) relate crypto-asset returns to investor demands, while Schwenkler and Zheng (2021) relate them to co-mentions in news. Biais, Bisiere, Bouvard, Casamatta and Menkveld (2018) relate Bitcoin prices with changes in transactional benefits and costs of bitcoin. Augustin, Rubtsov and Shin (2020) studies the impact of introducing the Bitcoin futures contract on the spot market. Foley, Karlsen and Putniņš (2019) study the illegal usage of cryptocurrencies. Our paper also touches upon the regulation of the crypto market in general, as in Li and Mann (2018).

Our paper is also another example of using leaked data to answer finance questions, see for example O'Donovan, Wagner and Zeume (2019). The section on evaluating indirect wash trading measures using our direct evidence also relates to similar approaches seen in other markets (see for example, Ellis, Michaely and O'Hara (2000), Ahern (forthcoming)). Finally, we contribute to studies on financial market manipulations in other markets, such as Allen and Gale (1992), Aggarwal and Wu (2006), Cumming, Johan and Li (2011), Griffin and Shams (2018) and more generally the empirical literature on financial misconducts, such as Ritter (2008), Zitzewitz (2012), Egan, Matvos and Seru (2019), Parsons, Sulaeman and Titman (2018), Bollen and Pool (2009), Dimmock, Gerken and Graham (2018), Liu (2016), Lie (2005), Kedia, Koh and Rajgopal (2015), Karpoff, Koester, Lee and Martin (2017), and

Chakrabarty, Moulton, Pugachev and Wang (2020).

The remainder of the paper is organized as follows: Section 1 describes the data. Section 2 presents direct evidence of wash trading. Section 3 investigates the conditions under which wash trading tends to arise and its market impact. Section 4 dives into the wash trader profiles to shed light on the perpetrators. Section 5 evaluates the accuracy of indirect wash trading measures using our direct evidence. Section 6 concludes.

1 Data description

Our data comes from the internal trading records of Mt.Gox, which is widely regarded as the first, and for a long time during its life, the only major bitcoin exchange in the world. Figure 1 presents the market shares of Mt. Gox in terms of trading volume over time. Individuals could deposit various fiat currencies to the exchange and purchase bitcoin, or conversely, deposit bitcoin to the exchange and sell for various fiat currencies of their own choices. Mt.Gox profits from collecting fees from all transactions. Within each fiat-bitcoin trade, the trader who buys bitcoin with fiat incurs a bitcoin fee (in units of bitcoin) and the one who sells bitcoin for fiat incurs a fiat fee (in units of the receiving fiat currency). The fee revenues to the exchange sums up both legs for each trade after currency conversions.

[Figure 1 about here.]

Figure 2 presents a series of snapshots of Mt.Gox's interface, including the welcome page, trading interface, fee calculator (which illustrates the bitcoin fees and fiat fees), the notice put forward when Mt.Gox collapsed in February 2014, and an account summary on a desktop terminal (which shows the parent company information, Tibanne Co. Ltd). The interfaces look similar to a typical online brokerage site.

[Figure 2 about here.]

Figure 3 plots the evolution of bitcoin prices during our sample period, and chronicles several major events throughout Mt.Gox's life. Of particular interests are June 19 2011, when Mt.Gox recovered from a hack that brought down its website for about a week. Wash trading first emerges immediately after this incident.

[Figure 3 about here.]

The data were first leaked by hackers in February 2014 upon the collapse of Mt.Gox. The entire leaked zip file in its original form contains 60 "trade" files (at least one for each month from April 2011 to November 2013). There are more trade files than the number of months because for some months by the end of sample, as bitcoin trading volume spiked there was one file for each week. Also for April 2013, there is an extra anonymized log in addition to the full log, which differ in two and only two places that we will elaborate further in Section 4 for identifying an exchange insider nicknamed "MagicalTux".

Each trade file includes second-by-second transaction records, with information that includes transaction ID, trader ID, amount (in both units of BTC and fiat currencies), time, order type (buy/sell), currency, JPY exchange rate, user location (country, state), a dummy for Japan or not (JP/NJP), and the two type of fees (fiat fees and Bitcoin fees). Figure 4 presents a snapshot of the original files.

[Figure 4 about here.]

Merging all the trade files renders approximately 17 million transactions (double counting each buy-sell pair) on Mt.Gox from April 1st, 2011 to November 30th, 2013. The original log however contains many duplicate transactions. We use a de-duplication process similar to that of Feder, Gandal, Hamrick and Moore (2017). Specifically, we remove duplicate transaction records with the same transaction ID, trader ID, transaction time, transaction type (buy/sell), and transaction amount.[5] Removing duplicates narrows the data to 16 million transactions, which is closer to the daily volumes reported on bitcoincharts.com than the original leaked data. To validate our de-duplication process, Figure 5 compares the daily volume calculated from our transaction records.

[Figure 5 about here.]

Following Nilsson (2014), we further exclude all "quartet transactions" prior to March 2013 by trader IDs TIBANNE_LIMITED_HKand THK, which are the only non-numerical IDs in our data, and are widely believed in the bitcoin community to belong to a "super user" created by Mt.Gox's parent company Tibanne.[6] It is argued that before March 2013, this super user's main role was to facilitate cross-currency trades.[7] Table 1 reports summary statistics for the cleaned sample. These transactions represent 108 million Bitcoins in trading volume (54 million Bitcoins buy/sell) from April 1st, 2011 to November 30th, 2013.

[Table 1 about here.]

Mt.Gox serves a global customer base from more than 160 countries and supports 17 fiat currencies. Mt.Gox users can opt to verify their accounts on Mt.Gox to shorten the waiting periods upon withdrawal. For verified users, Federal Information Processing Standard (FIPS) state codes will be entered into the User_Country and sometimes User_State variables in the

[5] The de-duplication procedure in Feder, Gandal, Hamrick and Moore (2017) also relies on fields including trader ID, transaction time, transaction type (buy/sell), and transaction amount, but does not include transaction IDs. We believe the inclusion of transaction IDs renders a more accurate result, as for example, a robot may post two identical orders, which are cleared at the same time, but going by with different transaction IDs. Both methods nevertheless generate similar results.

[6] Tibanne registered a company in Hong Kong on May 19th 2011, creating a "super user" with ID TIBANNE_LIMITED_HK, and later changed to THK in March 2013. This ID enjoys an exceptionally high privilege as it can trade without incurring any fees. For example, its first trade on August 27th 2011 7:48 bought 1BTC with USD and then at the same time sold 1 BTC to JPY, neither incurring any fees. The ID THK traded in total 2.8M BTC and 350M USD till the shutdown of Mt.Gox. The Post by pseudo-name HTCFOX is the first source known to us documenting THK, and the post is further popularized in a Reddit post.

[7] Since March 2013, which was regarded as the start of the 2013 bitcoin bull run, however, all of THK's trades were BUYs and no SELLs.

corresponding trade file entries. The two variables remain empty for unverified users, take values of "!!" for users who failed verification, and for a small number of anomalous trader IDss (about 50) take value of "??". The "??" trader IDs will be used later for exchange insider detection. Table 2 breakdowns the country origins of all the transactions.

[Table 2 about here.]

Similarly, Table 3 breakdowns the currencies used by all the transactions. As Table 2 and 3 show, while Mt.Gox is headquartered in Japan and uses JPY as the clearing currency internally, US is the largest market with about 34% of total transactions, and USD is used among more than 85% of the transactions.

[Table 3 about here.]

The trader IDs included in the data are crucial as they enable us to find wash trades. There is an important caveat, however: Since there is no mandatory know-your-customer (KYC) requirement imposed by Mt.Gox (as user verification is optional), the same person may open as many accounts as possible. In another word, it is possible that multiple different trader IDs in our data are controlled by the same person or entity. This data feature biases us toward underestimating the magnitude of wash trading on Mt.Gox, but it is acceptable for our purpose of detecting direct evidence of wash trading as it only prevents us from finding results. Our later robustness analysis in Section 4 will further take this data feature into account.

2 Direct evidence of wash trading

We define a wash trade as a transaction in which both sides (buy and sell) have the same trader ID. Figure 6 plots the daily wash trading volume on Mt.Gox during our entire sample (April 1$^{\text{st}}$, 2011 to November 30$^{\text{th}}$, 2013).

[Figure 6 about here.]

As Figure 6 demonstrates, wash trading emerges immediately after the 2011 Mt.Gox hack which brought down the exchange for about a week, an incident illustrated by Figure 7.[8]

[Figure 7 about here.]

Wash trading has been prevalent ever since this incident, until a sudden stop in 2013 following the U.S. Department of Homeland Security's seizure of $5 million assets from Mt.Gox's U.S. branch. Therefore, in the rest of the paper, unless otherwise noted, we will focus our attention on the period in which wash trading was active – that is, from June 26th, 2011 to May 20, 2013, a period with 11,097,734 transactions and a total volume of 87,030,784 BTC. These numbers will be used in subsequent analyses as the denominators for calculating various percentages of wash trading. Figure 8 visualizes the order size distribution of all wash trades.

[Figure 8 about here.]

Table 4 presents the main result of this section: We find in our sample 230,550 transactions that have both sides (buy and sell) with the same trader ID, accounting for 2.1% of the total number of transactions and 1.4% of the total amount of bitcoin traded during the period between June 26th, 2011 and May 20th, 2013. Our emphasis here is on the direct evidence of wash trading presence rather than the the quantities. This is because the numbers in Table 4 are likely underestimates of the actual extent of wash trading on Mt.Gox and are better off interpreted as lower-bounds, as Mt.Gox does not impose mandatory know-your-customer (KYC) verification on its users, and thus the same person or entity may potentially register multiple trader IDs. In other words, if a person or entity conducts wash trading with

[8]See Roy (2018) for a vivid narration, where several famous figures in the bitcoin community, including Jesse Powell, who later created a major crytocurrency exchange Kraken, and Roger Ver, who later serve as the CEO of Bitcoin.com and co-created Bitcoin Cash, helped with bringing the the site back online.

multiple trader IDs it controls, such trades will not be included in our calculation. In Section 4, we will conduct several robustness analysis taking this consideration into account.

[Table 4 about here.]

The magnitude of wash trading identified in this paper is significantly lower than those estimated in other studies (e.g. Cong, Li, Tang and Yang (2020) and Bitwise (2019)). There are multiple reasons for the discrepancy: 1) Different sample period: Our sample period is within 2011-2013, which is now widely regarded as the early days in Bitcoin's history, while other studies mostly focus on periods after 2018, when Bitcoin has received wide attentions; 2) Different market environments: Mt.Gox was a dominant, and for a significant proportion of our sample period, the only Bitcoin exchange, while other studies are conducted under fierce market competitions with hundreds of exchanges. Amiram, Lyandres and Rabetti (2020) empirically establish the relationship between wash trading and market competition; 3) Different methodologies: We direct observe wash trades from leaked internal records from the exchange, and to the extent that wash trades may be conducted among several trader IDs controlled by the same individual or entity, we provide an underestimate, while other studies rely on indirect statistical inferences, which are effectively anomalous trades that may include manipulations other than explicit wash trades. That said, the main focus of this paper is not on the magnitude *per se*, but rather to provide direct, indisputable evidence of wash trading to support the many indirect inferences in other studies. We also demonstrate that wash trading is an "original sin" with Bitcoin exchanges that has been in existence far earlier than it becomes aware to the community.

Table 4 also quantifies the number of transactions among all wash trading perpetrators. Specifically, we find 2,887 unique trader IDs involved in the 230,550 wash trades and define them as *wash trader IDs*. We then find 3,642,732 transactions among these wash trader IDs during the June 26$^{\text{th}}$, 2011 and May 20$^{\text{th}}$, 2013 period. These trades account for about 32.8%

of the total number of transactions and 32.5% of the total amount of bitcoin traded during the same period. This percentage reveals active interactions among the 2,887 wash trader IDs, as they are only a tiny fraction in number out of the 98,391 total number of unique trader IDs during the same period.

3 Consequences of wash trading

Having provided direct evidence of wash trading, we further analyze the market condition that foster wash trading activities and the impact these activities in turn have on the market. Given suspicions in the Bitcoin community that exchanges may pump up volumes to either boost fee collection or to look more attractive to traders, we are particularly interested in whether wash trades tend to occur following periods of low transaction fees collected by (i.e. revenues to) Mt.Gox, and whether wash trading affects subsequent transaction fee revenues. These questions concern the joint dynamics between wash trading and the fee revenues to Mt.Gox, and we thus conduct a vector autoregression (VAR) analysis.[9] Specifically, we estimate the following model:

$$Y_t = \sum_{i=1}^{p} A_i Y_{t-i} + \epsilon_t, \text{ where}$$

$\forall i \in \{1, \cdots, p\}, A_i$ is a 2×2 matrix to be estimated, Y_t is a 2×1 vector corresponding to 1) the log wash trading volume (in units of BTC) and 2) the log fee revenues collected from non-wash transactions (in units of BTC). Figure 9 plots the time series. Both time series are stationary: they contain no time trends, and significantly reject the unit root hypothesis by the Dickey-Fuller test.

[Figure 9 about here.]

[9]The VAR approach has been widely used in financial market research related to the impact of trades on market conditions (e.g. Hasbrouck (1991)).

Figure 10 plots the impulse response functions from the VAR analysis, while Table 5 presents the VAR coefficient estimates and Granger causality test results. The VAR contains two lags based on the Bayesian information criterion and stability test. A one standard deviation increase in fee revenue tends to precede a significant over 1% standard deviation drop in wash trading. Conversely, a one standard deviation increase in wash trading is followed by a significant about 5% standard deviation increase in fee revenues. Both effects persist for the next 60 minutes.

[Figure 10 about here.]

Table 5 presents the VAR regression results. As Column 1 summarizes, wash trading Granger causes fee revenues, and vice versa. The other columns present robustness results with alternative definitions of wash trading, which will be further explained in Section 4.

[Table 5 about here.]

Overall, the evidence suggests that whenever the fee revenues to the exchange drops, wash trading appears, which boosts subsequent fee revenues. Given that the exchange itself benefits directly from a higher fee revenue, one may expect exchange insiders to directly engage in wash trading. Section 4 will follow up and confirm the exchange's involvement in wash trading.

4 Exchange insiders' engagement in wash trading

This section continues to study whether the exchange owner and other associated interested parties are among the perpetrators of wash trading. We analyze the 2,887 wash trader IDs identified in Section 2. Even though there are no data available to help us exactly match all trader IDs to their corresponding real persons or entities behind, we nevertheless are able to associate a few trader IDs with Mt.Gox itself based on additional sources.

