Towards Understanding Cryptocurrency Derivatives: A Case Study of BitMEX

Kyle Soska
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
ksoska@cmu.edu

Jin-Dong Dong
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
jd0@cmu.edu

Alex Khodaverdian
University of California, Berkeley
Berkeley, California, USA
alexkhodaverdian@berkeley.edu

Ariel Zetlin-Jones
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
azj@cmu.edu

Bryan Routledge
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
routledge@cmu.edu

Nicolas Christin
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
nicolasc@cmu.edu

ABSTRACT

Since 2018, the cryptocurrency trading landscape has evolved from a collection of spot markets (fiat for cryptocurrency) to a hybrid ecosystem featuring complex and popular derivatives products. In this paper we explore this new paradigm through a study of BitMEX, one of the first and most successful derivatives platforms for leveraged cryptocurrency trading. BitMEX trades on average over 3 billion dollars worth of volume per day, and allows users to go long or short Bitcoin with up to 100x leverage. We analyze the evolution of BitMEX products—both settled and *perpetual* offerings that have become the standard across other cryptocurrency derivatives platforms. We additionally utilize on-chain forensics, public liquidation events, and a site-wide chat room to describe the diverse ensemble of amateur and professional traders that forms this community. These traders range from wealthy agents running automated strategies, to individuals trading small, risky positions and focusing on very short time-frames. Finally, we discuss how derivative trading has impacted cryptocurrency asset prices, notably how it has led to dramatic price movements in the underlying spot markets.

CCS CONCEPTS

• **General and reference** → **Measurement**; • **Applied computing** → **Digital cash**; Electronic funds transfer.

KEYWORDS

Bitcoin, Cryptocurrency, Finance, Markets, Trading, Derivatives, BitMEX

ACM Reference Format:
Kyle Soska, Jin-Dong Dong, Alex Khodaverdian, Ariel Zetlin-Jones, Bryan Routledge, and Nicolas Christin. 2021. Towards Understanding Cryptocurrency Derivatives: A Case Study of BitMEX. In *Proceedings of the Web Conference 2021 (WWW '21), April 19–23, 2021, Ljubljana, Slovenia.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3442381.3450059

WWW '21, April 19–23, 2021, Ljubljana, Slovenia

ACM ISBN 978-1-4503-8312-7/21/04.
https://doi.org/10.1145/3442381.3450059

1 INTRODUCTION

Cryptocurrency trading has undergone a powerful shift since 2018 away from traditional spot[1] markets to a mixture of both spot and derivatives[2] markets.

Launched in November of 2014, BitMEX is a cryptocurrency exchange that trades exclusively in cryptocurrency derivatives and has been at the center of this paradigm shift.

After a slow start, BitMEX's popularity grew considerably in late 2017, with the retail frenzy surrounding Bitcoin, to over 600,000 trader accounts. Despite its success, BitMEX has become the target of criticism, earning the nickname *Arthur's Casino* after one of its co-founders, due the high amount of leverage and risky nature of trading it facilities. This criticism is not without merit; in 2016, Arthur Hayes, co-founder and CEO of BitMEX gave a talk [5] about the origins of BitMEX where he said:

> "There are people who offer similar types of products but are focusing on degenerate gamblers, aka retail traders in Bitcoin, so why don't we do the same? [...] we are going to create the world's highest leveraged Bitcoin/USD product and [...] enable anyone who has Bitcoin to trade financial derivatives. [...] You can trade Bitcoin with 100x leverage on the most volatile asset in the history of the world, it's a lot of fun."

Since then BitMEX has doubled down on its efforts to appeal to the entertainment side of trading by implementing public leaderboards that track the most successful traders on the platform, and commands in a site-wide chat room that allow users to share ground-truth facts about their trades with each other.

BitMEX's recipe for success has become the blueprint for many other exchanges such as Binance [7], Bitfinex [8], Bybit [12], Deribit [17], FTX [19], Huobi [22], Kraken [28], and OKEx [44] which together form a nearly 30-billion dollar [3] futures market as of February 2021. All of these exchanges have since implemented

[1] A *spot market* is a public financial market in which the assets are traded for immediate delivery. In the case of Bitcoin this is typically a market that exchanges Bitcoin for either traditional, "fiat" currency (USD, EUR, JPY, etc.) or a "stablecoin" (USDT [45], USDC [13], DAI [31], etc.), pegged to a fiat currency.

[2] In a derivatives market, rather than trading assets, participants exchange contractual agreements whose payoffs are determined by the price of the underlying asset.

their own[3] derivatives products based on BitMEX's most successful instrument, the perpetual future. Many of these exchanges have also implemented leaderboards and other elements of gamification BitMEX pioneered. Thus, even though BitMEX has recently been under US regulators' scrutiny due to its alleged failure at being compliant with U.S. securities laws [2], and, resultingly, may be facing serious headwinds, most of the changes BitMEX brought to the cryptocurrency trading ecosystem are likely here to stay.

In this paper, we explore the recent trend of derivatives trading in the cryptocurrency ecosystem through a deep dive into BitMEX. We use on-chain forensics, public liquidation events, and logs of the site-wide chatroom to provide a descriptive analysis of BitMEX and the users who trade there. We make the following contributions:

(1) A detailed description of the structure of BitMEX and the history of products it has traded.
(2) An evaluation of the size and impact of BitMEX using unforgeable on-chain data.
(3) A characterization of the traders on BitMEX, of the kinds of risks they take, and how they engage with the exchange.
(4) A discussion of the impact that highly leveraged derivatives have had on the cryptocurrency markets.

Additionally, we have built a public website[4] that keeps a live record of BitMEX and provides real-time access to our analysis platform. All the code used in this paper and most of the data[5] collected are open source and publicly available.

The remainder of the paper is structured as follows. In section 2, we provide background on cryptocurrency, BitMEX's structure and policies as well as the products that they trade. In section 3, we describe the on-chain and off-chain datasets that we have curated for analyzing the platform. We detail the mechanics of our on-chain analysis in section 4, and present the results of this analysis in section 5 before diving into the off-chain evaluation in section 6. Finally, we discuss our findings and limitations in section 7, related work in section 8 and closing remarks in section 9.

2 BACKGROUND

In this section, we provide background information for BitMEX, by first describing the properties of Bitcoin relevant to our exposition. Bitcoin is both the main asset traded on BitMEX and the currency used as collateral. We then discuss the instruments that BitMEX trades and specific details such as the on-boarding process and account management.

2.1 Bitcoin and Modern Cryptocurrencies

Bitcoin, first proposed in 2008 [39], is a decentralized peer-to-peer payment system. It functionally serves as a decentralized currency and store of value, and has spawned a number of alternative currencies that provide variations in terms of features and design choices.

Almost all modern cryptocurrencies share a few key traits that are important for this work. They each contain a supply of tokens that is both discrete and finite at any point in time (although algorithmically they may eventually grow unbounded), and use a public ledger of transactions ("blockchain") that anybody can inspect. Additionally, currency owners are able to transfer custody of the tokens amongst each other. These properties have lead to the emergence of markets whereby users exchange tokens for fiat currency, either with the assistance of centralized exchanges or through some peer-to-peer process.

Transactions involving cryptocurrency can either be on-chain or off-chain. An *on-chain* transaction is one that takes place natively in the cryptocurrency network and is logged into the public ledger, while an *off-chain* transaction is not directly recorded on the ledger. Due to the relatively costly process of embedding transactions on-chain, there have been various proposals (e.g., [46]) to use on-chain transations primary as a settlement layer (i.e., to record a number of transactions as a compound) rather than to record each individual transaction. For example an off-chain transaction occurs when an exchange matches orders between its customers and updates an internal database of each customer's holdings without settling this information to any public ledgers.

2.2 Customer Accounts

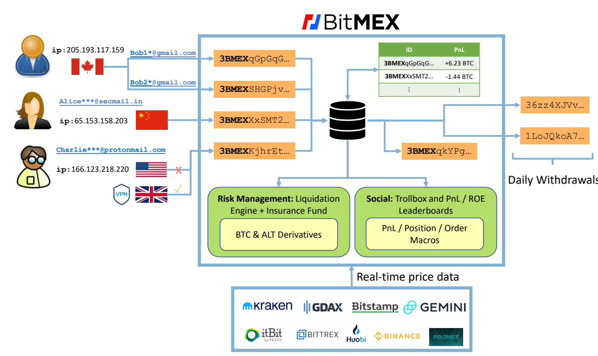


Figure 1: System Overview of Bitmex

Figure 1 shows an overview of the BitMEX exchange. Customers sign up for BitMEX by registering an account using an email address. This email could belong to a provider such as Gmail or to an email service wth stronger anonymity guarantees, such as ProtonMail [4].