For example, the Willy report has thoroughly gone through the leaked data and singled out a handful of trader IDs based on unusual trading behaviors and anomalies in data recording. These trader IDs have been nicknamed as Willy, Markus, MagicalTux, and Hijacker. In a related study, Gandal, Hamrick, Moore and Oberman (2018) focus on these insider IDs and provide ample evidences on their numerous price manipulations following Mt.Gox's loss of 850,000 bitcoins in a 2011 security breach, which contributed to the bitcoin price bubble in late 2013-early 2014.

We follow the spirit of Willy report and identify a sequence of (increasingly inclusive) sets of suspicious insider IDs (with decreasing confidence). We first list the following special trader IDs that have been singled out in the Willy report:

- Willy: Willy is a nickname given to a bot (or several bots) who buys a random number between 10 and 20 of bitcoins every 5-10 minutes, non-stop, for at least a month until the end of January, 2014. A number of traders reportedly first began to suspect the presence of Willy in December 2013, and a brief technical glitch that brought down the Mt.Gox site in early January 2014 finally confirmed Willy's existence (Reddit).

  In our data, all Willy's trader IDs are detected in the following two steps as instructed by Willy report:

  1. IDs 817985, 825654, 832432 are first detected by matching patterns with recognized bot behaviors. These IDs also have unusual User_Country and User_State fields: Normally, these fields contain country/state FIPS codes for verified users, or are empty ("`!!`") for unverified users (users who failed verification). However, the three detected IDs all have "`??`" in these fields.

  2. We then single out all trader IDs that have "`??`" for their User_Country and User_State fields.[10] These IDs indeed behave like what the community has iden-

[10]Specifically, 807884, 658152, 659582, 661608, 665654, 683148, 689932, 693122, 694306, 695340, 697722,

tified of Willy: They never performed a single sell, and their trades seamlessly connected to each other – when one user became inactive, the next was created usually within a few hours and actively market-buy coins until some very exact amount of USD ($2,500,000 being the most common) was spent. Such serial trading activities went back all the way to September 27th, 2013. In total, about $112 million was spent to buy close to 270,000 BTC – the bulk of which was bought in November 2013 driving the bitcoin price bubble. The very first "??" users is also a "time-traveler" – it has an usually high ID of 807884 even though regular accounts at that point only went up to around 650000.

Although we will remove Willy from our sample as its active periods are outside of June 26th, 2011 to May 20, 2013 – our interested time period during which wash trading was active (recall that the first appearance of Willy was on Sept. 27th, 2013), Willy is nevertheless important for identifying the following trader IDs that are suspected to be linked to Mt.Gox's owner:

- Markus: Inspired by Willy's "time-traveler" behavior, another time-traveler is singled out (698630, with a registered country and state: "JP", "40" – the FIPS code for Tokyo, Japan, where Mt.Gox is headquartered) and nicknamed Markus. After being active for close to 8 months, Markus became completely inactive 7 hours before the first Willy account became active, and has thus been been singled out by Nilsson (2014) and Gandal, Hamrick, Moore and Oberman (2018) to be controlled by the same entity behind Willy.

  Markus also has two peculiar attribute that further strengthens its insider status: First, it always pays zero fees. Second, related to the price correction discussed above

698233, 698232, 698234, 698235, 698236, 698237, 698238, 711137, 714565, 716004, 718998, 722068, 724340, 726910, 730861, 734205, 739116, 742432, 746069, 757154, 764692, 769205, 770436, 774567, 783273, 787018, 790503, 790667, 791191, 791965, 793833, 796081, 796083, 804879, 809401, 817985, 825654, and 832432.

in Section 1, Markus' fiat spent when buying coins always carries forward from the previous trade's fiat spent, regardless of the actual volume of BTC bought, and thus generating seemingly completely random prices paid per bitcoin. The Willy report conjectured that for Markus, the "Money" spent field is in fact empty, and the program that generates the trading logs simply carries forward whatever latest value was already there before. Finally, like Willy, Markus' trader ID 698630 was also out of place.

- MagicalTux: MagicalTux refers to the trader ID 634, singled out for the following reason: As we have pointed out in the data description in Section 1, for some months in 2013 (e.g. April 2013), there are two versions of trading logs in the leaked database: a full log and an anonymized log that differ in two and only two ways:

  1. User hashes and country/state codes in the anonymized log are removed.[11]
  2. More importantly for our analysis, Markus' out-of-place trader ID (698630) in the full log is changed to a small number (634) in the anonymized log, and its strange fixed "Money" values are corrected to the expected values.

  The second point above links ID 634 to Markus. To add to its insider status, from a 2011 leaked account list referenced in Willy report, the user with ID 634 has username "MagicalTux", which is a pseudo-name widely used by Mark Karpelès, the owner and CEO of Mt.Gox, across various venues, including for a now defunct blog of his, his Reddit account, Bitcointalk.org forum username, and his current Twitter handle.

- #1: Other than the occasional exceptions mentioned above, trader IDs are assigned to new traders in chronicle orders of registration. Therefore, traders with small IDs, and especially trader #1, the first registered trader on Mt.Gox, are most likely the exchange owner or at least closely related to the exchange owner.

[11] The Willy report conjectured that the anonymized log was created to send off to auditors or investors to show some internals.

In the rest of the paper, we will refer to the above three IDs (Markus, MagicalTux, and #1) as "Known Insiders". This category is our most conservative and confident insider designation, as they have been used in various other papers for the same purpose.

The facts that Markus pays zero fees in all his transactions, that Markus has its location in Mt.Gox's headquarter Japan, and that trader IDs are assigned incrementally as new users register, inspire us to further single out the following special trader IDs.

- Zero-fee Traders: 1032 trader IDs that never pay fees, unlike most other trader IDs.
- Hijacker: The Willy report has highlighted a subset of Zero-fee Traders, who in addition to paying zero fees in all transactions, also have location being "Japan" and trader IDs being small (<1000). These 285 IDs are so named by the Willy report as their trades demonstrate patterns suggesting that they are not executed by their original account holders, but rather "hijacked" in some way (see details here).
- Double Users: 73 trader IDs (all of which are smaller than 1000) have the same "User_ID" but different "User" values in the raw data.

In summary, we consider an increasingly inclusive sequence of insiders: 1) Known Insiders, 2) Hijackers, 3) Double Users, and 4) Zero-Fee Traders. Table 6 compares these various categories, and for each category summarizes the number of unique trader IDs in it, the number of IDs who has transactions with peers in the same category, the number of wash trader IDs in it, the number of transactions among traders in this category as well as its percentage out of all transactions in the June 26$^{\text{th}}$, 2011 to May 20$^{\text{th}}$, 2013 subperiod), and the amount of bitcoins traded among traders in this category as well as its percentage out of all transactions in the June 26$^{\text{th}}$, 2011 to May 20$^{\text{th}}$, 2013 subperiod.

Based on the insider definitions above, we confirm that exchange insiders are indeed involved in wash trading, consistent with our conjecture of the exchange itself engaging in wash trading to boost fee revenues. For example, out of the three Known Insiders, two (MagicalTux and #1) are found to participate in wash trading.

[Table 6 about here.]

We conclude this section with a robustness analysis using the various insider definitions in Table 6: Because the exchange itself may control multiple different trader IDs to clear orders among themselves, we augment the wash trades identified in Section 2 with transactions among the insiders. To the extent that those insider IDs are controlled by Mt.Gox itself, the augmented set will further include what are commonly known as painting-the-tape trades and thus more accurately reflect the extent to which wash trading prevails.

Specifically, with an increasing inclusiveness and decreasing confidence, we define various groups of (augmented) wash trades: Group 1 (G1) is the set union of all wash trades and all transactions among wash trader IDs who are also Known Insiders. G2 to G4 are similarly defined. Specifically, Group 2 (G2) is the union of G1 and all transactions among wash trader IDs who are also Hijackers; Group 3 (G3) is the union of G2 and all transactions among wash trader IDs who are also Double-Users; and Group 4 (G4) is the union of G3 and all transactions among wash trader IDs who are also Zero-Fee Traders. For ease of exposition, we also refer to the set of all wash trades as G0.

[Table 7 about here.]

Table 7 compares the various groups of augmented wash trading transactions over several dimensions, including the number of transactions counts in each set and the percentage share out of all transactions, as well as the amount of bitcoins traded within each set and the percentage share out of all transactions.

[Figure 11 about here.]

Figure 11 repeats the VAR analysis and plots the impulse responses functions as in Figure 10. Wash trading are augmented with painting-the-tape transactions among insiders, using various insider definitions explained above. Robust findings emerge that a lower fee revenue tends to be followed by more wash trading activities and an increase in wash trading precedes a higher fee revenue.

5 Evaluating indirect inference techniques

Because of the wide interest in the cryptocurrency community over volume manipulations on exchanges, in addition to our direct evidence on wash trading, a few other studies have also looked into this topic. Without access to internal trading records as we do, these other papers all resort to indirect inferences. It is then natural to ask how accurate are these indirect inferences, especially given that indirect inferences based on statistical pattern may detect patterns not solely driven by wash trading but various other trading strategies. Our internal trading records from the exchange then provides a unique opportunity to evaluate the merit of those proposed indirect inference techniques.

One indirect inference technique that has been used in both Cong, Li, Tang and Yang (2020) and Amiram, Lyandres and Rabetti (2020) takes advantage of the Benford' law (Benford (1938)), which argues that the first significant digit of order sizes should follow a logarithmic distribution benchmark. The authors reason that the extent to which the empirical distribution deviates from the benchmark (measured by a χ^2 statistic) may be useful in detecting anomalous trading behaviors. Buidiing on this assumption, they infer that no regulated exchanges yet a majority of unregulated exchanges exhibit anomalous activities, and suggest that less prominent unregulated exchange are more likely to be engaged in wash trading.

While the Benford's law has been used widely in other settings, without an evaluation against direct evidence, the validity of such indirect inferences in identifying wash trading remains an empirical question. We fill this void by taking advantage our data. Specifically, we apply the Benford's law and compare the sample of all trades and wash trades. Figure 12 illustrates the results. Consistent with the Benford's law inference, we find that in the sample of all trades, the empirical distribution does not significantly deviate from the Benford law, while the wash trades sample does.

[Figure 12 about here.]

Figure 13 repeats the analysis for different groups of insider wash trades as defined in Section 4 and find consistent results.

[Figure 13 about here.]

Overall, we provide support for the effectiveness of inferences based on the Benford's law. That said, just like any statistical inferences that permit type I and type II errors, we can identify wash trades by some individual wash traders that satisfy the Benford's law.[12] We also find non-wash traders who trades violate Benford's law.[13]

6 Conclusion

The major contribution of this paper is to provide first-hand direct evidence of wash trading on cryptocurrency exchange. While the community has only in recent years become suspicious of such market manipulations, we find that this practice actually has a much longer history. In fact, wash trading has been taking place on the world's first Bitcoin exchange, Mt.Gox, back in the early days of cryptocurrency trading. During June 26$^{\text{th}}$, 2011 to May

[12]For example, trader ID 337 has performed more than 6,332 wash trade transactions, yet its orders fully respect the Benford's law.

[13]For example, trader ID 105,326 has performed more than 35,347 transactions and no wash trade, however, its orders significantly violate the Benford's law.

20th, 2013, a relative small set of (2,887 out of 98,391, or 2.93%) trader IDs in Mt.Gox were found to be involved in 230,550 wash trading transactions. These trader IDs includes ones who have close ties with the exchange itself, and the number of transactions among them account for a disproportionately 32.8% of all transactions on Mt.Gox during the same period. Wash trading first emerges after a one-week halt of Mt.Gox's service in 2011, a time when it needs to rebuild liquidity, and stops when regulators step in for an investigation. Over time, wash trading tends to rise following low fee revenues to the exchange, and also significantly increases subsequent fee revenues. The impact is nevertheless short-lived: It typically dies out after 60 minutes.

Overall, our findings suggest that exchange insiders are involved in wash trading to boost fee revenues. This is finding is consistent with conjectures in the community. Importantly, we uncover the inter-temporal relationship between wash trading and benefits to the exchange in a market without fierce competition among exchanges, complementing the cross-sectional inferences in Amiram, Lyandres and Rabetti (2020).

We further evaluate some indirect inference techniques proposed in the literature. In particular, we evaluate the Benford's law, which has been in both Cong, Li, Tang and Yang (2020) and Amiram, Lyandres and Rabetti (2020). Future work can further evaluate other indirect inferences proposed in the literature.[14] We hope our first-hand direct evidence can shed new light on the community's suspicions over fake volume in cryptocurrency exchanges as well as the heated debate on how to regulate cryptocurrency exchanges, and to help develop ways to safeguard the healthy development of the cryptocurrency market.

[14] In general, we believe any future indirect inferences of wash trade would benefit from being evaluated against the Mt.Gox data as we do in our paper, given that our data is so far the only individual-level information available to the community.

References

Aggarwal, Rajesh K, and Guojun Wu. 2006. "Stock market manipulations." *The Journal of Business*, 79(4): 1915–1953. 7

Ahern, Kenneth R. forthcoming. "Do Proxies for Informed Trading Measure Informed Trading? Evidence from Illegal Insider Trades." *The Review of Asset Pricing Studies*. 7

Alexander, Gordon J, and Mark A Peterson. 2007. "An analysis of trade-size clustering and its relation to stealth trading." *Journal of Financial Economics*, 84(2): 435–471. 3

Allen, Franklin, and Douglas Gale. 1992. "Stock-price manipulation." *The Review of Financial Studies*, 5(3): 503–529. 7

Amiram, Dan, Evgeny Lyandres, and Daniel Rabetti. 2020. "Competition and Product Quality: Fake Trading on Crypto Exchanges." *Available at SSRN 3745617*. 3, 13, 21, 23

Augustin, Patrick, Alexey Rubtsov, and Donghwa Shin. 2020. "The impact of derivatives on cash markets: Evidence from the introduction of bitcoin futures contracts." *Available at SSRN*. 7

Benetton, Matteo, and Giovanni Compiani. 2020. "Investors' Beliefs and Asset Prices: A Structural Model of Cryptocurrency Demand." *University of Chicago, Becker Friedman Institute for Economics Working Paper*, , (2020-107). 7

Benford, Frank. 1938. "The law of anomalous numbers." *Proceedings of the American philosophical society*, 551–572. 21

Biais, Bruno, Christophe Bisiere, Matthieu Bouvard, Catherine Casamatta, and Albert J Menkveld. 2018. "Equilibrium bitcoin pricing." *Available at SSRN 3261063*. 7

Bitwise. 2019. "Bitwise Asset Management Presentation to the U.S. Securities and Exchange Commission." 3, 13

Bollen, Nicolas PB, and Veronika K Pool. 2009. "Do hedge fund managers misreport returns? Evidence from the pooled distribution." *The Journal of Finance*, 64(5): 2257–2288. 7

Chakrabarty, Bidisha, Pamela C Moulton, Leonid Pugachev, and Xu Frank Wang. 2020. "Catch Me If You Can: Improving the Scope and Accuracy of Fraud Prediction." *Available at SSRN 3352667*. 8

Choi, Kyoung Jin, Alfred Lehar, and Ryan Stauffer. 2020. "Bitcoin microstructure and the kimchi premium." *Available at SSRN 3189051*. 7

Cong, Lin William, Xi Li, Ke Tang, and Yang Yang. 2020. "Crypto Wash Trading." *working paper*. 3, 13, 21, 23

Cumming, Douglas, Sofia Johan, and Dan Li. 2011. "Exchange trading rules and stock market liquidity." *Journal of Financial Economics*, 99(3): 651–671. 7

Daian, Philip, Steven Goldfeder, Tyler Kell, Yunqi Li, Xueyuan Zhao, Iddo Bentov, Lorenz Breidenbach, and Ari Juels. 2020. "Flash Boys 2.0: Frontrunning in Decentralized Exchanges, Miner Extractable Value, and Consensus Instability." *2020 IEEE Symposium on Security and Privacy (SP)*, 566–583. 3

Dimmock, Stephen G, William C Gerken, and Nathaniel P Graham. 2018. "Is fraud contagious? Coworker influence on misconduct by financial advisors." *The Journal of Finance*, 73(3): 1417–1450. 7

Egan, Mark, Gregor Matvos, and Amit Seru. 2019. "The market for financial adviser misconduct." *Journal of Political Economy*, 127(1): 233–295. 7

Ellis, Katrina, Roni Michaely, and Maureen O'Hara. 2000. "The accuracy of trade classification rules: Evidence from Nasdaq." *Journal of Financial and Quantitative Analysis*, 529–551. 7

Feder, Amir, Neil Gandal, J.T. Hamrick, and Tyler Moore. 2017. "The impact of DDoS and other security shocks on Bitcoin currency exchanges: evidence from Mt. Gox." *Journal of Cybersecurity*, 3, 2: 137–144. 9, 10, 34

Foley, Sean, Jonathan R Karlsen, and Tālis J Putniņš. 2019. "Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies?" *The Review of Financial Studies*, 32(5): 1798–1853. 7

Gandal, Neil, JT Hamrick, Tyler Moore, and Tali Oberman. 2018. "Price manipulation in the Bitcoin ecosystem." *Journal of Monetary Economics*, 95: 86–96. 4, 6, 16, 17

Griffin, John M, and Amin Shams. 2018. "Manipulation in the VIX?" *The Review of Financial Studies*, 31(4): 1377–1417. 7

Griffin, John M, and Amin Shams. 2020. "Is Bitcoin really untethered?" *The Journal of Finance*, 75(4): 1913–1964. 6

Hamrick, JT, Farhang Rouhi, Arghya Mukherjee, Amir Feder, Neil Gandal, Tyler Moore, and Marie Vasek. 2018. "The economics of cryptocurrency pump and dump schemes." *CEPR Discussion Paper No. DP13404*. 7

Hasbrouck, Joel. 1991. "Measuring the information content of stock trades." *The Journal of Finance*, 46(1): 179–207. 14

Karpoff, Jonathan M, Allison Koester, D Scott Lee, and Gerald S Martin. 2017. "Proxies and databases in financial misconduct research." *The Accounting Review*, 92(6): 129–163. 7

Kedia, Simi, Kevin Koh, and Shivaram Rajgopal. 2015. "Evidence on contagion in earnings management." *The Accounting Review*, 90(6): 2337–2373. 7

Kuo, Wei-Yu, Tse-Chun Lin, and Jing Zhao. 2015. "Cognitive limitation and investment performance: Evidence from limit order clustering." *The Review of Financial Studies*, 28(3): 838–875. 36

Lie, Erik. 2005. "On the timing of CEO stock option awards." *Management Science*, 51(5): 802–812. 7

Li, Jiasun, and Guanxi Yi. 2019. "Toward a factor structure in crypto asset returns." *The Journal of Alternative Investments*, 21(4): 56–66. 7

Li, Jiasun, and William Mann. 2018. "Digital tokens and platform building." 7

Li, Tao, Donghwa Shin, and Baolian Wang. 2019. "Cryptocurrency pump-and-dump schemes." *Available at SSRN 3267041*. 7

Liu, Xiaoding. 2016. "Corruption culture and corporate misconduct." *Journal of Financial Economics*, 122(2): 307–327. 7

Liu, Yukun, Aleh Tsyvinski, and Xi Wu. 2019. "Common risk factors in cryptocurrency." National Bureau of Economic Research. 7

Liu, Yukun, and Aleh Tsyvinski. forthcoming. "Risks and Returns of Cryptocurrency." *The Review of Financial Studies*. 7

Makarov, Igor, and Antoinette Schoar. 2020. "Trading and arbitrage in cryptocurrency markets." *Journal of Financial Economics*, 135(2): 293–319. 7

Nilsson, Kim. 2014. "The Willy Report." *https://willyreport.wordpress.com/*. 4, 10, 17

O'Donovan, James, Hannes F Wagner, and Stefan Zeume. 2019. "The value of offshore secrets: evidence from the panama papers." *The Review of Financial Studies*, 32(11): 4117–4155. 7

Parsons, Christopher A, Johan Sulaeman, and Sheridan Titman. 2018. "The geography of financial misconduct." *The Journal of Finance*, 73(5): 2087–2137. 7

Ribes, Sylvain. 2018. "Chasing Fake Volume: A Crypto-Plague." *https://medium.com/@sylvainartplayribes/chasing-fake-volume-a-crypto-plague-ea1a3c1e0b5e*. 3

Ritter, Jay R. 2008. "Forensic finance." *Journal of Economic Perspectives*, 22(3): 127–47. 7

Roy, Levin. 2018. "Mt Gox Hack Explained: Full History & Information Guide." *https://totalcrypto.io/mt-gox/*. 12

Schwenkler, Gustavo, and Hannan Zheng. 2021. "Peer Co-Movement in Crypto Markets." 7

Shams, Amin. 2020. "The structure of cryptocurrency returns." *Charles A. Dice Center Working Paper*, , (2020-11). 7

Tibeaudou, M. 2019. "Problems with the Bitwise Fake Volume Report." *https://hedgetrade.com/bitwise-exchange-report-raises-questions/*. 3

Underwood, BD. 2018. "Virtual Market Integrity Initiative Report." *Report, Office of the New York State Attorney General, Albany.* 2

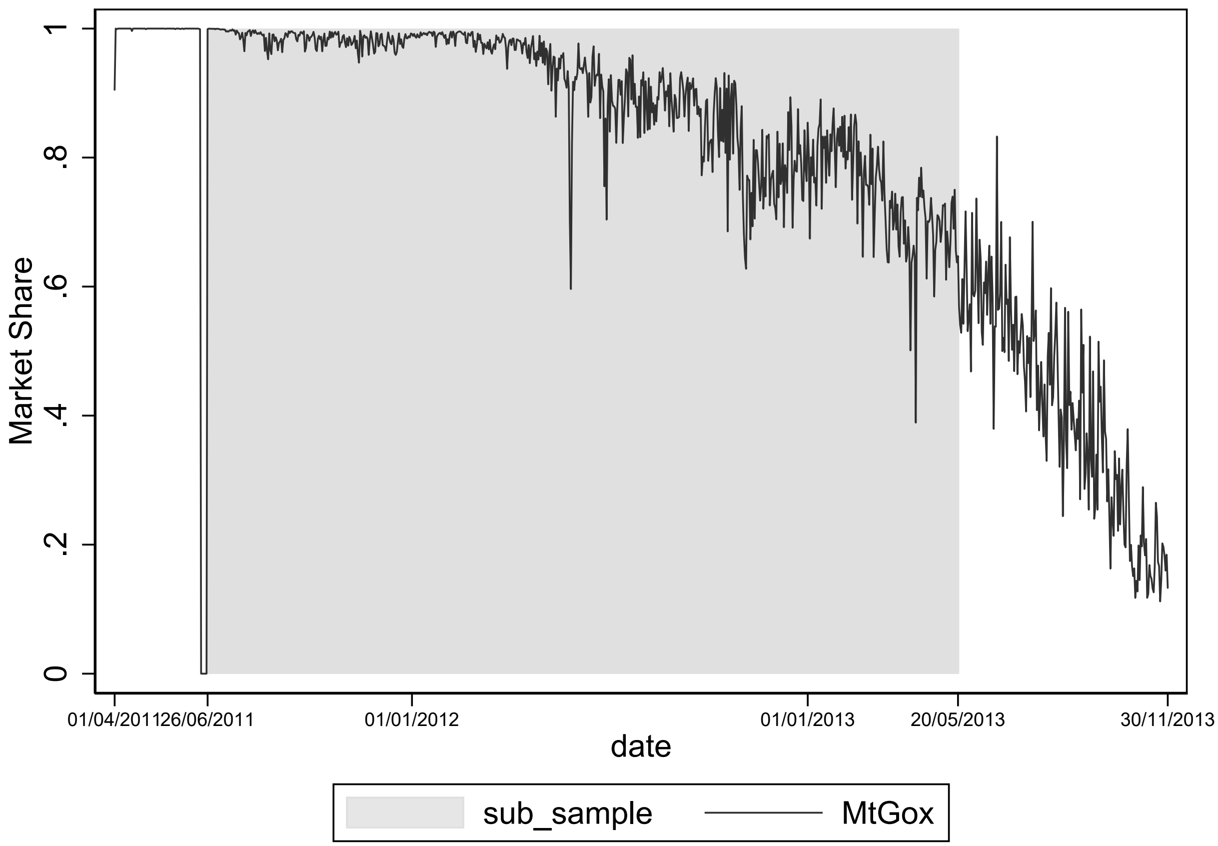
Xu, Jiahua, and Benjamin Livshits. 2019. "The anatomy of a cryptocurrency pump-and-dump scheme." *Proceedings of 28th USENIX Security Symposium*, 1609–1625. 7

Yu, Yang Gloria, and Jinyuan Zhang. 2020. "Flight to Bitcoin." *Jinyuan, Flight to Bitcoin (Oct 26, 2020)*. 7

Zitzewitz, Eric. 2012. "Forensic economics." *Journal of Economic Literature*, 50(3): 731–69. 7