Although BitMEX has not had any strict forms of Know-Your-Customer (KYC) policies—e.g., verification of government-issued identification documents—in the past, it did actively check the geo-location of customer IP addresses against a blacklist of prohibited locations. At the time of writing, this blacklist includes the United States, Quebec, Cuba, Crimea, Sevastopol, Iran, Syria, North Korea, and Sudan. If an account is ever accessed from an IP in the blacklist, the customer is given a grace period to close their open positions and withdraw their funds before the account is terminated. Discussions seen on Reddit and Twitter suggest that traders frequently use VPNs or other obfuscation techniques to circumvent this policy. User verification has since been enforced [9] and began to take effect on November 5, 2020.

[3] A few exchanges such as OKEx offered different derivatives products, namely their Bitcoin quarterly futures before BitMEX's sucess.
[4] http://cryptotrade.cylab.cmu.edu
[5] BitMEX's terms of service restrict the re-hosting and distribution of some data.

After registering an account with BitMEX, a unique vanity[6] address is created for deposits. Any funds sent to this on-chain address are credited to the corresponding account. At that point, a database internal to BitMEX maintains the state of user accounts including their profits and losses from trades. This process can be seen in Figure 1 where each user account corresponds to precisely one Bitcoin address. Due to the lack of restrictions however, a single user or entity may possess several BitMEX accounts. At any point users may withdraw their balance from BitMEX, however BitMEX only processes withdraws once per day, typically at 11am UTC.

BitMEX is structurally different from most exchanges because it operates exclusively with Bitcoin as opposed to fiat currencies or "stablecoins," that is, cryptocurrencies pegged to a fiat currency. On the left-hand side of Figure 1, users deposit funds via on-chain transactions in the Bitcoin network. On the right-hand side, all customer withdrawals are also processed in Bitcoin on-ledger. This is particularly important from a regulation and policy perspective, since at no point in a user's interaction with BitMEX will they ever convert their holdings into fiat.

2.3 Futures and Derivatives

Derivatives are financial products whose future cash-flows depend on an underlying asset's value (e.g., a stock, a commodity, a currency). These derivatives can be used to mitigate risk (hedging), or increase exposure to price movements (speculation). Traditional derivatives include forward contracts that commit to the underlying asset's future delivery at the agreed-to price. Other derivatives, options for example, specify that one party has the right but not the obligation to deliver the underlying. Derivatives contracts can be "physically settled" where the terms of the derivative are executed at maturity or "cash-settled", where the terminal value of the contract is calculated as the financial equivalent of delivery.

BitMEX offers a wide range of derivatives contracts whose value depends on the performance of an underlying cryptocurrency. By far, their most successful product is their *Perpetual Contract*[7], a product that shares similarities to cash-settled futures contracts. However, its details differ quite a bit from traditional financial futures markets. We focus on the *Perpetual Bitcoin Contract*, XBTUSD. This contract allows traders to enter levered positions that appreciate or depreciate with movements in an *index price* that represents the USD spot price of Bitcoin as measured on a variety of other cryptocurrency exchanges (see Figure 1).

For simplicity, let us first describe BitMEX's Perpetual Contract without paying attention to any maintenance fees. Consider a trader who enters a long position (i.e., they are betting the price of the underlying asset is going to increase). "Long" here is relative to a price that represents the USD price of one bitcoin. Assume that, at time t, the trader goes long on USD X worth of contracts. For instance, a trader could decide to invest USD 10,000 in these contracts, betting the Bitcoin price will rise. Given the XBTUSD price at t, P_t, the trader chooses an amount of leverage, L, and then she posts M_t bitcoins in her margin account on the exchange where

$$M_t = \left(\frac{X}{L}\right)\frac{1}{P_t}\ . \tag{1}$$

Each Perpetual Contract has a notional value of USD 1 worth of bitcoins. Hence, an entry position of USD X in Bitcoin is X/P_t bitcoins. With a leverage ratio of L, the trader must post M_t bitcoins on margin. BitMEX accepts the trade if this initial margin is at least 1% of the entry position, or $M_t \geq 0.01(X/P_t)$. Hence, the initial margin limits acceptable leverage to $L \leq 100$.

Returning to our example, if at time t, 1 bitcoin is worth USD 10,000, without any leverage, the entry position would be $X/P_t = 1$. With a leverage ratio $L = 100$, M_t would be 0.01 bitcoin. This implies the trader would only need to post USD 100 worth of bitcoin.

While a position is open, it is subject to *funding* and *minimum maintenance* requirements. Funding is paid or charged to positions every 8 hours by the exchange. Typically, the funding rate reflects a short-term interest rate. However, the funding rate includes a premium that reflects differences between the current trading price of the perpetual contract and the current index price. The funding rate explicitly links the performance of the derivative to the index.

Assume that time is divided in discrete periods, $t, t+1, \ldots$ and let r_{t+s}, with $s \geq 0$, denote the funding rate at (discrete) time $(t+s)$, with the convention that when $r_{t+s} > 0$, long positions pay short positions. At each period $(t+s)$ the position is open, the trader's margin account (BitMEX refers to this as the "wallet balance") updates according to

$$M_{t+s} = M_{t+s-1} - r_{t+s}\left(X\frac{1}{P_{t+s}}\right)\ . \tag{2}$$

If, as in our example, $s = 1$, $r_{t+1} = 0.001$ (0.1%) and $P_{t+1} = P_t = 10000$ (we assume here the price has not moved at all), we then have $M_{t+1} = 0.01 - 0.001 = 0.099$ bitcoins. In other words, unless the price increases, the margin is losing value.

Then the *equity value* of the position, V_{t+s}, margin plus unrealized gains or losses, fluctuates with the index price and funding:

$$V_{t+s} = M_{t+s} + X\left(\frac{1}{P_t} - \frac{1}{P_{t+s}}\right). \tag{3}$$

(In our running example, V_{t+1} and M_{t+1} are identical since the price did not move.) As long as the trader's position is open, it is also subject to a minimum maintenance margin requirement:

$$V_{t+s} \geq \theta\frac{X}{P_{t+s}} \tag{4}$$

where for XBTUSD, $\theta = 0.0035$. When a trader's equity value fall below the maintenance margin requirement, her position is liquidated by the exchange and she receives zero.

Also, notice that even if the funding were zero, $r_{t+s} = 0$, then (1)–(4) imply that the price at which the trader is liquidated satisfies

$$P^{\texttt{liquidation}} = \frac{\theta + L}{1 + L}P_t. \tag{5}$$

The exchange will liquidate the position of a long trader before her entire margin account is fully depleted ($P^{\texttt{liquidation}} > P_t L/(1+L)$) and, thus, even if a trader chooses not to use leverage, a long position will be liquidated before the price is zero.

In addition to the novel perpetual instruments, BitMEX has also offered several different instrument designs that trade exposure to

[6] A vanity Bitcoin address is one chosen to intentionally include specific characters, typically a prefix. In the case of BitMEX, addresses begin with the prefix `3BMEX`.
[7] BitMEX refers to this as a *Leveraged Perpetual Swap* while other services have used other terms such as *Perpetual Future*, *Inverse Perpetual Future* or simply *Perp*.

dozens of underlying assets. For simplicity we group the offerings from BitMEX into the the following broad derivatives categories: `Perpetual Bitcoin`, `Settled Bitcoin`, `Perpetual Ethereum`, `Settled Ethereum`, `Perpetual Altcoins`, `Settled Altcoins`.

3 DATA

Price Data: We collected simple price and volume data for all 265 instruments (and 212 indexes) traded on BitMEX since its inception. This information was provided by the exchange API and with only a few exceptions is comprised of 1-minute intervals containing the open, high, low and close price of the interval as well as the volume that was traded in terms of contracts. For each instrument and index we also recorded its full set of specifications that includes, among others, information such as the maker and taker fees,[8] contract sizes, listing and settlement times, initial and maintenance margin requirements and tick sizes.

We also collected price data for the Bitcoin markets on Coinbase [15], Kraken [28] and Bistamp [11] which have served as foundation for the Bitcoin index. Because some instruments have used foreign currencies as a basis, we also grabbed daily snapshots [43] of the ratio between USD and the Korean Won (KRW), Japanese Yen (JPY) and Chinese Yuan (CNY) to normalize trading volumes to USD. In total we collected over 97 million data points which represents 11.9 GB of data.

Trollbox and Liquidations: BitMEX implements a site-wide chat room with dedicated channels for English, Chinese, Korean, Russian, Spanish, French and Japanese where traders discuss the market in real time. This chat room is furnished with special commands that allow users to publicly and verifiably share information such as their profits and losses (PnL), orders and positions.