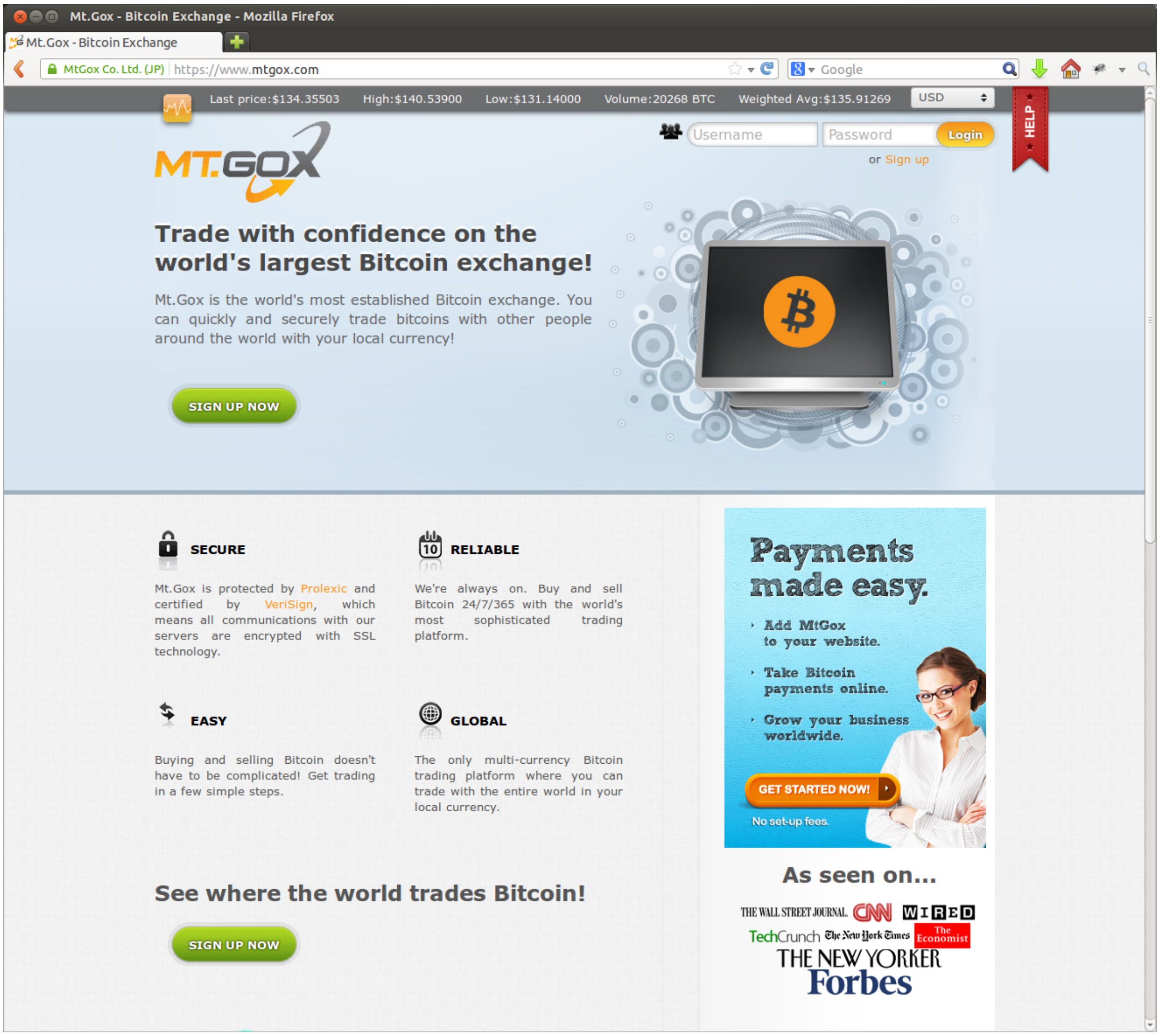
Figures

Figure 1: Mt.Gox's Market Share over Time



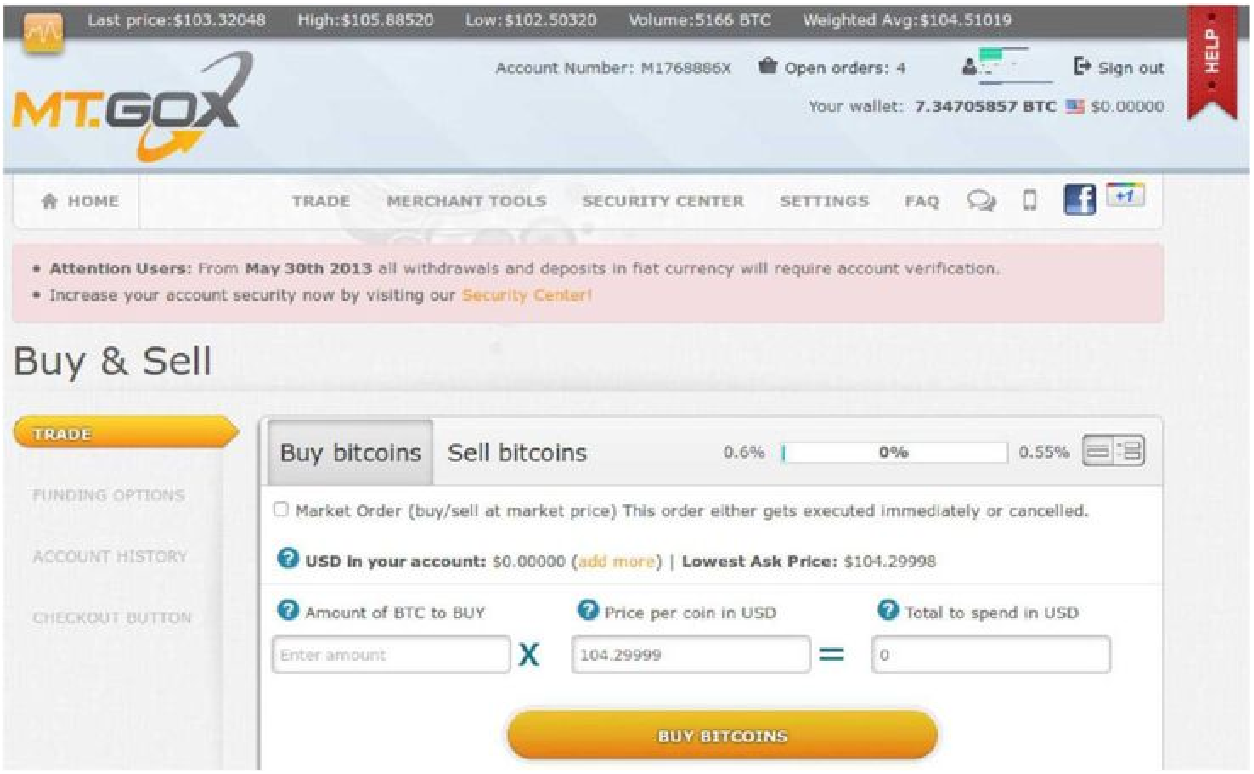
This figure presents Mt. Gox's market shares by trading volume over our entire sample period (April 1st, 2011 to November 30th, 2013). The shaded area corresponds to the sub sample period from June 26th, 2011 to May 20th, 2013, during which wash trading was active. We notice that Mt.Gox is the largest, and for about half of our sample period, the monopolistic bitcoin exchange. Data for daily trading volume are obtained from Bitcoinity. However, following a June 2011 hack of Mt. Gox (as indicated by the sudden drop of market share to 0 right before the shaded sub-period), Bitcoinity reports missing volumes for many days, and we thus replace the Mt. Gox trading volume data from June 26th, 2011 to Novermber 29th, 2011 using data from BitcoinCharts. Our calculations are consistent with third-party estimates, such as those reported in Medium.

Figure 2: The Mt.Gox Interface



This figure illustrates Mt.Gox's user interface with various screenshots, including the welcome page, trading interface, fee calculator (which illustrates the bitcoin fees and fiat fees), the notice put forward when Mt.Gox collapsed in February 2014, and an account summary on a desktop terminal (which shows the parent company information, Tibanne Co. Ltd). Mt.Gox effectively operates like a limit order book, and its user interface is similar to those for trading other assets (e.g. online stock brokerages).

Figure 2: The Mt.Gox Interface, Continued



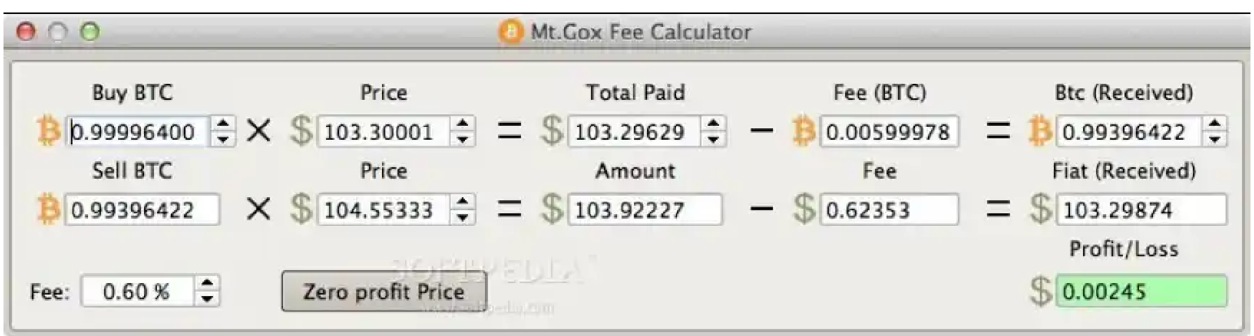


Figure 2: The Mt.Gox Interface, Continued

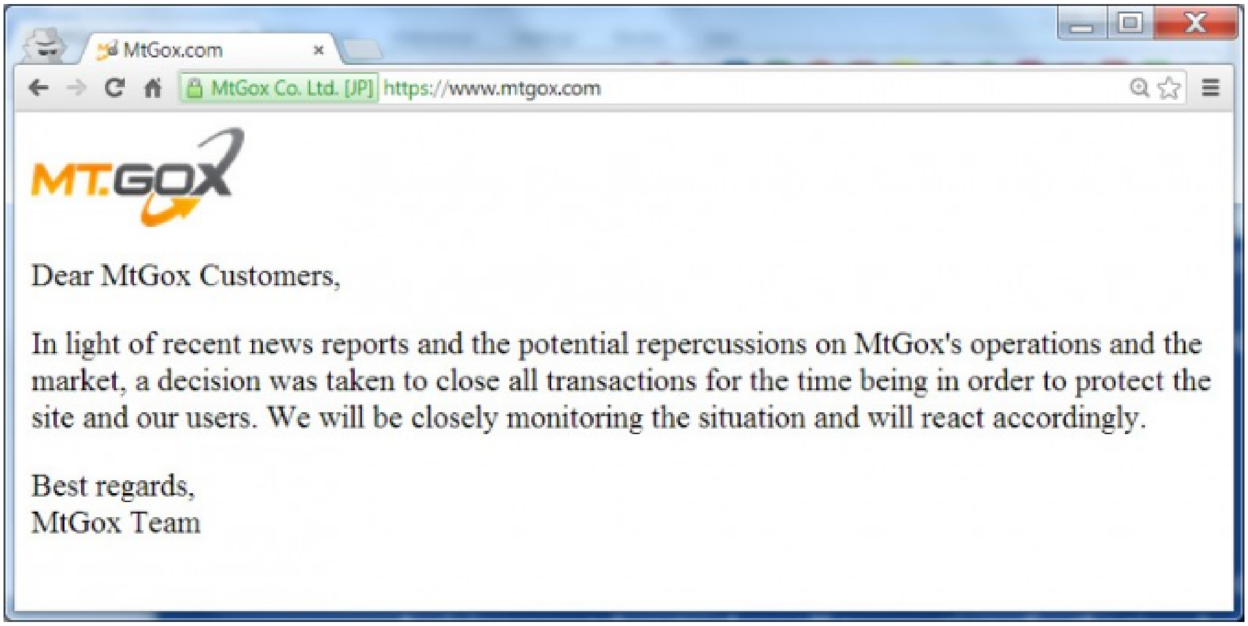
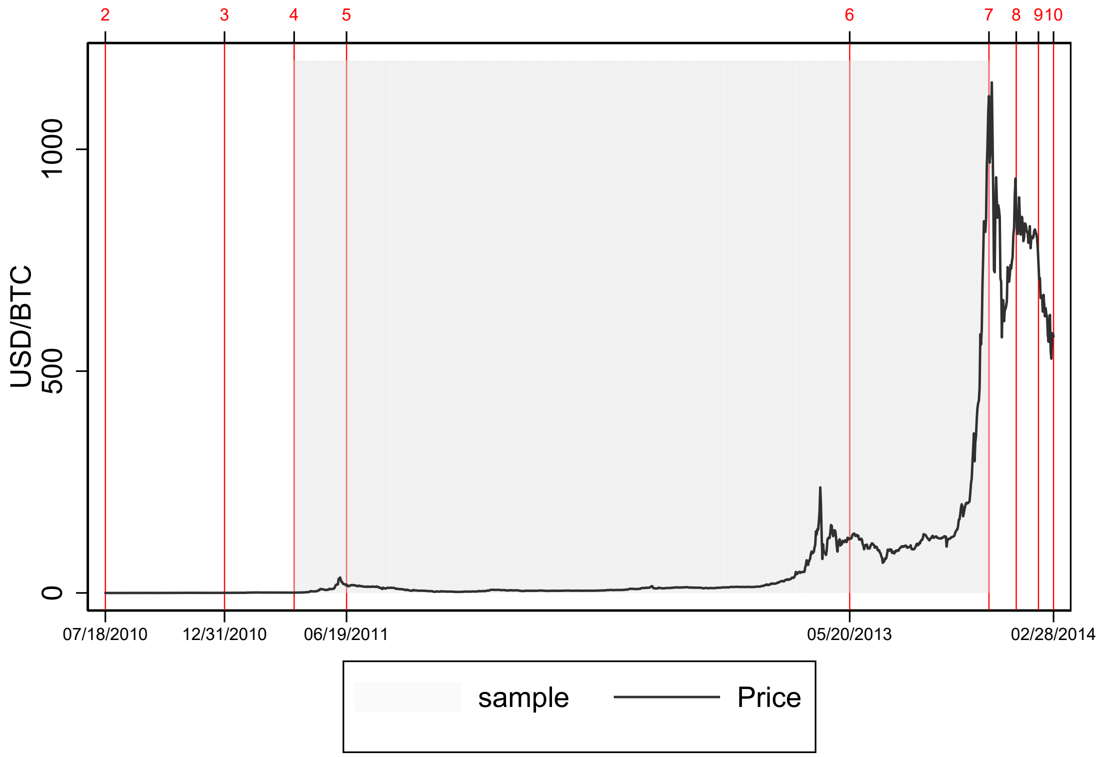


Figure 3: Timeline and Historical Events



This figure plots the evolution of bitcoin prices during our sample period. Our sample covers the period corresponding to the shaded area, which by now has been viewed as the early days in the history of bitcoin trading. We also list several major historical events surrounding our sample (red vertical dashed lines), including:

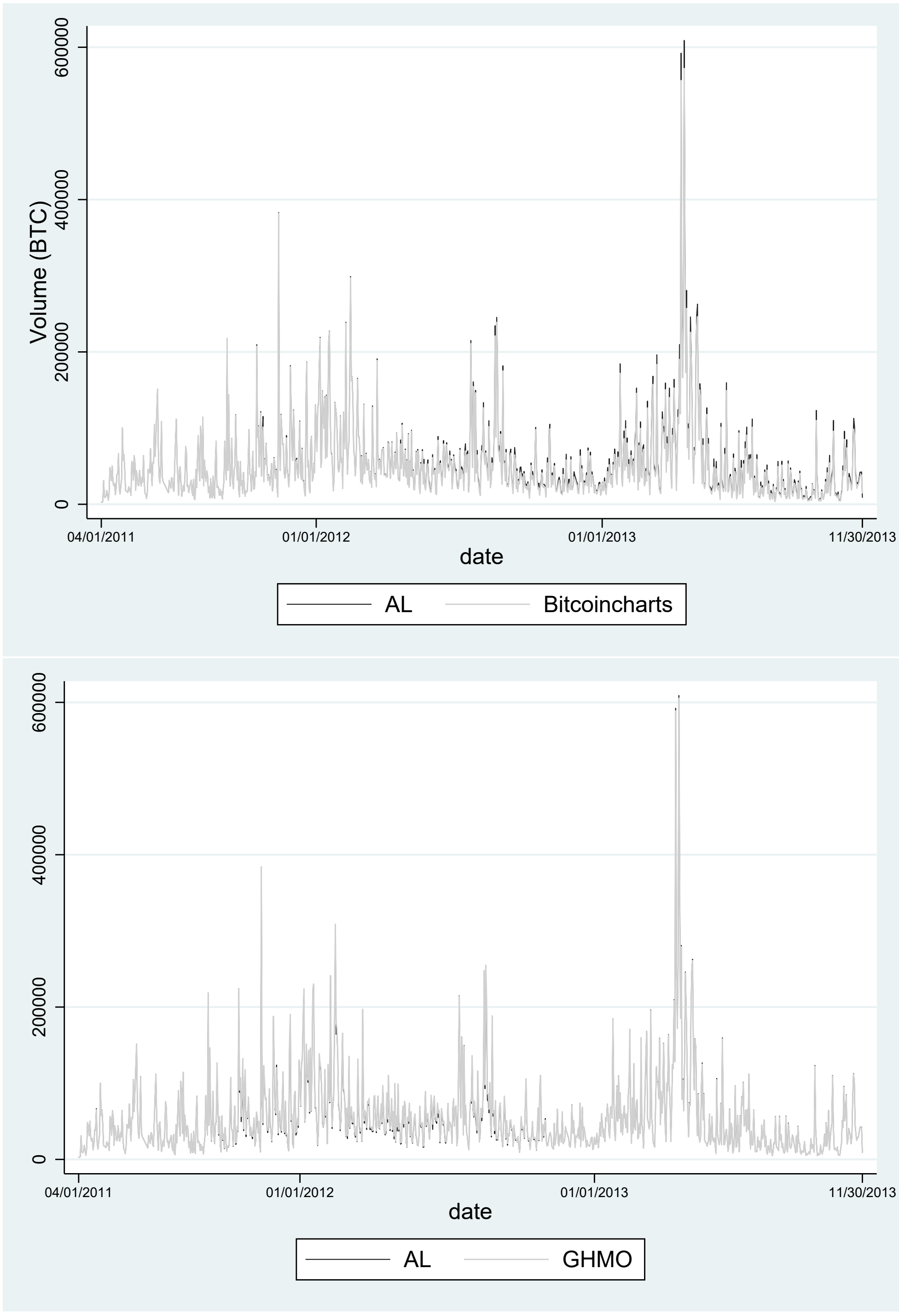
1. late 2007: Jed McCaleb founded "Magic: The Gathering Online Exchange" (Mt.Gox) as a trading venue for the card game "Magic: The Gathering".
2. July 18, 2010: Mt.Gox started quoting prices of Bitcoin. Popularity boomed.
3. *circa* the end of 2010: Mark Karpelés bought 88% share of Mt.Gox from McCaleb and revamped the website. McCaleb later went on to create the payment startup Ripple.
4. April 1$^{\text{st}}$, 2011: Sample starts.
5. June 19 2011: Mt.Gox hacked; site offline for about a week (further illustrated in Figure 7).
6. May 20, 2013: Coinlab sued Mt.Gox claiming \$75M over a non-materialized deal on the exchange's US based customers, followed by seizure of around \$5M from the company's bank accounts by U.S. Department of Homeland Security in the investigation.
7. November 30$^{\text{th}}$, 2013: end of the sample.
8. January 7, 2014: the presence of trading bot Willy was confirmed during a Mt Gox site glitch.
9. February 7, 2014: Mt.Gox froze all Bitcoin withdrawals.
10. February 28, 2014: Mt.Gox filed for bankruptcy, with 744,408 bitcoins belonging to customers and 100,000 to the company "lost".