Prior to March 13, 2020, BitMEX exposed an API that allowed for the enumeration of the entire trollbox history. On March 13th BitMEX claimed to have been a victim of a computational DDOS attack [10] that exploited an inefficient API implementation and removed all trollbox history up to that point. Since our previous collection occured on March 2, 2020, we have an archive of the trollbox from its creation until March 2, 2020, and then from March 13, 2020 until the time of writing, with an 11-day gap in the middle.

The trollbox archive contains 57.8 million messages from over 149,000 unique accounts with over a million ground-truth data points about account positions and orders. Our copy of the trollbox including meta data spans over 48 GB.

Included in the trollbox are messages from a bot which was is run by BitMEX and goes by the username "`REKT`." This bot echos a live feed of liquidation events into the trollbox and includes information such as the product that the position was taken on, the size and direction of the position, and the price that the liquidation engine assumed control of the position at. We also have the approximate time that the liquidation occurred at based on the time-stamping of the message. We have collected over 425,000 liquidation events on 205 instruments totaling 60 billion dollars in value.

[8]In a standard trade, the *maker* is the party that places an offer to buy or sell an asset, security, or contract and the *taker* is the party that accepts this offer resulting in a trade. Typically, the fees for makers are smaller than the fees for takers.

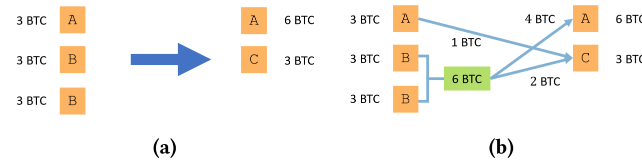


Figure 2: A hypothetical Bitcoin transaction (a) with three inputs from two addresses which generates two outputs (fees ignored) and the corresponding flow decomposition (b).

4 LEDGER METHODOLOGY

4.1 Detection and Filtering

Upon registration, every BitMEX account is assigned a unique corresponding Bitcoin address for receiving customer deposits. This unique address is owned by BitMEX and is generated with a `3BMEX` vanity prefix. To the best of our knowledge, account holders cannot change their deposit address.

Such vanity addresses are a necessary but not sufficient condition for identifying customer deposit addresses. Indeed, some unrelated, randomly-generated, Bitcoin addresses may end up with with a `3BMEX` prefix by chance; some may be intentionally crafted to imitate the exchange addresses.

To address this issue, we first filter out all `3BMEX` addresses active on the Bitcoin blockchain before BitMEX launched in November of 2014. Second, we discovered that the exchange frequently spends coins from multiple addresses to fulfill a withdrawal request, but never mixes inputs from `3BMEX` addresses with other, non-vanity addresses that it might own. Thus, we filtered out any `3BMEX` addresses that appear as a transaction input with non-`3BMEX` addresses.

In addition to customer deposit accounts that exhibit typical on-chain behaviors, some `3BMEX` addresses never directly receive funds from an external address. The functional role of these addresses is unclear: They could represent new customer accounts funded from existing accounts, or they could be internal BitMEX addresses that do not represent customer activity. We have also never observed any address besides `3BMEX` vanity addresses play a functional role in the exchange's on-chain presence. For example, we have never seen a customer withdrawal fulfilled from a non-vanity address; we have not seen any non-vanity address to seemingly serve as long-term storage. In short, we believe that vanity addresses—once filtered with our above heuristics—represent the totality of BitMEX's on-chain presence. In the remainder of this paper, we will denote the set of Bitcoin addresses identified as BitMEX addresses as *internal* addresses; and will call all other Bitcoin addresses *external* addresses.

4.2 Flows

We decompose Bitcoin transactions into input-output flows using Möser et al. [37]'s taint analysis intuition. For a transaction with a total input value of N bitcoins, an input address that contributes a fraction α of the input generates a flow of $\alpha\beta N$ coins to an address that receives a fraction β of the total output. We ignore "reflexive" flows where an address appears as both an input and an output of a transaction. In general, a transaction with a unique input addresses and b unique recipient addresses (not counting the implicit fee to miners) that share c elements decomposes into $ab - c$ flows. Figure 2 shows an example of flow decomposition.

Decomposing transactions into flows allows us to reason about where an address receives funds from and sends funds to, for instance, to compute which fraction of an account's deposits come from various known hot wallets [9] or from other BitMEX accounts. We also leverage flows to determine the role of a specific address.

4.3 BitMEX account clustering

As discussed in Section 2.2, absent KYC restrictions, a single entity may operate many BitMEX accounts. We thus want to detect and cluster instances in which a user owns several accounts to infer accurate customer demographics.

The structure of customer deposit addresses on BitMEX allow us to improve on traditional blockchain clustering heuristics [34]. The key insight is that, with a few exceptions, the entity sending funds to a deposit address on BitMEX is also the owner of the corresponding account. An exception to this rule arises when a third party deposits funds into a BitMEX account on behalf of the user. This occurs when a user makes a deposit through a mixing service, or when an exchange uses their hot-wallet to send funds on the customer's behalf. Another problem comes from *dusting attacks*, in which an external entity sends a very small amount of bitcoins to an address, hoping the recipient will spend it in a way that degrades their anonymity.

Service detection: We mitigate these exceptions as follows. Service addresses, such as exchange hot-wallets and dusters, are typically present in a large number of flows, to a diverse group of destinations. We thus consider the number of unique BitMEX accounts accessed, the number of bitcoins transacted and the distribution of transaction sizes to infer whether an address belongs to a service.

Iterative clustering: We then cluster BitMEX accounts together, by iterating over all flows from external addresses into BitMEX accounts and building up a constraint set as follows. We first apply the rule that two non-service external addresses with flows into the same deposit address are owned by the same entity. This captures the notion that only the BitMEX account owner would ever deposit money into their account, so that deposits from two distinct non-service addresses must actually belong to the same owner. We then apply a second constraint that two BitMEX accounts that receive deposits from the same non-services external address are owned by the same entity. The second constraint simply extends the idea of ownership from external Bitcoin addresses to the BitMEX accounts that are being funded. The result is a set of constraints on BitMEX accounts that induce a clustering.

Community detection: A few services remain undetected by our service detection heuristic, which causes the formation of a few very large and loosely connected clusters. To break down these clusters, we use community detection techniques, specifically, Label Propagation [48]. The algorithm works by first assigning every node in the graph a unique label before repeatedly updating each node's label to be the label that appears the most in its neighbors. The algorithm terminates when each node has the label that appears most frequently among the neighbors. Nodes with identical labels form a single community and are our final clusters. There are numerous community detection algorithms (see, e.g., [52]). However, most of them are computationally too expensive for clustering Bitcoin addresses. Label propagation is suitable, even with our large dataset (>4M nodes), due to its linear-time computation and our expectation of very dense connections within communities.

We only use deposit transactions from external addresses to internal addresses for clustering. Indeed, other transactions (internal-internal, or internal-external) cannot help, in general, without additional knowledge of how BitMEX internally moves funds.

5 LEDGER ANALYSIS

The Bitcoin ledger provides us with a view of the addresses that are owned by BitMEX including over 610,000 addresses[10] that are used to receive customer deposits which we utilize in this section to study the behavior of traders.

As discussed above, we cannot generally infer how many bitcoins are credited to each customer account at any point in time since that information is maintained using an internal database and is not synchronized with the Bitcoin ledger. The on-chain flows into customer deposit addresses do however provide us with ground-truth information regarding the funding of these accounts.

We spot checked the deposit and withdrawal history of a few BitMEX accounts which were provided to us by anonymous contributors, and found that BitMEX appears to prioritize using funds from the customer's deposit address to fulfill their withdrawals. Accounts that made unprofitable trades and whose balance fell below its on-chain value saw their deposit address used as a source of funds for fulfilling the withdrawals of other customers. This collection of observations suggests that withdrawal processing causes the on-chain representation of accounts to converge to the internal database. The velocity with which account balances may change and the relative infrequency of withdrawals means that the on-chain balance of an individual account is not particularly meaningful, but the collective distribution of on-chain account balances may still yield important insights about the distribution of wealth on the platform.

5.1 Account Activity

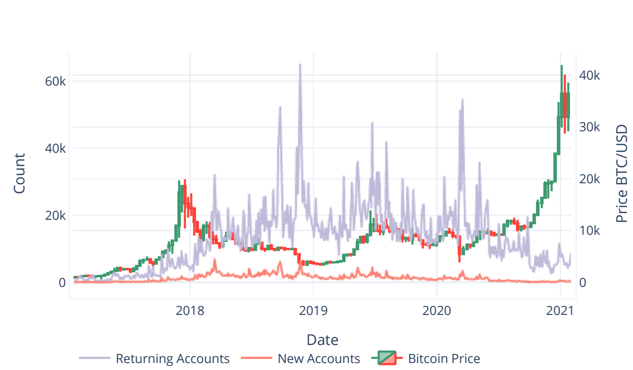


Figure 3: The number of new and existing BitMEX customer accounts that received deposits within a 432-block (approximately 3-day) period.

[9] A hot wallet is a Bitcoin address used by a service such as an exchange to process withdrawals on behalf of many customers.

[10] As of February 8th, 2020.

We usually cannot tell when a customer of BitMEX is actively trading, but we can still approximate activity by observing the on-chain deposits made to customer accounts. Figure 3 shows the number of customer deposit addresses that received funds on-chain in a rolling 432-block (approximately 3-day) window. When a customer's address receives funds for the very first time it is recorded as a new account otherwise it is recorded as an existing account. Unlike many services in cryptocurrency, BitMEX's popularity increased dramatically with the decline in Bitcoin's price in 2018, reaching a crescendo in November 2018 when the price tumbled to just over USD 3,000 per coin. One possible explanation for this trend is that the derivatives on BitMEX allow customers to gain short exposure to Bitcoin (i.e., make money when the price is going down), which, in the 2018 environment of steadily declining prices, was an attractive feature that very few other exchanges offered.

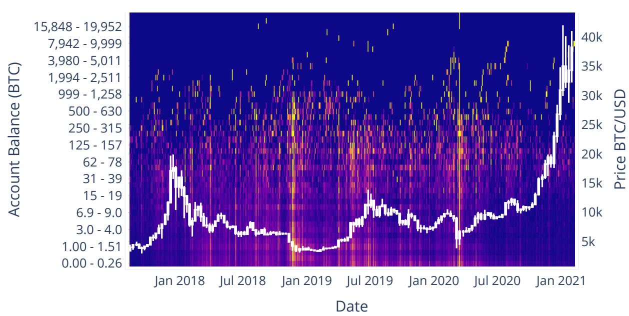


Figure 4: Heatmap of the inflows of Bitcoin to BitMEX by time, broken down by the wealth tier of the recipient account batched by 3-day blocks. Bright yellow indicates a flow of 400 bitcoins or more.

We explore how customers behave by considering the wealth of thier accounts denominated in bitcoin. To do this we define the tier of an account with an on-chain balance of b bitcoins as:

$$tier = \lfloor \log_{10}(b) * 10 \rfloor \tag{6}$$

We compute the tier of an account just before the funds are sent to it and then aggregate that inflow with other inflows from accounts of the same tier. We then partition the history of BitMEX into 4 hour blocks to produce the heatmap of Figure 4 that displays the volume of deposits that customers make to their accounts. To clearly observe the relationship between customer deposits and the price of bitcoin, we overlay the historical bitcoin price on the heatmap using the secondary y-axis.

The first insight Figure 4 provides, is that significant movements in price are followed by a corresponding increase in inflows from all tiers of accounts that tends to last roughly 2–3 days at a time. Large inflows to BitMEX also appear to correspond with a temporary reversal in the trend of Bitcoin's price, marking either a local high or low. Another detail to notice is that the aggregate inflows of wealthy accounts tend to look more random as opposed to the less wealthy accounts which appear structured which is a result of there being far fewer wealthy accounts on the platform.

Entity	Type	Coins	% Supply	Value (Bn. USD)
Coinbase [21] [11]	Exchange	944,039	5.18%	10.384
Grayscale [23] [12]	Fund	395,507	2.14%	4.351
Huobi [21]	Exchange	357,256	1.94%	3.930
Binance [21]	Exchange	273,838	1.48%	3.012
BitMEX	**Exchange**	**215,476**	**1.17%**	**2.370**
OKEx [21]	Exchange	210,428	1.14%	2.315
Kraken [21]	Exchange	135,143	0.73%	1.487
Bitstamp [21]	Exchange	125,329	0.68%	1.379
Bittrex [21]	Exchange	105,781	0.53%	1.164
Gemini [21]	Exchange	96,084	0.52%	1.057
HitBTC [21]	Exchange	71,754	0.39%	0.789
Bitfinex [21]	Exchange	66,942	0.36%	0.736

Table 1: The number of bitcoins held by several significant entities as of July 31, 2020.

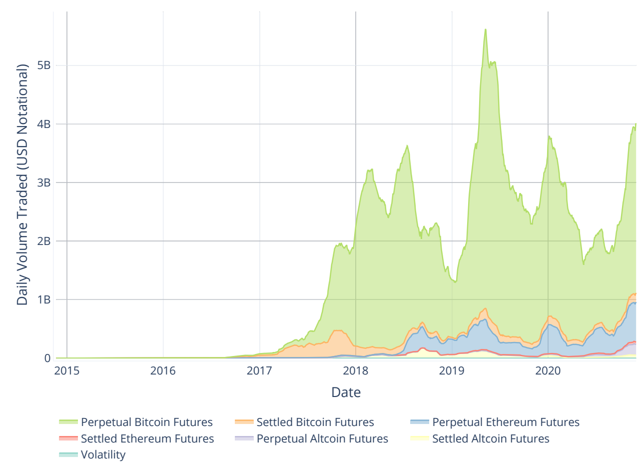


Figure 5: The absolute volume traded in various instrument categories on BitMEX over time smoothed using a 3-day simple moving average.

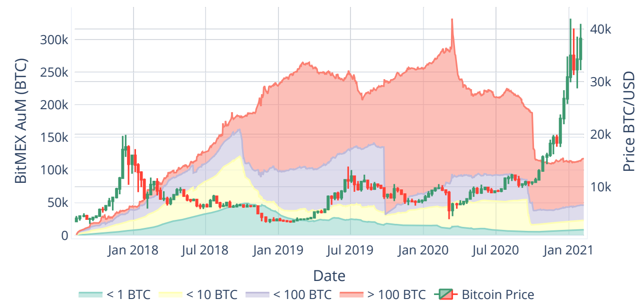


Figure 6: The total number of bitcoins held by BitMEX over time broken down by the tiers of addresses holding those coins.

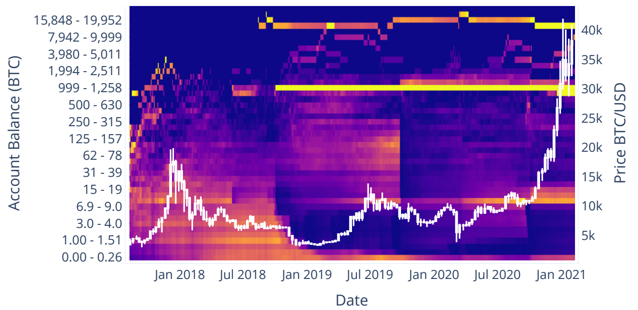


Figure 7: A heatmap of the distribution of wealth on BitMEX batched by 3-day blocks. Bright yellow indicates 10% ownership or more.

5.2 Size and Wealth Distribution

By aggregating the bitcoin held by all on-chain addresses from BitMEX, we can compute the total amount of coins in custody at any point in time. Table 1 compares BitMEX against some of the largest exchanges and known actors in the ecosystem as of July 31, 2020. BitMEX ranks fourth among exchanges and fifth overall with over 1.1% of the total supply of bitcoins which at the time was valued at over 2.3 billion US dollars.

Figure 6 shows the number of bitcoins that have been custodied by BitMEX over time, decomposed by the value of the customer addresses that are holding them. BitMEX thrived following the collapse in the price of Bitcoin in 2018, growing its assets until the summer of 2019 where it briefly dipped before peaking around 310,000 bitcoins on March 13, 2020. In September 2020, the United States Department of Justice indictment of BitMEX [2] lead to a material decline in bitcoins held by the exchange. These trends mirror what we observe with respect to trading volumes as shown in Figure 5 where the traded volume of products on BitMEX really exploded in popularity through 2018 and into 2019. A number of efforts to identify wash trading of popular cryptocurrency exchanges [29, 40, 49] have failed to find any on BitMEX and consistently rank it among the most transparent exchanges.