Figure 4: Snapshots of Original Files

Trade_Id	Date	User_Id	Japan	Type	Currency	Bitcoins	Money	Money_Rate	Money_JPY	Money_Fee	Money_Fee_Rate	Money_Fee_JPY	Bitcoin_Fee	Bitcoin_Fee_JPY	User	User_Id_Hash	User_Country	User_State
1322952858139990	03/12/2011 22:54	634	JP	buy	USD	34274.76049	99396.80543	77.89948931	7742960.382	0	77.89948931	0	89.11437728	21896.31892				
1322952858139990	03/12/2011 22:54	1	NJP	sell	USD	34274.76049	99396.80543	77.89948931	7742960.382	0	77.89948931	0	0	0	3fff3b8a-8bb0-457b-9f90-50dacbe13e75		US	
1321241568352270	14/11/2011 03:32	634	JP	buy	USD	20000	50000.2	77.35531136	3867781.039	0	77.35531136	0	54	12438.61257				
1321241568352270	14/11/2011 03:32	634	JP	sell	USD	20000	50000.2	77.35531136	3867781.039	135.00054	77.35531136	10443.0088	0	0				
1328428881285630	05/02/2012 08:01	634	JP	buy	USD	20000	114000	76.21580547	8688601.824	0	76.21580547	0	56	25618.24401				
1328428881285630	05/02/2012 08:01	89169	NJP	sell	USD	20000	114000	76.21580547	8688601.824	285	76.21580547	21721.50456	0	0				
1324405083671670	20/12/2011 18:18	4425	NJP	buy	USD	12127.8276	47298.52763	77.8587315	3682603.363	0	77.8587315	0	36.38348279	9819.622951	73232106-96a9-41b0-95fc-5323a9fa84bc		US	CA
1324405083671670	20/12/2011 18:18	1	NJP	sell	USD	12127.8276	47298.52763	77.8587315	3682603.363	0	77.8587315	0	0	0	3fff3b8a-8bb0-457b-9f90-50dacbe13e75		US	
1318007613288910	07/10/2011 17:13	27551	NJP	buy	USD	10029.47837	44882.91864	76.77292938	3445793.143	0	76.77292938	0	30.0884351	10997.87902	dacee2cc-a4a2-476e-b7f9-f7bed86fd320		US	CA
1318007613288910	07/10/2011 17:13	1	NJP	sell	USD	10029.47837	44882.91864	76.77292938	3445793.143	0	76.77292938	0	0	0	3fff3b8a-8bb0-457b-9f90-50dacbe13e75		US	
1312674598617870	06/08/2011 23:49	440	NJP	buy	USD	10000	60000	78.59413635	4715648.181	0	78.59413635	0	30	13255.61999	64c56917-3a4d-4cbf-9f24-a01faa84df5e		JP	40
1312674598617870	06/08/2011 23:49	6771	NJP	sell	USD	10000	60000	78.59413635	4715648.181	180	78.59413635	14146.94454	0	0				
1314383542606940	26/08/2011 18:32	440	NJP	buy	USD	10000	85000	77.16999445	6559449.529	0	77.16999445	0	29	12820.67678	64c56917-3a4d-4cbf-9f24-a01faa84df5e		JP	40
1314383542606940	26/08/2011 18:32	1	NJP	sell	USD	10000	85000	77.16999445	6559449.529	0	77.16999445	0	0	0	3fff3b8a-8bb0-457b-9f90-50dacbe13e75		US	
1342490441668310	17/07/2012 02:00	74129	NJP	buy	USD	10000	90000	79.042	7113821.138	0	79.042	0	26	16535.615	11f1848d-bab2-4c31-b8e4-cf257e6ca239		US	TX
1342490441668310	17/07/2012 02:00	74129	NJP	sell	USD	10000	90000	79.042	7113821.138	234	79.042	18495.935	0	0	11f1848d-bab2-4c31-b8e4-cf257e6ca239		US	TX
1321238761494800	14/11/2011 02:46	89169	NJP	buy	USD	9994.348607	25985.30638	77.35531136	2010101.466	0	77.35531136	0	25.98530638	5985.577008				
1321238761494800	14/11/2011 02:46	634	JP	sell	USD	9994.348607	25985.30638	77.35531136	2010101.466	70.16033	77.35531136	5427.274172	0	0				
1332197220238560	19/03/2012 22:47	105211	NJP	buy	USD	9943.130713	9943.130713	83.72217139	4468406.727	0	83.72217139	0	24.85782678	11171.01681				
1332197220238560	19/03/2012 22:47	105211	NJP	sell	USD	9943.130713	9943.130713	83.72217139	4468406.727	113.98855	83.72217139	9543.36892	0	0				
1321241066569310	14/11/2011 03:24	1	NJP	buy	USD	9701.376317	24253.53781	77.35531136	1876139.969	0	77.35531136	0	0	0	3fff3b8a-8bb0-457b-9f90-50dacbe13e75		US	
1321241066569310	14/11/2011 03:24	634	JP	sell	USD	9701.376317	24253.53781	77.35531136	1876139.969	65.48455	77.35531136	5065.577754	0	0				
1328240519702430	03/02/2012 03:41	991	NJP	buy	USD	9468.334363	56431.2728	76.11119597	4295051.663	0	76.11119597	0	26.51133622	12338.85849	6c9081f1-e6e1-4184-963c-4ee50c565669		US	UT
1328240519702430	03/02/2012 03:41	634	JP	sell	USD	9468.334363	56431.2728	76.11119597	4295051.663	158.00756	76.11119597	12026.14436	0	0				
1321516000458840	17/11/2011 07:46	634	JP	buy	USD	9104.08065	20028.97743	76.957876	1541387.561	0	76.957876	0	23.67060969	4471.933962				
1321516000458840	17/11/2011 07:46	231	NJP	sell	USD	9104.08065	20028.97743	76.957876	1541387.561	0	76.957876	0	0	0	0837e260-175b-40a6-9094-00e32899eabf		JP	!!
1321654820470670	18/11/2011 22:20	634	JP	buy	USD	8952.435932	18800.11546	76.97329377	1447106.81	0	76.97329377	0	23.27633342	4113.208024				
1321654820470670	18/11/2011 22:20	231	NJP	sell	USD	8952.435932	18800.11546	76.97329377	1447106.81	0	76.97329377	0	0	0	0837e260-175b-40a6-9094-00e32899eabf		JP	!!
1329108193108430	13/02/2012 04:43	74129	NJP	buy	USD	8938.393787	46569.03163	77.66320419	3616700.212	0	77.66320419	0	24.13366322	10726.88706	11f1848d-bab2-4c31-b8e4-cf257e6ca239		US	TX
1329108193108430	13/02/2012 04:43	634	JP	sell	USD	8938.393787	46569.03163	77.66320419	3616700.212	125.73639	77.66320419	9765.09093	0	0				
1327618928008280	26/01/2012 23:02	634	JP	buy	USD	8744.03875	47217.80925	78.05594189	3685630.575	0	78.05594189	0	25.35771238	11936.23831				
1327618928008280	26/01/2012 23:02	6288	NJP	sell	USD	8744.03875	47217.80925	78.05594189	3685630.575	188.87124	78.05594189	14742.52254	0	0	fb1f5b82-b1d6-4ace-9848-1ac33ff8315c		HR	21
1337306425429550	18/05/2012 02:00	6534	NJP	buy	USD	8538.5196	43546.44996	80.24759502	3494497.881	0	80.24759502	0	24.76170684	9931.030168	22fa9cb7-edfa-475e-a598-7b6b0d147ea5		!!	
1337306425429550	18/05/2012 02:00	104510	NJP	sell	USD	8538.5196	43546.44996	80.24759502	3494497.881	187.24973	80.24759502	15026.3405	0	0				

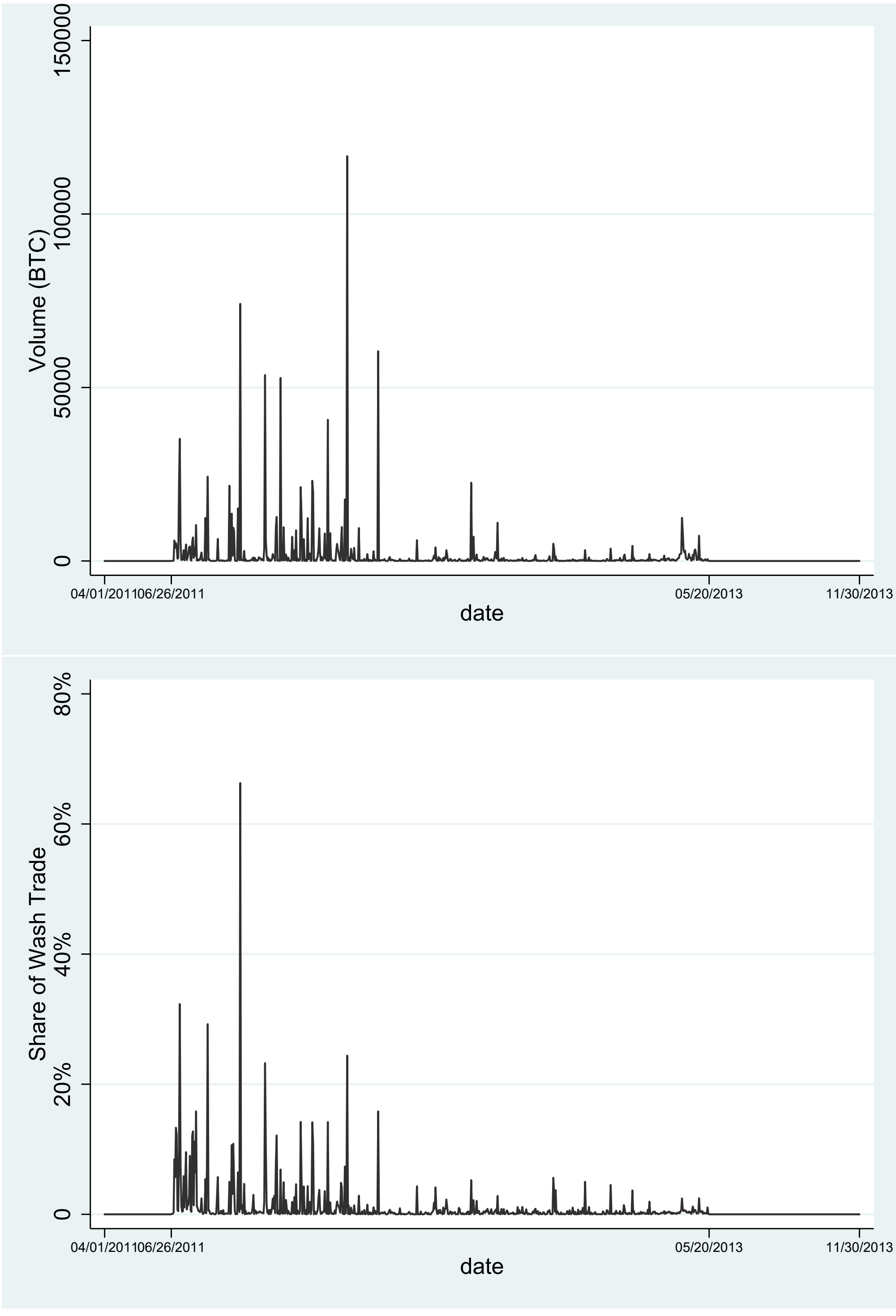
This figure presents a snapshot of the original data sorted by order sizes. The raw data contain information for each transaction a trade identifier, date and time, the user ID of the trader, whether the trade is in Japanese or not, whether it is a buy or sell leg, fiat currency involved, order size (in units of bitcoins and fiat currencies, as well as the converted value in Japanese Yen), the amount of fees paid (either in bitcoin or in fiat currency), country and location of the trader (for verified users only). We highlight several examples of wash trades among the largest transactions in red. These pairs of transactions share the same trade identifier, and have the same user ID as both the buyer and seller at two legs of the same transaction.