Figure 7 is a heatmap of how wealth is distributed on BitMEX accounts over time. In 2017 at the height of the retail mania, most of the wealth on BitMEX was concentrated into accounts that held 10 bitcoins or less. As we discuss in section 6.2, November 2018 culminated in a massive liquidation event of long contracts that simultaneously shifted the wealth demographics towards higher-tier accounts holding the majority of coins while many lower tier accounts were wiped out. This pattern appears to have occured again in September 2019; however, further inspection indicates that in this event, BitMEX seemingly confiscated tens of thousands of bitcoins and placed them into special vanity accounts that had never received any external deposits before. One possibility is that these accounts constitute the insurance fund that the exchange maintains, and the movement simply consolidated coins that had been earmarked for the insurance fund. Curiously this shift of funds occurred within moments of a sharp decline in the price of Bitcoin of over 20%, and further research is needed to determine if this played a causal role in the price movement or if it was merely a coincidence.

[11]This includes coins in custody after their acquisition of Xapo [14] in 2019.

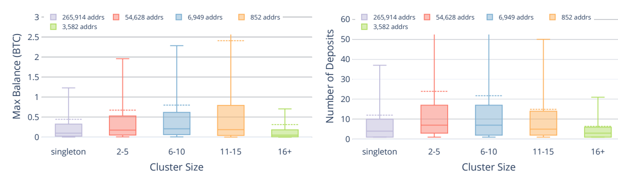
[12]Also includes the Grayscale Large Cap Fund [24].

5.3 Trader Sophistication

The on-chain activity of accounts suggests that some actors are engaging with BitMEX in sophisticated ways. We first derived clusters of accounts using the methodology described in Section 4.3.

Someone may choose to interact with BitMEX through multiple accounts for a few reasons. First, BitMEX's risk management restricts the leverage of large positions (> 200 bitcoins) which can be circumvented by splitting a position across multiple smaller accounts. Second, sophisticated traders may use the rate-limited API service provided by BitMEX to perform automated trading. A trader can multiplex their commands through multiple accounts to increase their effective rate limits or run separate algorithms and strategies on different accounts altogether. Third, trading bitcoin is unique because the flows that traders create between exchanges are public, and so sophisticated traders may wish to obfuscate these movements to mitigate the impact of being front-run.

After applying both our rule-based and community detection algorithms for clustering, we identified that about 90% of accounts are not part of a cluster while less than 1% belong to clusters of 5 or more accounts. We did however discover hundreds of prolific clusters, the largest of which include 50 accounts or more.



(a) Maximum balances **(b) External deposits**

Figure 8: The 5th, 25th, 50th, 75th and 95th percentile of the average maximum account balance (resp. average number of deposits) for accounts belonging to clusters of various sizes. The dashed line is the mean value for clusters of each size.

Figure 8 characterizes the clusters by account balance, and average number of deposits. In Figure 8(a), we see that cluster sizes of 6 to 10 accounts on average appear to have a higher amount of wealth per account than clusters of other sizes, albeit not by a large margin. The singleton clusters, on the other hand, are significantly lower than all the other sizes of clusters. This plateau suggests that the account wealth may be intentionally limited as discussed before and large deposits are scaled horizontally forming larger clusters. Figure 8(b) suggests that larger clusters also engage in a higher number of deposits than the singleton clusters. Large numbers of on-chain transactions may be a sign that the account is being used as part of an arbitrage strategy where the trader manages accounts on multiple exchanges that are frequently reconciled using on-chain transactions. These observations suggest that, in addition to retail speculators, BitMEX is utilized by highly sophisticated traders which echoes the claims made by a professional market marker [18] about the usefulness of derivatives in cryptocurrency.

An example of a large cluster would be the one rooted from Bitcoin address 1KiJkugknjgW6AHXNgVQgNuo3b5DqsVFmk, which owns 86 BitMEX accounts. This address has sent approximately 13,900 bitcoins to BitMEX but has extracted over 72,100 bitcoins from it.

6 USER EVALUATION

We complement our on-chain evaluation of BitMEX with an analysis of its users. We first look at the site-wide IRC-like chatroom known as *the trollbox*, before analyzing leveraged positions.

6.1 Trollbox Analysis

Trollbox users consist of a mix of traders, administrators and automated bots that post information such as a live feed of position liquidations on the platform. The trollbox also supports macros such as /position, /orders, /pnl, /rpnl, which display unforgeable facts about the account of the user issuing them.

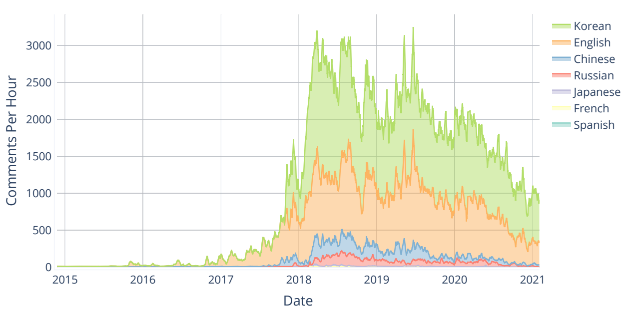


Figure 9: 7 day moving average of the number of comments left in the trollbox broken down by language.

6.1.1 General statistics. The BitMEX trollbox is a highly prolific messaging system, with 57.8 million messages from November 14th, 2014 to February 10th, 2021. Figure 9 shows a 7-day moving average of the trollbox message volume, broken down by language. Since 2018, the trollbox has sustained an average of over 2,000 messages per hour with frequent spikes above 3,000 messages per hour. The popularity of the trollbox closely mirrors the total trading volume on BitMEX shown in Figure 5. This far surpasses other mediums of cryptocurrency discussion such as the popular cryptocurrency subreddits /r/cryptocurrency, /r/bitcoin, /r/bitcoinmarkets, /r/ethfinance and /r/ethereum which average just above 200 comments per hour.

Likely owing to South Korea's cryptocurrency frenzy [38], Korean became in mid-2018 the most popular language, followed by English; Chinese is a distant third.

In Figure 10, we organize the messages by time of day into one-hour buckets and normalize each bucket by volume. A surprisingly small amount of temporal correlation occurs among languages. Russian and Chinese exhibit patterns where the most common hour of the day is more than twice as prolific as the least common hour of the day. This is understandable since a significant concentration of people who speak these languages live in a few consecutive timezones. English messages on the other hand are

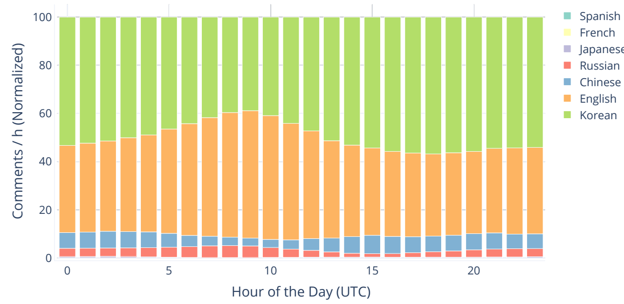


Figure 10: The number of comments left across the different language channels in the trollbox by time of day, normalized by volume.

relatively time invariant and likely reflects the global distribution of English-speaking traders.

The relative invariance of the Korean message volume to the time of day is far more surprising. While 94% of Korean speakers live in the GMT+9 timezone [1], the most prolific hours of the day for the Korean language only contain around 50% more messages than the least popular hour. This is in dramatic contrast to the trends observed in hobbies such as videogames [51] where the ratio between peak and troughs is regularly 3 or higher. Unlike traditional financial markets, cryptocurrency markets are active 24/7. Korean traders seem to be active at all hours of the day, indicating that trading may be an all-consuming activity for many of them.

6.1.2 Sentiment. To further our understanding of BitMEX traders, we next describe a sentiment analysis of the trollbox messages. The influence of Bitcoin price fluctuations on user mood should indeed reveal the timeframes on which traders operate.

Trollbox messages are similar to sentences, and average around eight words per message. Messages however contain a lot of slang, profanity, emojis, ASCII-art and community specific terms such as asset tickers,[13] which makes pre-trained sentiment models poorly suited; likewise, the absence of any ground-truth label makes training a new model difficult. However, a key insight is that the average mood of the trollbox is still likely correlated with the price action of Bitcoin, and that correlation allows us to extract some signal.

Thus, we first automatically assign labels to trollbox messages based fluctuations in the price of Bitcoin. Parameterizing the label assignment algorithm allows us to adjust the time-frame considered for the label. We then take this labeled data and use it to train a convolutional neural network following the approach of Kim [26] using the CoreNLP [32] open sourced natural language processing package and its Python variant Stanza [47]. A labeling of messages drawn from time-frames synchronized with the mood of BitMEX traders should produce a higher performance model than one produced by labels drawn from orthogonal time-frames.

Intuitively, we want to assign labels to messages to capture trader excitement when the price is going up rapidly, and despair

[13] A ticker symbol is an arrangement of characters, typically letters, which represents a particular asset or market which is traded publicly.

or capitulation when it is going down. *Technical indicators*[14] allow us to mathematically describe price fluctuations.