Figure 5: Verification of the De-duplication Procedure



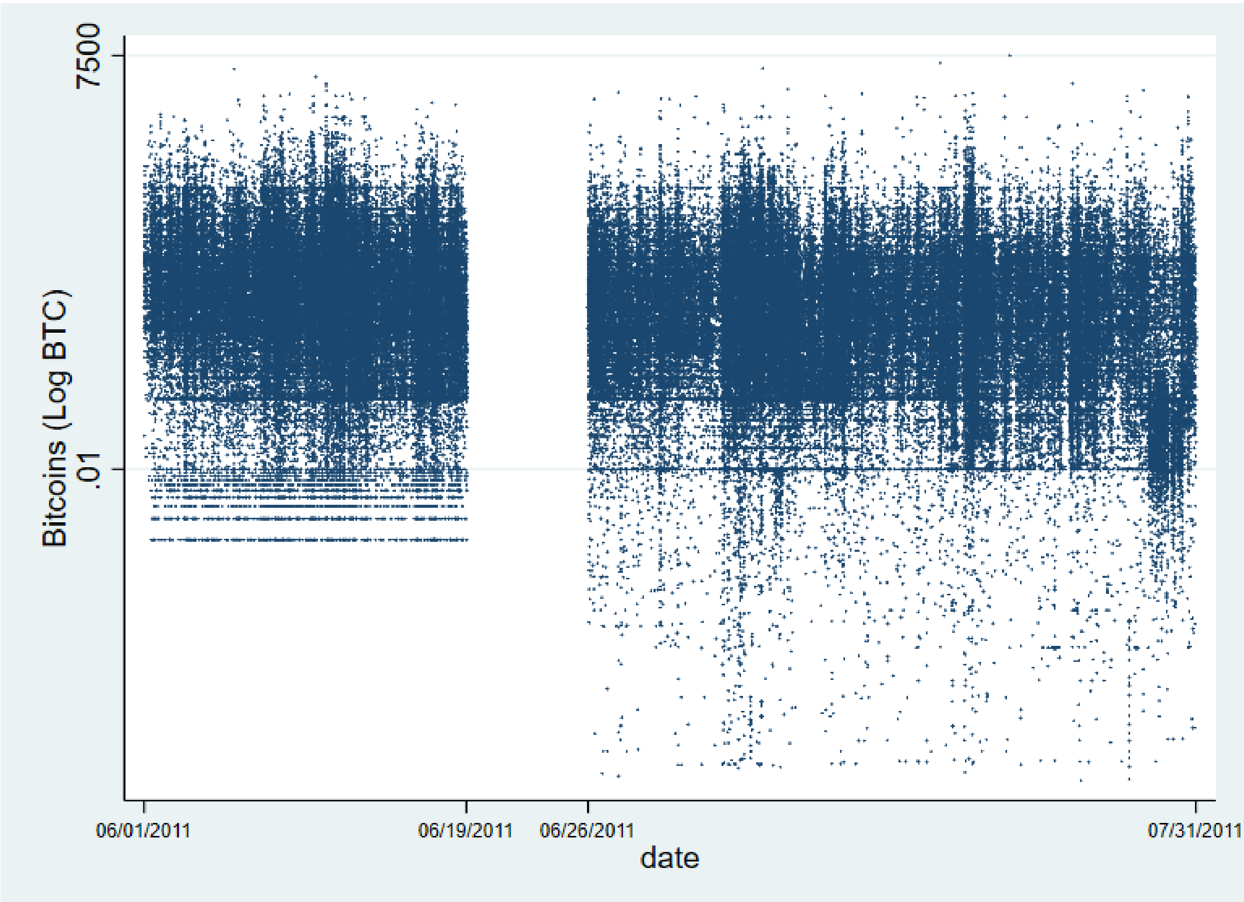
This figure compares the daily volume calculated from our de-duplication process (AL) with 1) that from Bitcoincharts in the top panel and 2) that from Feder, Gandal, Hamrick and Moore (2017) (FGHM) in the bottom panel, and confirm that our de-duplication process leads to volumes closely approximating those from external sources and FGHM.

Figure 6: Wash Trading Overtime



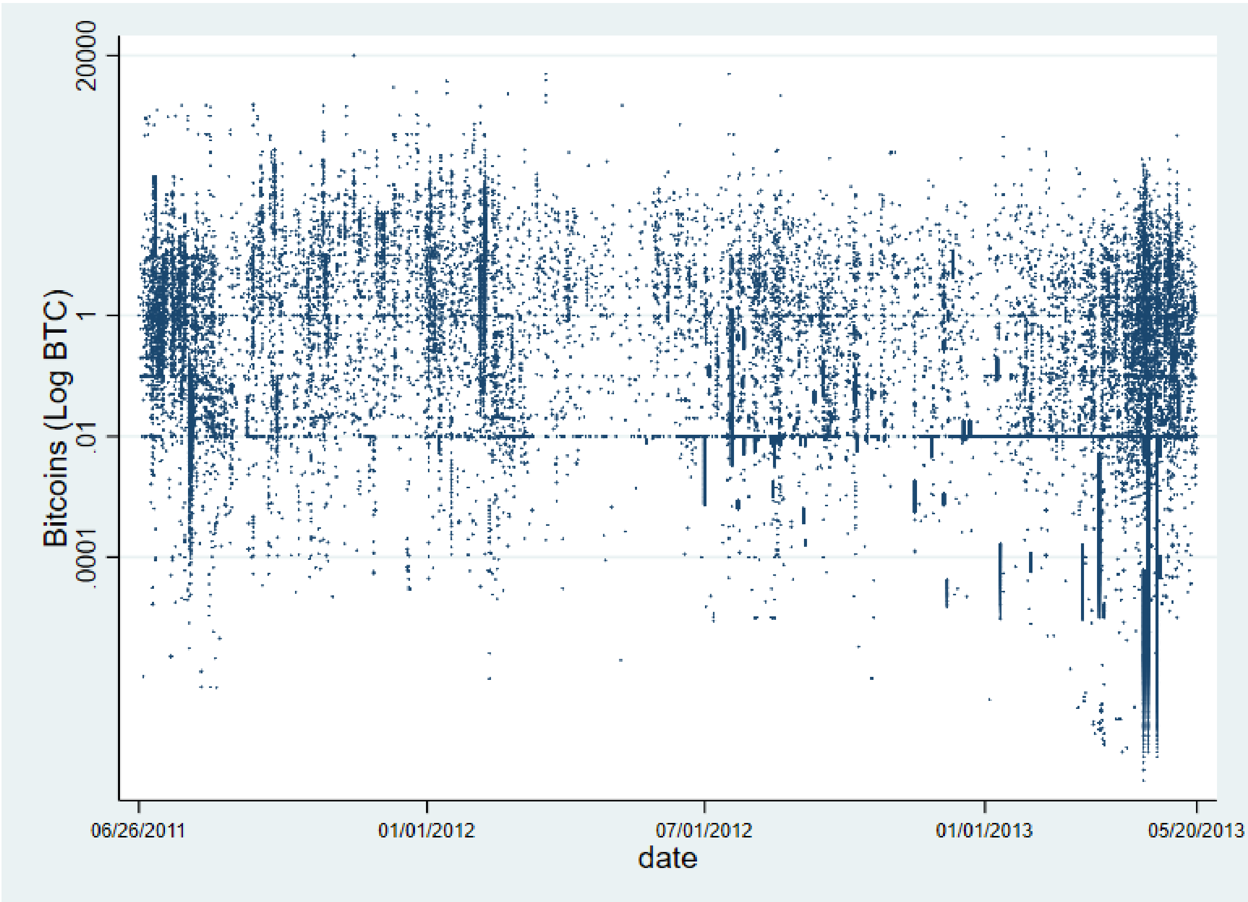
This figure plots the daily dollar volume of wash trading over time. The top panel plots the daily wash trading volume (in BTC) on Mt.Gox, and the bottom panel plots the daily share of wash trading volumes within total daily trading volumes on Mt.Gox. We see that wash trading first arise circa June 26$^{\text{th}}$, 2011 (after a technical glitch brought down Mt.Gox for about a week). Wash trading is prevalent until May 20, 2013, when Mt.Gox faced an investigation from the regulator.

Figure 7: The June 19th 2011 Hack on Mt.Gox that Halts Trading for a Week



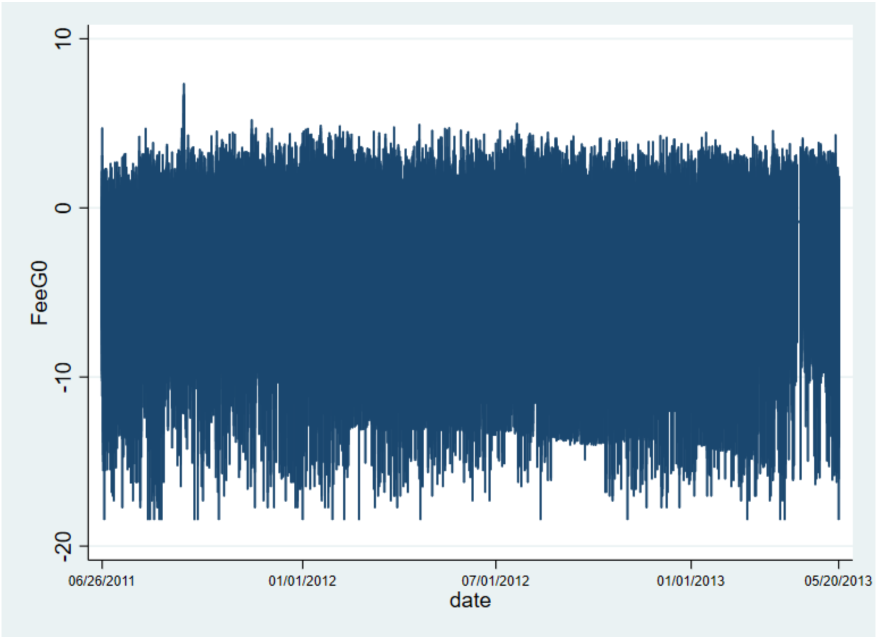
This figure visualizes the order size distribution surrounding the June 19th 2011 hack. The horizontal axis is time and the vertical axis is order size in logarithm scale. Trading halts for about a week following the hack. It is after this hack that wash trading starts to emerge on Mt.Gox. Note that there are significant trade size clustering, which is also a well-recognized trader behavior in other markets (See e.g. Kuo, Lin and Zhao (2015)).

Figure 8: Wash Trading Patterns

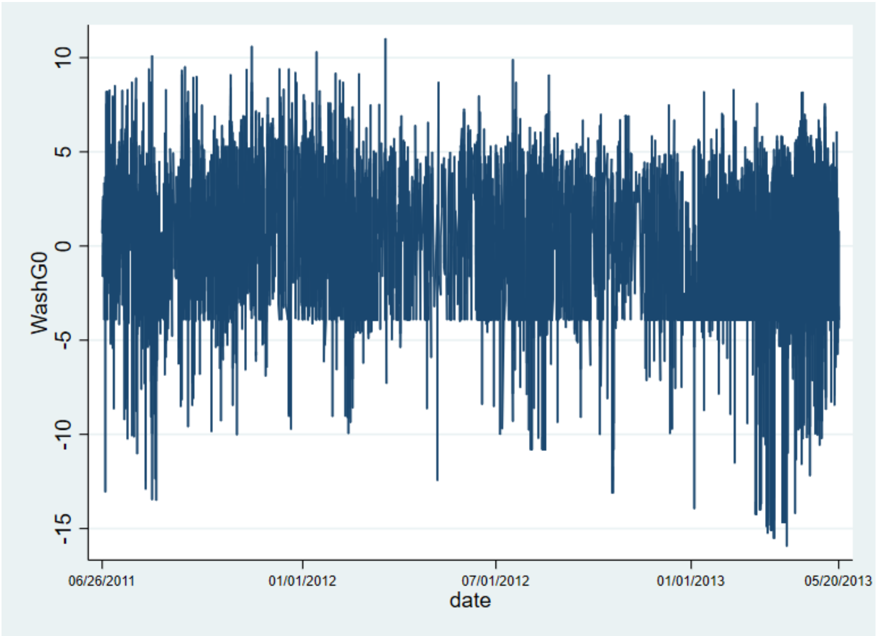


This figure visualizes the order size distribution of all wash trades over time. The horizontal axis is time and the vertical axis is order size in logarithm scale. Wash trading started on June 26th 2011 (one week after the attack on Mt Gox on June 19th 2011) and ended on May 20th 2013 (five days after issuing a warrant by the US Department of Homeland Security (DHS) to seize money from Mt.Gox's U.S. subsidiary's account with payment processor Dwolla. Note that there are still significant order size clustering even among wash trades, suggesting that some wash traders are smart to mimic regular trader behaviors.