In particular, the Relative Strength Index (RSI) [50] takes as input the price history p of an asset and a time parameter σ and outputs a value in the range $[0, 100]$ to describe the momentum of that asset's price at time t, based on fluctuations over (roughly) the previous 15σ. (We refer the reader to Wilder [50] for a formal definition.)

We score the sentiment of each message by first computing the RSI value at the time the message appears in the trollbox. We partition the space of RSI values into five ranges, $[0, 30]$, $(30, 43]$, $(43, 57]$, $(57, 70]$, $(70, 100]$, which we map to the sentiment labels $[0, 4]$. We chose this specific partitioning as it nearly balanced the number of messages assigned each label when using $\sigma = 1$ hour. The higher the score, the more positive the sentiment. We then tag the message with the corresponding sentiment label. For instance, a message issued when $RSI = 37$ is tagged with sentiment value 1.

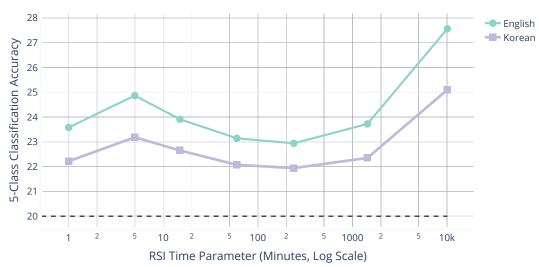


Figure 11: CNNs accuracy for sentiment prediction when trained using trollbox data with labels derived from RSI, using different time parameters.

Besides labeling, we removed all automated messages, macros and known bots. We sanitized the remaining messages by removing special characters, URLs, and usernames. We also sanitized numbers to avoid situations where the price of Bitcoin might be used to influence the model; however we did leave in punctuation as that may be influential in the sentiment of messages. We trained a separate model for each language and used language-specific precomputed word vector mappings [35][15] for each. For each model, we balanced the training and testing data across classes by re-sampling the minority classes to match the majority class.

Figure 11 shows the performance of the classifiers when trained with labels derived from the RSI using a time-span parameter ranging from one minute to one week (10,080 minutes). A randomly labeled dataset expectedly produces a classifier with just 20% accuracy in the 5-class prediction task, so all RSI labellings produced a signal that encodes some information about sentiment.

The local maxima at $\sigma = 5$ minutes implies that conversation in the trollbox is largely focused on price action from the last hour or so ($15\sigma = 75$ minutes). Manual inspection confirms that users who have recently made profitable trades are disproportionately prolific in the trollbox relative to those who have not. The quality of the trained models falls off until $\sigma = 1,440$ minutes or 1 day and really takes off at $\sigma = 10,080$ minutes or 1 week. This suggests that the sentiment of the trollbox is also largely impacted by the trend of the market over the previous 15 weeks. We suspect that this is due to survivor bias where traders whose (bearish or bullish) outlook on the market has been supported by the price trend are prolific, while many traders whose outlook has been contradicted by the market trend have dropped out of the platform.

[14] A technical indicator is a heuristic or pattern-based signal that is produced by the price, volume, and/or open interest of a security or contract and used by traders who follow technical analysis.

[15] These mappings are publicly available: https://code.google.com/archive/p/word2vec/.

6.2 Leverage Analysis

One of the best records on the leverage used by traders at BitMEX comes from a blog post [25] by BitMEX CEO Arthur Hayes where he details ground-truth information about the leverage traders applied to their positions between May 2018 and April 2019. His analysis took snapshots on the last day of each month and calculated the effective leverage of each position on the XBTUSD perpetual Bitcoin futures instrument. Over this time period, Hayes shows that the weighted average leverage of long and short positions was around 20–35x with short positions briefly averaging under 20x leverage in November and December 2018. Additionally, the average effective leverage of long positions is on average higher than that of short positions, but there is significant volatility, and short positions were more leveraged during three of the twelve months analyzed.

We supplement Hayes' analysis by covering all activity on BitMEX in continuous time up to August of 2020 and extending our investigation to all traded instruments. Unfortunately, we generally cannot know the leverage of a trader's position, so we cannot directly replicate Hayes' experiment. Instead we explore trader leverage and risk by looking at *liquidation* events.

Without user verification, BitMEX was unable to know the identities of traders on its platform. As a result, when a trader's account becomes overdrawn, BitMEX had no recourse to seek additional funds from the user though a traditional margin call process. Instead, BitMEX took over the risky position in a process called *liquidation*.

Liquidation events are broadcast publicly through both an API feed and via an automated "REKT" bot in the trollbox. These public events include the instrument that the position was taken on, the size of the position, and the liquidation timestamp.

Liquidations over time. Figure 12 shows a 7-day moving average of the daily volume of contracts liquidated on BitMEX, adjusted to US dollars, and compares it to the Bitcoin price. As expected, the amount of daily liquidated contracts picked up with the trading volume in 2018 following the market top and spiked with increases in price volatility, peaking as high as 1 billion dollars in aggregate in a single week in November 2018. Most of 2018 was characterized by significant liquidation events (> USD 100M) every few weeks which coincides with the price fluctuating in rapid discrete jumps, a pattern referred to by the community as *barts*. Although barts share a strong correlation to these liquidation events, further research is needed to determine if leveraged Bitcoin trading plays a causal role in barts or if these liquidations are merely a symptom of the price action. Also note that after contacting BitMEX about our research in November of 2020, the REKT bot was disabled in the trollbox until eventual being re-enabled in January 2021.

As Figure 12 shows, significant liquidations tend to disproportionately occur to the long side of contracts with aggregate long liquidations regularly spiking above short liquidations. Curiously,

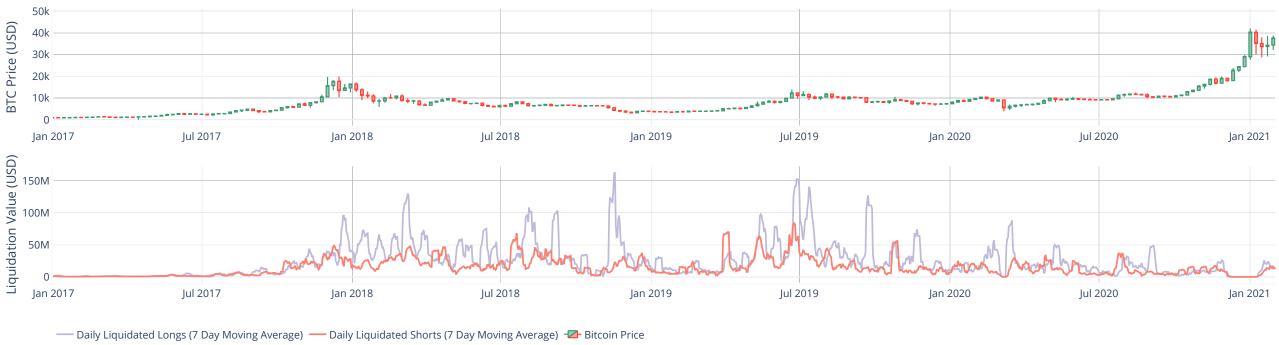


Figure 12: 7-Day moving average of total daily liquidations on BitMEX from January 2017 to February 2021.

this observation holds even when the price of Bitcoin is trending up as seen in July 2019. Two notable exceptions occured. On April 1, 2019 a large coordinated purchase of Bitcoin took place on the BTCUSD spot markets at Coinbase, Kraken and Bitstamp, and resulted in a 25% increase in the spot (and therefore index) price of Bitcoin, causing the liquidation of over USD 400M of short contracts. On October 24–25 2019, PRC president Xi Jinping declared that China aspires to become a world leader in blockchain technology, which triggered a short-lived bull run.

Liquidations over instruments. Table 2 aggregates liquidations by product category to illuminate any trends specific to a particular instrument class. As Figure 12 suggested, there is an asymmetry among the volume in liquidated long and short contracts. This trend is present regardless of the instrument and underlying asset that is being traded, however the ratio between long and short liquidation volume is somewhat unstable.

The fraction of liquidations over total trading volume is a proxy for evaluating the risk of an instrument: higher volume-normalized liquidation denote instruments with riskier positions. In all cases, settled futures appear riskier than perpetual swaps. Additionally, Bitcoin and altcoin futures have very similar volume-normalized characteristics while Ethereum settled futures appear riskier for both longs and shorts, and Ethereum perpetual futures appear to be safer. These results are interesting since the instruments in these categories support different amounts of leverage, with Bitcoin allowing up to 100x leverage, Ethereum allowing up to 50x, and altcoins being a mix that typically ranges from 20x to 33.33x.

Liquidations over position sizes. Figure 13 hints at the how much liquidation volume is contributed by positions of different sizes. This plot was formed by partitioning position sizes into USD 50,000-buckets and plotting the cumulative value of all liquidated positions up to a particular value. 50% of long liquidations come from positions of USD 1.6M and under, while 50% of short liquidations come from positions USD 950K and under. The monotonely decreasing slope of both curves implies that a disproportionate fraction of total liquidations comes from smaller position sizes. This could potentially be due to better risk management and lower personal risk tolerance of traders who manage larger positions or a systematic fallacy of traders who are reluctant to sell their losers [41]. Additionally, the difference between the cumulative long and short liquidations forms a (black dashed) curve with positive slope at all

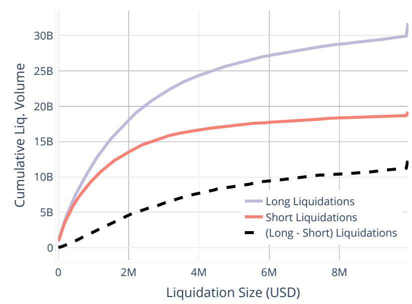
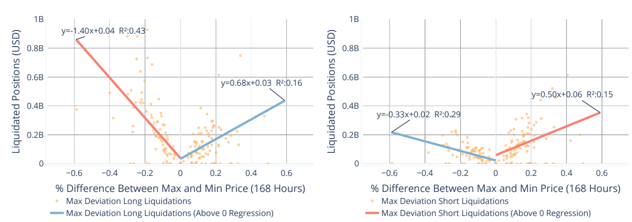


Figure 13: Cumulative XBTUSD perpetual future liquidations by increasing position size.

points. This shows there is always a greater liquidation volume of long positions regardless of size.



(a) Long liquidations **(b) Short liquidations**

Figure 14: Linear regression of the volume of liquidated contracts on BitMEX vs the difference between the maximum and minimum Bitcoin price traded within a 1-week (168 hour) period.

Liquidations over price fluctuations. One key to understanding why liquidations occur is to study the price action of Bitcoin leading up to a liquidation event. In Figure 14 we partitioned the history of BitMEX into 1 week (168 hour) sections and computed the maximum and minimum price that was traded on the XBTUSD perpetual Bitcoin instrument [16] within each section along with

[16] We used the Bitcoin spot price on Coinbase to analyze data before the XBTUSD instrument existed.

Contract Type	Long Liquidations (USD, in billions)	Short Liquidations (USD, in billions)	Total Liquidations (USD, in billions)	Long/Short Liquidation Ratio	Volume-Normalized Long Liquidations	Volume-Normalized Short Liquidations
Perpetual Bitcoin	30.48	18.30	48.78	1.67	1.26%	0.76%
Settled Bitcoin	2.52	1.04	3.56	2.42	2.06%	0.85%
Perpetual Ethereum	1.29	0.90	2.19	1.43	0.61%	0.43%
Settled Ethereum	0.29	0.10	0.39	2.86	2.64%	0.92%
Perpetual Altcoins	0.06	0.02	0.08	3.24	1.23%	0.38%
Settled Altcoins	1.31	0.55	1.87	2.38	2.13%	0.98%

Table 2: The USD value of liquidated contracts aggregated by instrument types on BitMEX in addition to the total liquidation volume normalized by total traded volume on the respective instruments up to September of 2020.

the total USD value of all long liquidations across all instruments. If the price at the beginning of the interval is lower than at the end, the difference is assigned a positive value, otherwise the difference is defined to be negative. In Figure 14(a), when we fit a linear regression to the distribution, the slope of liquidations of long contracts is unsurprisingly steeper when the price is trending down. What is less intuitive is that the volume of long liquidations is positively correlated with increases in price, that is, as the price of Bitcoin trends up and the gap between the minimum and maximum price traded within a week expands, the volume of liquidated long contracts increases. This could potentially be explained by an increase in volatility during weeks with significant price expansion. Late June–early July 2019, as seen in Figure 12, is a good example: while the price trended up, significant volumes of long liquidations were observed. Although we restrict here our analysis of liquidated position sizes to the XBTUSD instrument, we noted nearly identical trends on the Ethereum and altcoin instruments.

By contrast, Figure 14(b) displays the same linear regression analysis for short contracts. Again as expected, the slope of the regression is less steep when the price is going down and steeper when the price is increasing and going against the position. Interestingly, the slope of the regressions for short contracts are significantly less steep then those for long liquidations, and as the price is increasing, we expect to observe higher volumes of long liquidations than short liquidations.

7 DISCUSSION

BitMEX—and for that matter, related cryptocurrency derivative markets—raises a number of important questions regarding whether the service it offers is a societally desirable, or even a net positive for cryptocurrency adoption. The demand for leveraged exposure to cryptocurrency from retail speculators and professional traders alike is clearly present, based on the level of activity we observed on the platform. However, community anecdotes [42], coupled to our own leverage and liquidation analysis suggests that products like those traded on BitMEX exacerbate large moves in underlying asset price. History has taught us that commodity speculation [33] using derivatives can have undesirable consequences: cryptocurrencies are simply the newest manifestation of this issue.

More specifically, the complexity of the derivative instruments offered, paired with the tremendous amount of liquidations we observe, particularly of modest size, suggests that not all small, "retail" traders fully understand the high risks involved. Similar concerns in the past have motivated policies to restrict certain financial offerings to accredited investors.

Limitations and Future Work. We did not study the impact of geo-fencing. This could be done by checking the on-chain flows for systematic differences before and after BitMEX implemented this policy (roughly in Nov./Dec.-2018). Several other exchanges such as Bybit [12] currently rely on geo-fencing, so understanding its efficacy could have profound consequences.

There may also be significant structure in the on-chain transactions that BitMEX generates for fulfilling withdrawals that could further enhance our understanding of trading behavior. Another potentially valuable signal we did not use lies in the millions of ground-truth position, order, and profits-and-losses datapoints that traders and bots posted in the trollbox along with the public leaderboard of the most profitable accounts.

BitMEX Statement. We reached out to BitMEX in November 2020 with a draft of the paper and the analysis website. BitMEX representatives responded with the following statement, without elaborating any further:

> "We will not provide specific comments on your paper as it contains various inaccurate and/or misleading statements that do not properly reflect the platform's structure and operations and also do not reflect the platform's user verification requirements that are in place for all customers."

8 RELATED WORK

While there is ample financial literature on the study of derivatives trading, cryptocurrency derivatives trading is novel enough to have remained mostly unexplored—save for Hayes' aforementioned analysis [25]. On the finance side, the work of Bhardwaj et al. [6] studies a history of commodity futures which mirrors our own efforts to study cryptocurrency futures. In cryptocurrencies, the closest related work comes from Gandal et al. [20] who performed a postmortem analysis of the Mt. Gox bitcoin exchange. They had the benefit of the exchange's back-end database, while we sourced various public signals to reconstruct BitMEX's history. Moore and Christin [36] looked at early cryptocurrency exchanges (2008–2013), and observed that anti-money laundering precautions were rare, and exchanges were frequently compromised. Our work, almost a decade later, shows that, while the financial instruments have become far more complex, cryptocurrency traders' risk appetite remains high. Decentralized exchanges have recently been the focus of a number of research papers, in particular, on how to attack them.

For instance, Daian et al. [16] examined various attempts at gaming decentralized platforms for profit; while important, such attacks are less relevant in the context of centralized platforms such as BitMEX. Last, from a methods standpoint, we build upon the methodology of Meiklejohn et al. [34] and Möser et al. [37] for clustering addresses and tainting flows. We were also inspired by Kogan et al. [27] and by Loughran and McDonald [30] for relating the price signals of a market to sentiment within textual data.

9 CONCLUSION

Through the innovation of its complex yet intuitive perpetual futures instrument, BitMEX became a multi-billion dollar exchange that transformed the landscape for cryptocurrency derivatives. While we cannot affirm that derivatives products like the ones offeed on BitMEX are responsible for the rapid price jumps that have become commonplace in Bitcoin, our analysis suggests that these derivatives, through excessive leverage and cascading liquidations are supportive of them. We also confirm that these derivatives instruments attract a culture of long-biased highly leveraged speculators. However, clustering shows that BitMEX is also home to many professional outfits that control thousands of Bitcoins and manage dozens of accounts. Smaller traders disproportionately account for liquidations, and chatbox evidence suggests that many users are obsessively trading 24/7. All of this raises concerns about the impact that derivatives have on BitMEX's customers and on the cryptocurrency ecosystem as a whole. The flip side of the coin is that these phenomena, and possible responses to interventions, are far easier to measure in the context of cryptocurrencies, than they are in more traditional markets.