Figure 9: Time-series for VAR



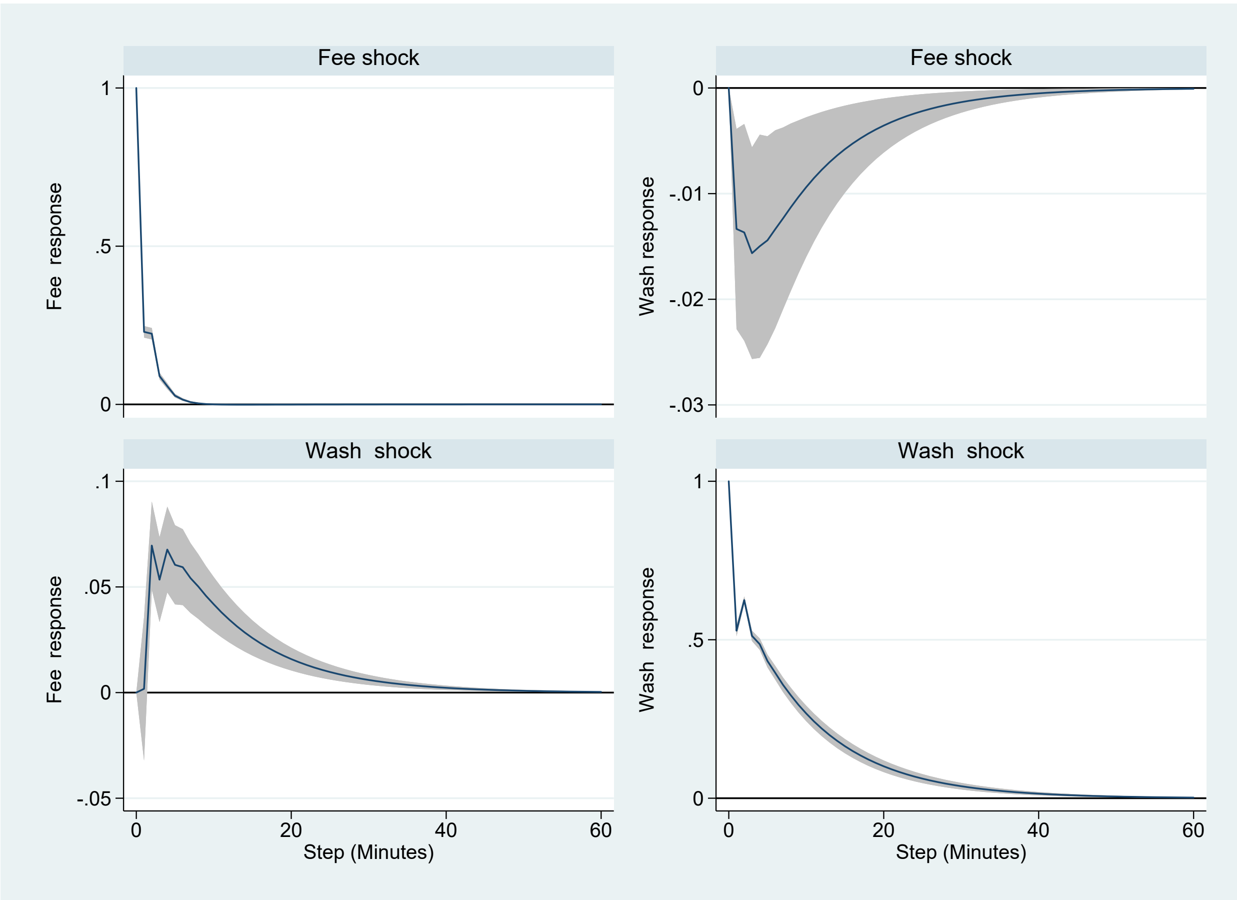
$\log(\text{Fee Revenue}_t)$



$\log(\text{Wash Volume}_t)$

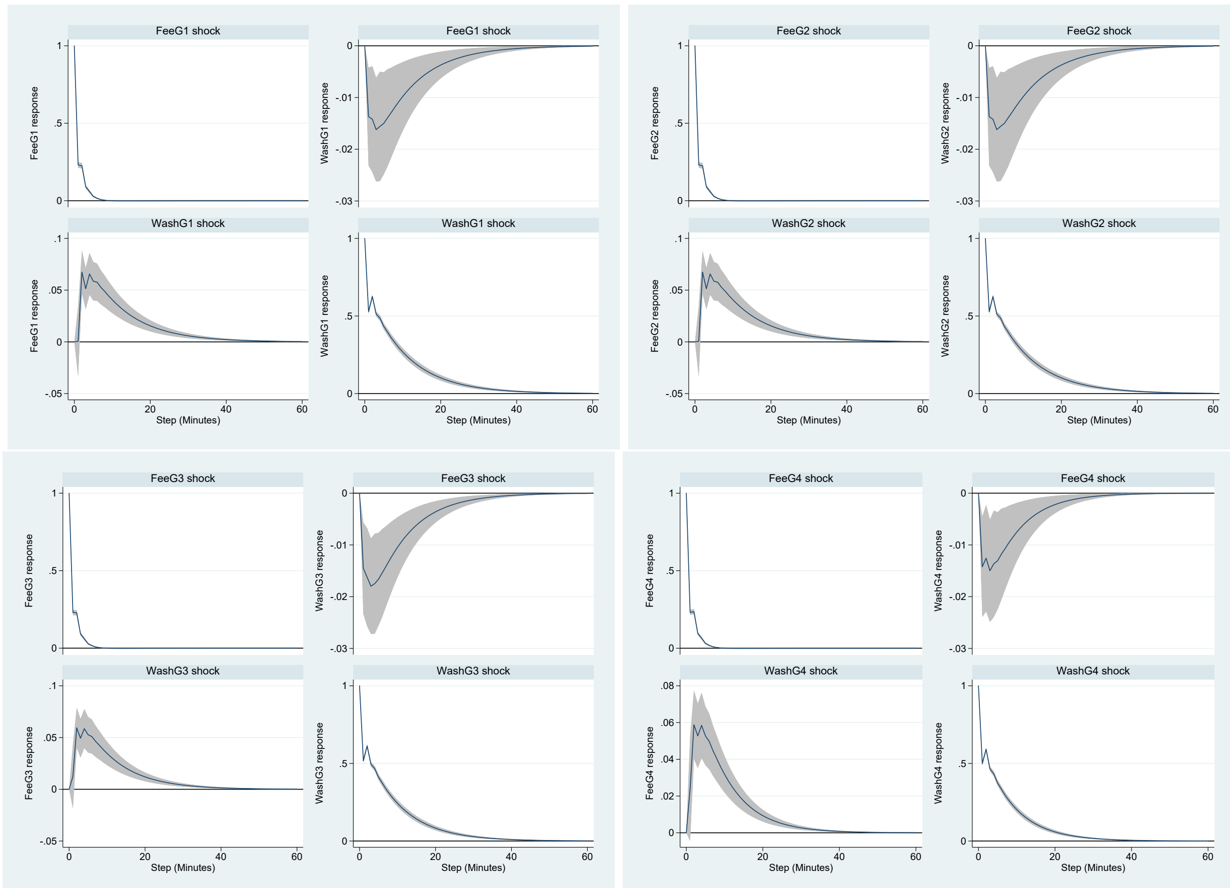
This figure plots the time series of 1) the log wash trading volume (in units of BTC) and 2) the log fee revenues collected from non-wash transactions (in units of BTC). Both time series are stationary: they contain no time trends, and significantly reject the unit root hypothesis by the Dickey-Fuller test.

Figure 10: Impulse response function



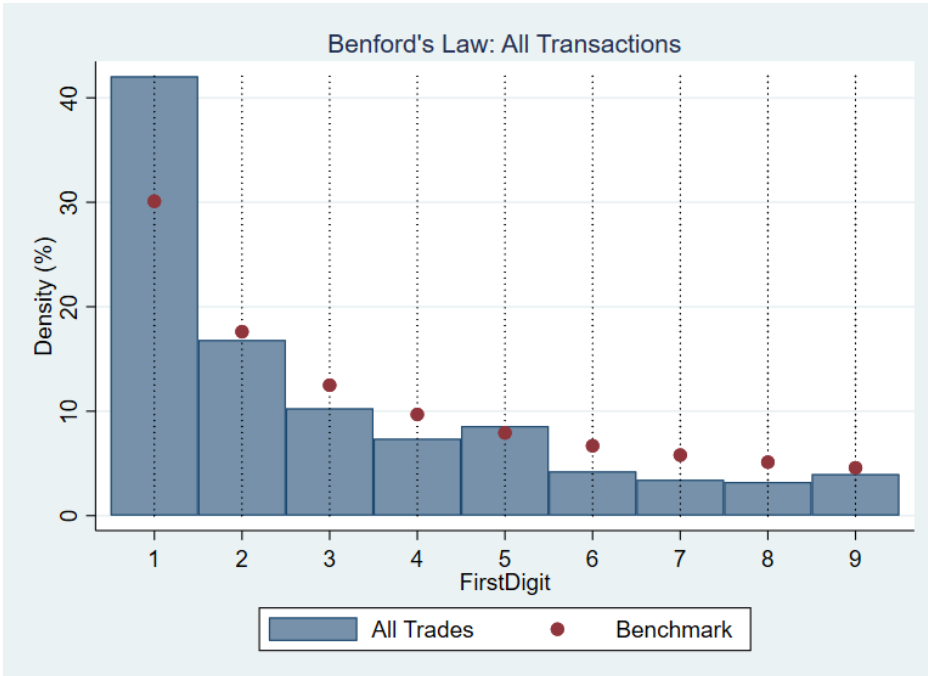
These figures plot impulse responses for the following VAR analysis: $Y_t = \sum_{i=1}^{p} A_i Y_{t-i} + \epsilon_t$, where $\forall i \in \{1, \cdots, p\}$, A_i is a 2×2 matrix to be estimated, Y_t is a 2×1 vector corresponding to 1) the log wash trading volume (in units of BTC) and 2) the log fee revenues collected from non-wash transactions (in units of BTC). The first row presents the evolution of fee revenue and wash trading volume in response to a one standard deviation change to fee revenue; and the second row the evolution of fee revenue and wash trading volume in response to a one standard deviation change to wash trading volume. A lower fee revenue tends to be followed by more wash trading activities. On the other hand, an increase in wash trading precedes a higher fee revenue. Both effects persist for the next 60 minutes. Shaded areas represent 95% confidence intervals.

Figure 11: Impulse Response Functions (Alternative Wash Trading Definitions)

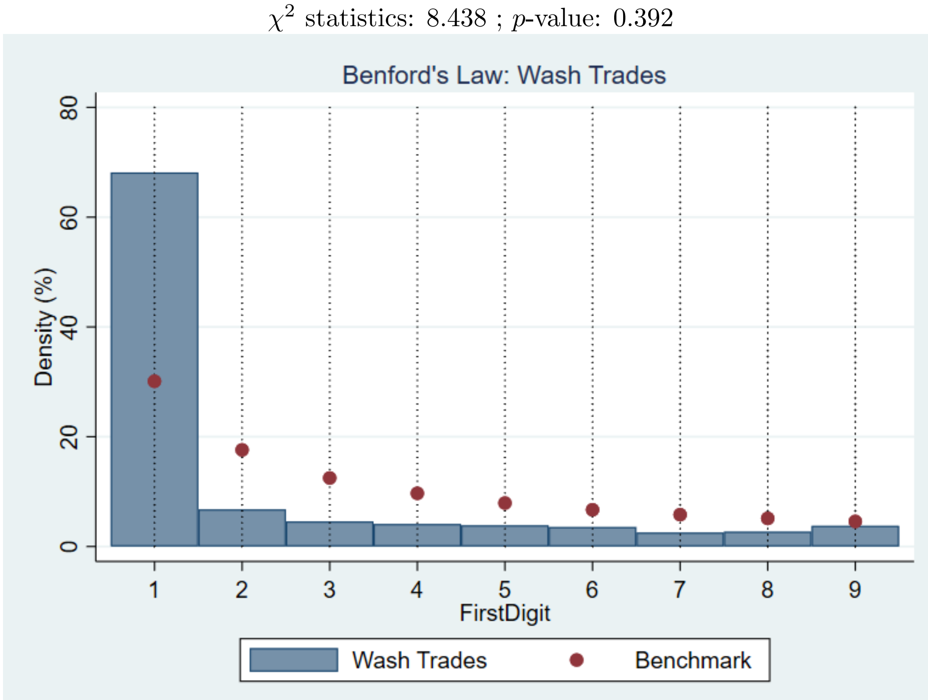


These figures repeat the impulse responses functions as in Figure 10: $Y_t = \sum_{i=1}^{p} A_i Y_{t-i} + \epsilon_t$, where $\forall i \in \{1, \cdots, p\}$, A_i is a 2×2 matrix to be estimated, Y_t is a 2×1 vector corresponding to 1) the log wash trading volume (in units of BTC) and 2) the log fee revenues collected from non-wash transactions (in units of BTC). With each panel, the first row presents the evolution of fee revenue and wash trading volume in response to a one standard deviation change to fee revenue; and the second row the evolution of fee revenue and wash trading volume in response to a one standard deviation change to wash trading volume. Wash trading are augmented with painting-the-tape transactions among insiders, defined differently as in Section 4. A robust finding across all panels is that a lower fee revenue tends to be followed by more wash trading activities and an increase in wash trading precedes a higher fee revenue. Both effects persist for the next 60 minutes. Shaded areas represent 95% confidence intervals.

Figure 12: Indirect Inferences by Benford's Law



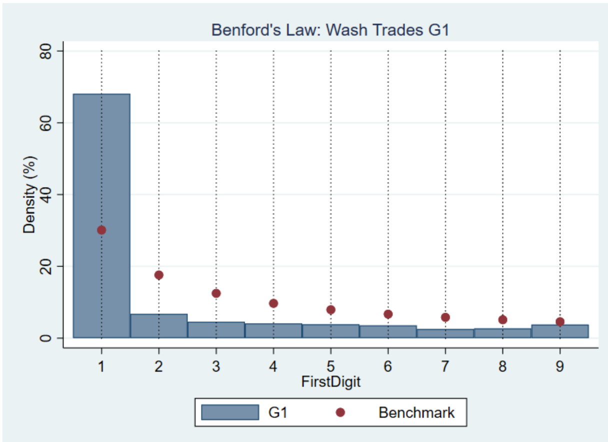
χ^2 statistics: 8.438 ; p-value: 0.392



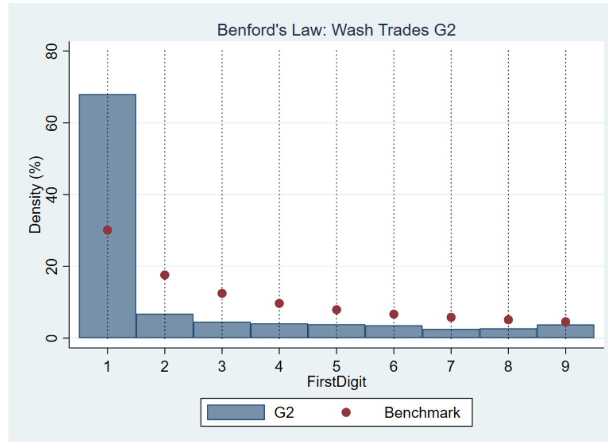
χ^2 statistics: 69.709; p-value: 0.000

This figure plots the distribution of the first significant digits of order sizes (bars), for all trades and only wash trades, respectively, and compare to the theoretical value suggested by the Benford's law. We find that the distribution of all trades does not significantly deviate from the Benford law, while it does for wash trades. Therefore, our direct evidence provides support for the effectiveness of the Benford's law for wash trading inference.

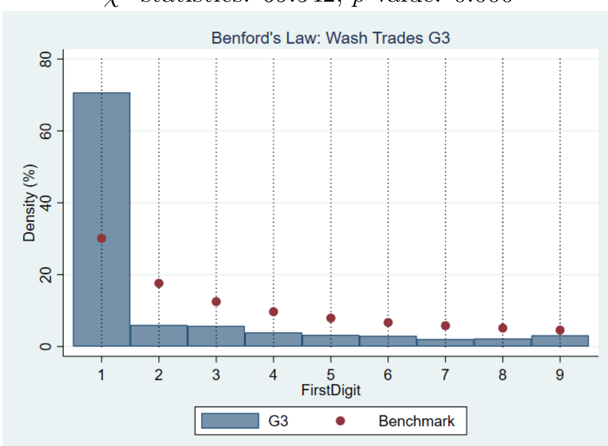
Figure 13: Testing Indirect Inferences by Benford's Law



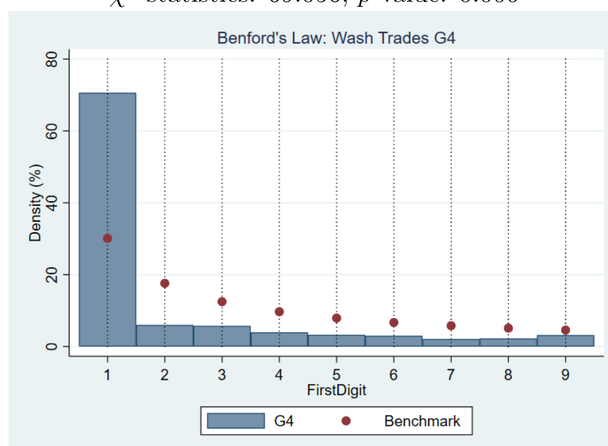
χ^2 statistics: 69.542; p-value: 0.000



χ^2 statistics: 69.056; p-value: 0.000



χ^2 statistics: 78.871; p-value: O.000



χ^2 statistics: 78.573; p-value: 0.000

This figure plots the distribution of the first significant digits of order sizes (bars) for the various groups of transactions (G1 to G4) between June 26$^{\text{th}}$, 2011 and May 20, 2013, and compare to the theoretical value suggested by the Benford's law. We find that the distribution of wash trades for all definitions significantly deviate from the Benford law. Again, our direct evidence provides support for the effectiveness of the Benford's law for wash trading inferences.