ACKNOWLEDGMENTS

We thank our anonymous reviewers for helpful feedback, and Flood (https://twitter.com/ThinkingUSD) for insightful discussions. This research was partially supported by Ripple's University Blockchain Research Initiative (UBRI) at Carnegie Mellon. Some of the authors hold non-negligible cryptocurrency positions, but none on the BitMEX platform.

REFERENCES

[1] 2013. Korean Language. *Ethnologue (17th ed.)* (2013).
[2] 2020. United States of America vs. Arthur Hayes, Benjamin Delo, Samuel Reed, and Gregory Dwyer. United States District Court, Southern District of New York. Indictment 20Cr.500.
[3] 2021. Global Charts for Cryptocurrency Futures Markets. https://coinalyze.net/futures-data/global-charts/.
[4] Proton Technologies AG. 2020. Secure email: ProtonMail is free encrypted email. https://protonmail.com/.
[5] Beyond10x. 2018. Bitmex CEO Arthur Hayes Crypto Millionaire Explains How He Created A Multi Million Dollar Company. https://www.youtube.com/watch?v=Ljw9ulT2NHE.
[6] G. Bhardwaj, R. Janardanan, and G. Rouwenhorst. 2019. The Commodity Futures Risk Premium: 1871–2018. *Available at SSRN 3452255* (2019).
[7] Binance. 2020. Bitcoin Exchange | Cryptocurrency Exchange. https://www.binance.com.
[8] Bitfinex. 2020. Cryptocurrency Exchange | Bitcoin Trading | Futures Trading | Margin Trading. https://www.bitfinex.com.
[9] BitMEX. 2020. Accelerating the BitMEX User Verification Programme. https://blog.bitmex.com/accelerating-the-bitmex-user-verification-programme/.
[10] BitMEX. 2020. How We Are Responding to the 13 March DDoS Attacks. https://blog.bitmex.com/how-we-are-responding-to-last-weeks-ddos-attacks/.
[11] Bitstamp. 2020. Buy and sell Bitcoin and Ethereum. https://www.bitstamp.com.
[12] Bybit. 2020. Bitcoin Ethereum Futures Trading | Cryptocurrency Exchange Platform. https://www.bybit.com.
[13] Circle Internet Financial. 2020. USDC: the fastest growing, fully reserved digital dollar stablecoin. https://www.circle.com/en/usdc.
[14] Coinbase. 2019. Coinbase Custody acquires Xapo's institutional business, becoming the world's largest crypto custodian. https://blog.coinbase.com/coinbase-custody-acquires-xapos-institutional-business-becoming-the-world-s-largest-crypto-2c1b46fc94c4.
[15] Coinbase. 2020. Buy & Sell Bitcoin, Ethereum, and more with trust. https://www.coinbase.com.
[16] P. Daian, S. Goldfeder, T. Kell, Y. Li, X. Zhao, I. Bentov, L. Breidenbach, and A. Juels. 2020. Flash Boys 2.0: Frontrunning in Decentralized Exchanges, Miner Extractable Value, and Consensus Instability. *Proc. 2020 IEEE Symp. Sec. Priv.* 910–927.
[17] Deribit. 2020. Bitcoin Futures and Options Trading. https://www.deribit.com/.
[18] A. Fifield. 2019. Leaving Wall Street, Entering Crypto Chaos Sam Bankman-Fried. https://youtu.be/gSDk5PAJss4?t=2669.
[19] FTX. 2020. Cryptocurrency Derivatives Exchange. https://www.ftx.com.
[20] N. Gandal, J. Hamrick, T. Moore, and T. Oberman. 2018. Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics* 95 (2018), 86–96.
[21] Glassnode. 2020. On-chain market intelligence. https://glassnode.com/.
[22] Huobi Global. 2020. Safe Bitcoin Ethereum & Litecoin Exchange. https://www.huobi.com.
[23] Grayscale. 2020. Bitcoin Trust. https://grayscale.co/bitcoin-trust/.
[24] Grayscale. 2020. Digital Large Cap Fund. https://grayscale.co/digital-large-cap/.
[25] A. Hayes. 2019. BitMEX Leverage Statistics, April 2019. https://blog.bitmex.com/bitmex-leverage-statistics-april-2019/.
[26] Y. Kim. 2014. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882* (2014).
[27] S. Kogan, D. Levin, B. Routledge, J. Sagi, and N. Smith. 2009. Predicting risk from financial reports with regression. *Proc. 2009 Annual Conf. North Am. Chap. Assoc. Comp. Linguistics.* 272–280.
[28] Kraken. 2020. Bitcoin & Cryptocurrency Exchange | Bitcoin Trading Platform. https://www.kraken.com.
[29] M. Lerner, M. Hougan, H. Kim. 2019. Economic and Non-Economic Trading In Bitcoin: Exploring the Real Spot Market For The World's First Digital Commodity. https://www.sec.gov/comments/sr-nysearca-2019-01/srnysearca201901-5574233-185408.pdf.
[30] T. Loughran and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance* 66, 1 (2011), 35–65.
[31] Maker Ecosystem Growth Holdings. 2020. A better money: Digital currency that can be used by anyone, anywhere, anytime. https://makerdao.com/en/.
[32] C. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. *Proc. 52nd Ann. Meeting Assoc. Comp. Ling.: System Demonstrations.* 55–60.
[33] M. Masters. 2008. Testimony before the committee on homeland security and governmental affairs. *US Senate, Washington, May* 20 (2008).
[34] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. Voelker, and S. Savage. 2013. A fistful of bitcoins: characterizing payments among men with no names. *Proc. 2013 Internet Measurement Conf.*. 127–140.
[35] T. Mikolov, Q. Le, and I. Sutskever. 2013. Exploiting similarities among languages for machine translation. *arXiv preprint arXiv:1309.4168* (2013).
[36] T. Moore and N. Christin. 2013. Beware the Middleman: Empirical Analysis of Bitcoin-Exchange Risk. *Proc. 2013 IFCA Financial Crypto.*. Okinawa, Japan.
[37] M. Möser, R. Böhme, and D. Breuker. 2013. An inquiry into money laundering tools in the Bitcoin ecosystem. *Proc. 2013 APWG eCrime*, 1–14.
[38] P. Mourdoukoutas. 2018. How To Find The Next Ripple, Ethereum, Monero and EOS - Try The Kimchi Premium. https://www.forbes.com/sites/panosmourdoukoutas/2018/04/29/how-to-find-the-next-ripple-ethereum-monero-and-eos-try-the-kimchi-premium/.
[39] S. Nakamoto. 2008. Bitcoin: A peer-to-peer electronic cash system.
[40] Nomics. 2021. Top Crypto Exchanges Ranked By Volume | Nomics. https://nomics.com/exchanges.
[41] T. Odean. 1998. Are investors reluctant to realize their losses? *The Journal of Finance* 53, 5 (1998), 1775–1798.
[42] FTX Official. 2019. Why Does Bitcoin's Price Surge and Plunge so Much? https://www.youtube.com/watch?v=WdApztsrn-E.
[43] OFX. 2020. Historical Exchange Rates Tool & Forex History Data. https://www.ofx.com/en-us/forex-news/historical-exchange-rates/.
[44] OKEx. 2020. Cryptocurrency Exchange | Bitcoin Exchange | Crypto Exchange | BTC Exchange. https://www.okex.com.
[45] Tether Operations. 2020. Digital money for a digital age. Global, fast, and secure. https://www.tether.to.
[46] J. Poon and T. Dryja. 2016. The bitcoin lightning network: Scalable off-chain instant payments. http://lightning.network/lightning-network-paper.pdf.
[47] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, and C. Manning. 2020. Stanza: A Python Natural Language Processing Toolkit for Many Human Languages. *Proc. 58th Ann. Meeting Assoc. Comp. Ling.: System Demonstrations.*
[48] U. Raghavan, R. Albert, and S. Kumara. 2007. Near linear time algorithm to detect community structures in large-scale networks. *Phys. Rev. E* 76, 3 (2007), 036106.

[49] Alameda Research. 2019. Investigation into the Legitimacy of Reported Cryptocurrency Exchange Volume. https://ftx.com/volume-report-paper.pdf.
[50] J. Wilder. 1978. *New concepts in technical trading systems*. Trend Research.
[51] Marlamin xPaw. 2020. Steam Charts and Stats - Concurrent Steam Players - SteamDB. https://steamdb.info/graph/.
[52] Z. Yang, R. Algesheimer, and C. Tessone. 2016. A comparative analysis of community detection algorithms on artificial networks. *Sci. Reports* 6 (2016), 30750.