Tables

Table 1: Summary Statistics for All Transactions

Panel A: Full Sample: April 1, 2011 to November 30, 2013

	Mean (1)	Min (2)	25st percentile (3)	Median (4)	75st percentile (5)	Max (6)	Standard deviation (7)
Order size (BTC)	6.65	0.00	0.04	0.50	3.00	34274.76	44.57
Price (JPY)	10833.08	0.00	570.46	4071.52	12089.74	150000.00	19852.68
Price (USD trades only; USD)	103.16	0.50	6.51	25.93	119.63	1333.33	190.36
BTC fee (sell orders; %)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
BTC fee (buy orders; %)	0.41%	0.00%	0.29%	0.43%	0.55%	100.00%	1.32%
Money fee (sell orders; %)	0.38%	0.00%	0.27%	0.43%	0.55%	100.00%	0.42%
Money fee (buy orders; %)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Panel B: Sub-sample: June 26th 2011 ad May 20th 2013

	Mean (1)	Min (2)	25st percentile (3)	Median (4)	75st percentile (5)	Max (6)	Standard deviation (7)
Order size (BTC)	7.84	0.00	0.05	0.57	3.99	34274.76	50.34
Price (JPY)	4039.30	0.00	436.46	952.14	8080.29	26379.39	5183.94
Price (USD trades only; USD)	39.77	1.00	5.26	11.40	74.02	397.21	51.06
BTC fee (sell orders; %)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
BTC fee (buy orders; %)	0.38%	0.00%	0.28%	0.40%	0.55%	100.00%	0.57%
Money fee (sell orders; %)	0.36%	0.00%	0.27%	0.40%	0.53%	100.00%	0.38%
Money fee (buy orders; %)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

This table presents summary statistics for the entire sample, which includes all transaction records on Mt.Gox from April 1^{st} 2011 to November 30^{th}, 2013, as well as the period with active wash trading, which includes all transaction records on Mt.Gox from June 26^{th} 2011 to May 20^{th}, 2013. The BTC and fiat fees contain outliers because minimum fiat fee is US$0.00001, so those who buys 1 Sathoshi pay at least $0.00001 for the BTC and US$0.00001 fees (if applies).

Table 2: Sample Breakdown by Countries

Panel A: The Entire Sample (April 2011 to November 2013)

	Counts	Order size (BTC)	Order size (JPY)	Price (JPY)	FX Fee (%Fiat)	BTC Fee (%BTC)	Wash Trader TX
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
US	2941786	3.99	43867	16893	0.14	0.15	97491
		21.36	250774	22332	0.25	0.01	
DE	633944	3.74	36832	18308	0.23	0.00	1875
		18.61	184051	23830	0.31	0.01	
HK	333963	5.47	69862	20088	0.02	0.00	61
		25.97	315253	24376	0.14	0.00	
GB	556926	3.21	40536	19649	0.20	0.00	5916
		16.19	212913	23753	0.39	0.01	
!!	201502	5.04	64386	20124	0.23	0.00	3202
		27.46	296008	24395	0.29	0.02	
JP	560026	6.09	56042	14487	0.23	0.00	4868
		38.41	325843	19140	0.38	0.01	
PL	86827	6.50	70830	25504	0.21	0.00	835
		34.03	276717	29005	2.06	0.02	
Other countries	3327676	4.21	53029	20395	0.21	0.00	16564
		20.96	256440	25232	0.35	0.01	
All countries	8642650	4.25	49133	18630	0.18	0.00	130812
		22.76	256279	23769	0.36	0.01	
All	15802148	6.65	30909	10833	0.19	0.00	230550
		44.57	197303	19853	0.35	0.01	

This table reports descriptive statistics of bitcoin transactions for (verified) trader IDs based on their country of origin. Country codes are empty for unverified users, and `!!` for traders who fail verification (unverified users go through extra delays when withdrawing funds, but are otherwise not restricted in trading). The first and last column report the Counts of all transactions and wash trades for a particular country, and other variables report both the means and standard deviations. The last row, "All" includes transactions from "All countries" (the previous row) and those with empty country information (the case with unverified users).

Table 2: Sample Breakdown by Countries, Continued

Panel B: Period with Active Wash Trading (June 26th 2011 to May 20th 2013)

	Counts	Order size (BTC)	Order size (JPY)	Price (JPY)	FX Fee (%Fiat)	BTC Fee (%BTC)	Wash Trader TX
US	1599854	5.32	38619	8105	0.14	0.15	97491
		25.76	212319	5414	0.26	0.31	
DE	302852	5.45	31385	8031	0.22	0.20	1875
		24.03	156702	5461	0.30	0.44	
HK	160935	6.72	51997	8956	0.02	0.02	61
		29.92	214358	5474	0.09	0.09	
GB	255563	4.56	34291	8946	0.19	0.23	5916
		21.72	167312	5225	0.44	0.61	
!!	114521	5.98	47768	9548	0.23	0.27	3202
		31.76	233773	4868	0.28	1.48	
JP	336037	7.63	52537	8258	0.23	0.20	4868
		45.21	286780	5557	0.33	0.39	
PL	26828	7.16	51669	9812	0.15	0.26	835
		28.14	161789	5080	0.42	0.28	
Other countries	1584761	5.77	42214	9114	0.21	0.24	16564
		26.51	194153	5183	0.35	0.76	
All countries	4381351	5.70	41040	8604	0.18	0.19	130812
		28.02	207483	5346	0.32	0.59	
All	11097734	7.84	21030	4039	0.18	0.19	230550
		50.34	141144	5184	0.33	0.45	

Table 3: Sample Breakdown by Currencies

Panel A: The Entire Sample (April 2011 to November 2013)

	Counts	Order size (BTC)	Order size (JPY)	Price (JPY)	FX Fee (%Fiat)	BTC Fee (%BTC)	Wash IDs TX
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AUD	191937	4.26	30954	12764	0.14	0.29	528
		20.12	113439	18498	0.23	1.68	
CAD	33850	3.88	28816	20191	0.16	0.25	738
		24.19	103670	25133	0.23	1.74	
CHF	9344	3.59	28717	16478	0.07	0.29	1026
		14.54	99559	23259	0.18	0.29	
CNY	2319	3.55	19793	16041	0.11	0.53	90
		13.58	80059	26180	0.22	5.08	
DKK	2105	2.49	15661	11328	0.07	0.29	136
		7.96	42665	13216	0.17	0.29	
EUR	1272907	3.55	29849	16536	0.22	0.26	7949
		18.85	143630	24240	0.26	0.79	
GBP	423008	4.08	17413	8998	0.15	0.28	9379
		20.96	91878	17354	0.24	0.60	
HKD	2495	4.57	33434	15616	0.04	0.36	131
		17.53	104139	20467	0.13	2.84	
JPY	115656	4.86	56863	21526	0.13	0.27	1465
		21.99	223170	27323	1.81	1.92	
NOK	2455	2.92	29679	18875	0.03	0.32	4
		7.53	94571	27766	0.13	2.03	
NZD	3679	3.28	25028	11872	0.06	0.28	254
		10.59	91617	14147	0.17	0.28	
PLN	163351	2.42	11821	9529	0.24	0.29	3576
		10.83	49979	16980	0.28	1.85	
RUB	3619	8.15	9734	10146	0.13	0.29	163
		51.22	36501	17829	0.22	2.36	
SEK	10730	3.18	25506	15713	0.12	0.30	232
		12.37	89853	24631	1.79	0.29	
SGD	3441	4.14	34131	14323	0.04	0.30	369
		19.68	115835	23116	0.14	0.36	
THB	437	1.75	14570	17674	0.07	0.26	8
		4.26	45262	20914	0.17	0.27	
USD	13560815	7.14	31453	10213	0.18	0.19	204502
		47.45	206083	19290	0.25	0.93	
All	15802148	6.65	30900	10835	0.19	0.20	230550
		44.57	197267	19857	0.30	0.96	

This table reports descriptive statistics of Bitcoin transactions for Mt Gox trades based on their denominated fiat currencies. The first and last column report the Counts of all transactions and wash trades with a particular currency, and other variables report both the means and standard deviations.

Table 3: Sample Breakdown by Currencies, Continued

Panel B: Period with Active Wash Trading (June 26th 2011 to May 20th 2013)

	Counts	Order size (BTC)	Order size (JPY)	Price (JPY)	FX Fee (%Fiat)	BTC Fee (%BTC)	Wash IDs TX
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AUD	116463	5.45	25386	6154	0.11	0.28	528
		24.96	106958	5847	0.22	0.30	
CAD	15462	6.26	18408	6567	0.18	0.21	738
		34.90	69445	6014	0.24	0.28	
CHF	6244	4.52	20810	7095	0.08	0.29	1026
		17.05	81455	5705	0.19	0.28	
CNY	1341	5.08	15705	5844	0.16	0.56	90
		16.90	74509	5767	0.25	5.45	
DKK	1645	2.77	13046	8420	0.08	0.29	136
		8.83	40579	6814	0.19	0.29	
EUR	743875	4.58	19811	5906	0.21	0.26	7949
		22.49	99381	5641	0.25	0.41	
GBP	342796	4.57	11842	3965	0.14	0.28	9379
		22.84	63069	4979	0.23	0.34	
HKD	1751	5.72	24248	8124	0.05	0.35	131
		20.45	57851	6677	0.15	2.40	
JPY	58232	6.64	37110	7522	0.06	0.23	1465
		27.41	159847	6075	0.63	0.51	
NOK	1409	3.60	13731	7500	0.03	0.30	4
		8.64	25477	5908	0.13	0.29	
NZD	2467	3.74	20047	7345	0.08	0.27	254
		11.02	62403	5797	0.19	0.28	
PLN	130615	2.71	8400	4712	0.25	0.27	3576
		11.86	39307	5576	0.28	1.11	
RUB	2769	10.31	7597	4628	0.16	0.26	163
		58.34	26911	5755	0.23	1.91	
SEK	7550	3.88	15843	5836	0.14	0.31	232
		14.53	60679	5978	2.14	0.29	
SGD	2525	4.35	22108	6540	0.05	0.32	369
		21.79	89908	6148	0.15	0.38	
THB	293	2.21	10803	7453	0.09	0.26	8
		4.99	21249	5415	0.19	0.28	
USD	9662297	8.33	21485	3830	0.18	0.18	204502
		53.23	147130	5083	0.23	0.43	
All	11097734	7.84	21027	4039	0.18	0.19	230550
		50.34	141136	5184	0.25	0.45	

Table 4: Direct Evidence of Wash Trading

	TX counts	BTC volume
	(1)	(2)
Wash trading (amount)	230,550	1,213,397
Trades among wash traders (amount)	3,642,732	28,270,590
Wash trading (% out of total)	2.1%	1.4%
Trades among wash traders (% out of total)	32.8%	32.5%
All trades	11,097,734%	87,030,784

This table presents the number of transactions and bitcoin volume, as well as the percentage share out of all transactions, for wash trades and trades among all wash traders, respectively during the sub-sample from June 26$^{\text{th}}$, 2011 to May 20$^{\text{th}}$, 2013 during which there is active wash trading on Mt.Gox. Wash trades are defined as trades with the same trader ID in both sides (buy/sell), and wash traders are IDs that show up in both sides (buy and sell) of at least one wash trade.

Table 5: VAR Results

	G0		G1		G2		G3		G4	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	Wash Trade	Fee	Wash Trade	Fee	Wash Trade	Fee	Wash Trade	Fee	Wash Trade	Fee
Wash Trade L1	0.53***	0.00	0.53***	0.00	0.53***	0.00	0.52***	0.01	0.50***	0.02*
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
Wash Trade L2	0.35***	0.07***	0.35***	0.07***	0.35***	0.07***	0.34***	0.05***	0.34***	0.04***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
Fee L1	-0.01***	0.23***	-0.01***	0.23***	-0.01***	0.23***	-0.01***	0.23***	-0.01***	0.23***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Fee L2	-0.00	0.17***	-0.00	0.17***	-0.00	0.17***	-0.01	0.18***	-0.00	0.18***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Constant	-0.49***	-1.36***	-0.49***	-1.36***	-0.49***	-1.36***	-0.50***	-1.45***	-0.54***	-1.41***
	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.04)	(0.02)	(0.04)
					Granger Causality					
Wash Trade		***		***		***		***		***
Fee	***		***		***		***		***	
Both	***	***	***	***	***	***	***	***	***	***

This table presents the regression results from the following VAR: $Y_t = \sum_{i=1}^{p} A_i Y_{t-i} + \epsilon_t$, where $\forall i \in \{1, \cdots, p\}, A_i$ is a 2×2 matrix to be estimated, Y_t is a 2×1 vector corresponding to 1) the log wash trading volume (in units of BTC) and 2) the log fee revenues collected from non-wash transactions (in units of BTC). We include results for wash trading (G0, or the first column), while the next four columns contains robustness analysis for expanded definitions of wash trades. We report coefficients (and standard deviations), and show that wash trading Granger causes fees collected, and vice versa. Both wash trading and fee revenue Granger cause each other.

Table 6: IDs Suspected to Be Associated with Exchange Insiders

Insider Type	# IDs	#IDs with in-group TXs	#Wash-trader IDs	#In-group TXs	(%) All	BTC sum	(%) All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Known Insiders	3	2	2	422	0.00%	378,169	0.43%
Hijackers	318	2	10	16	0.00%	40,562	0.05%
Double-Users	71	54	7	133,144	1.20%	111,651	0.13%
Zero-Fee Traders	1,032	356	5	6,120	0.06%	107,241	0.12%

This table compares the variously-defined categories of likely exchange insiders. Each column (from left to right) reports the number of IDs within each category, the number of IDs within each category who ever trade with other insiders within the same category, the number of wash trader IDs within each category, the number of transactions within each category (as well as percentage within all transactions in the June 26th, 2011 to May 20th, 2013 subperiod), and the amount of bitcoins traded within each category (as well as percentage within all transactions in the June 26th, 2011 to May 20th, 2013 subperiod).

Table 7: Suspected Wash Trades Associated with Exchange Insiders

	TX counts	(%)	BTC sum	(%)
	(1)	(2)	(3)	(4)
G0	230,550	2.08%	1,213,397	1.39%
G1	230,766	2.08%	1,412,615	1.62%
G2	230,776	2.08%	1,412,807	1.62%
G3	294,474	2.65%	1,776,736	2.04%
G4	307,446	2.77%	1,915,865	2.20%

This table compares the variously-defined sets (G1-G4) of wash trades and painting-the-tape trades by exchange insiders. G1 is the set union of all wash trades and all transactions among wash trader IDs who are also Known Insiders. G2 is the union of G1 and all transactions among wash trader IDs who are also Hijackers; G3 is the union of G2 and all transactions among wash trader IDs who are also Double-Users; and G4 is the union of G3 and all transactions among wash trader IDs who are also Zero-Fee Traders. Each column (from left to right) reports the number of transactions in each set and the percentage share out of all transactions, as well as the amount of bitcoins traded within each set and the percentage share out of all transactions.