JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 3(2), 2017, 127–135
doi: 10.1093/cybsec/tyx007
Advance Access Publication Date: 1 August 2017
Research paper

Research paper

The price of anonymity: empirical evidence from a market for Bitcoin anonymization

Malte Möser[1,*] and Rainer Böhme[2]

[1]Department of Computer Science, Princeton University, 35 Olden Street, Princeton, NJ 08540, USA and [2]Department of Computer Science, Universität Innsbruck, Technikerstrasse 21A, 6020 Innsbruck, Austria

*Corresponding author: E-mail: mmoeser@princeton.edu

Received 13 May 2017; accepted 15 June 2017

Abstract

We present the first measurement study of JoinMarket, a growing marketplace for more anonymous transfers in the Bitcoin ecosystem. Our study reveals that this market is funded with multiple thousand bitcoins and generated a turnover of almost 29.5 million USD over the course of 13 months. Assessing the resilience of the market against a well-funded attacker, we discover that in a typical scenario, a selective attack with a 90% success rate requires an investment of 14 000–54 000 USD (which is recoverable after the attack). We present economic arguments to explain the existence of this novel market for anonymity and underpin the hypothesis of heterogeneous time preference with empirical data.

Key words: economics of privacy; measurement; anonymity; Bitcoin; CoinJoin

Introduction

Anonymity and economics are an odd couple. Most microeconomic models assume agents without name, and fail to predict outcomes if agents become identifiable [1, 2]. Anonymity can be defined as the state of being "not identifiable within a set of subjects, the anonymity set" [3]. With this definition, a simple measure of the quality of anonymity is the size of this set, as the probability of successful identification by random guessing is inversely proportional to the set size.

If there exist situations where the state of being anonymous improves an agent's wealth, one would expect a market for anonymity to develop. Yet anonymity is an unconventional economic good. To produce it, other agents must behave in an indistinguishable way to an observer of the agent who seeks anonymity. This is nicely summarized in the expression "anonymity loves company" [4]. In the language of economics, the production of anonymity generates positive externalities because all agents who contribute to the supply of anonymity also receive the good in demand. Production and consumption are hard to tell apart. So, what should the market price for anonymity be?

While the economics of privacy [5] and personal data [6] have been studied for long, and empirical research has estimated price information (see [7] and [8] for reviews), remarkably little is known about the price of anonymity. Acquisti *et al.* [9] study the incentives to participate in anonymous communication systems based on mix networks, such as Tor. Their analysis is comprehensive, includes attacker behavior and adoption dynamics, but remains theoretical. Spiekermann [10] interprets survey data collected by self-selection among the early-adopters of an academic anonymous communication system (JAP). Köpsell [11] uses technical measurements in an experimental setup of the same system to approximate the value of anonymity by observing users' aggregate willingness to trade performance for anonymity. (Time-trade-offs were used subsequently to quantify the value of security, e.g., [12]). All these works contribute interesting observations, but barely scratch the surface of the puzzles associated with markets for anonymity.

This article makes a modest next step by leveraging the cryptographic currency Bitcoin and its ecosystem as a "social science laboratory" [13]. We present a longitudinal measurement study of a market designed to match supply and demand of anonymous value transfers. In principle, Bitcoin transactions can be traced and histories inspected for known identifiers that allow informed parties to associate the initially pseudonymous account numbers with real-world identities. JoinMarket, our object of study, offers a clandestine

marketplace to arrange a special kind of transaction that mixes transfers of many different parties, thereby exponentiating the complexity of deanonymization attempts. All parties involved in such a transaction roughly form an anonymity set as used in the above definition of anonymity.

Our approach draws on a combination of methods to answer a number of research questions. We collected price information from the public order book of this market between June 2015 and June 2016. Moreover, we obtain volume information by matching changes in the order book with likely trades in Bitcoin's public ledger. To validate our method with ground truth, we participated on the supply side of the market at selected points in time using a minimal invasive trading strategy. This combination of methods allows us to quantitatively describe the market development over time. Our second contribution is on the economics of security. We study the theoretical possibility of an attacker participating in the market and estimate the cost of deanonymization over time. This cost is expressed in terms of capital employment and as a function of the targeted probability of success. We note that adversaries may even profit from launching attacks, generated from fees paid for the (then empty) promise of better anonymity. To better understand this and other anomalies of markets for anonymity, we formulate stylized economic models of supply and demand. The first model uses time preferences and the second model uses qualitative differentiations to explain the existence and price formation on this market. Where possible, we underpin the underlying hypotheses with data from JoinMarket.

The remainder of this article is organized as follows. We offer the necessary technical background about Bitcoin, CoinJoin transactions, and JoinMarket in Section "Background". Our measurement approach, descriptive results, and the hypothetical attack scenario are presented in Section "Measurements". Section "Economics of JoinMarket" discusses economic models devised to explain the observed anomalies. The article closes with a brief discussion (Section "Discussion") and concludes in the last section.

Background

Bitcoin in a nutshell

Bitcoin is the first, and to this date most popular, instance of a decentralized cryptographic currency [14]. A key characteristic of this first generation of cryptographic currencies is their public ledger, replicated on every "full" node of a peer-to-peer network [15, 13]. This ledger, called blockchain, contains the records of all transactions that have ever taken place in the system. Since only transactions reassign ownership of bitcoins and each transaction references all relevant previous transactions, it is possible to validate the state of the public ledger by following the references backwards. A probabilistic consensus protocol resolves conflicting updates at the end of the ledger. Its design, incentive mechanisms, and security properties are vital for the system but irrelevant for this article.

The Bitcoin protocol assigns value, denominated in bitcoins (BTC), to addresses. Bitcoin addresses serve like account numbers. They are derived from the public keys of an asymmetric encryption system. Ownership of accounts is controlled by the knowledge of the corresponding private keys. As anyone can generate fresh key pairs, Bitcoin users enjoy a degree of pseudonymity. However, with all addresses and the associated transactions stored in the public blockchain, an observer can identify relations between them. If this knowledge is enriched with auxiliary information on the real-world identities behind addresses, it becomes possible to deanonymize selected users [16–19].

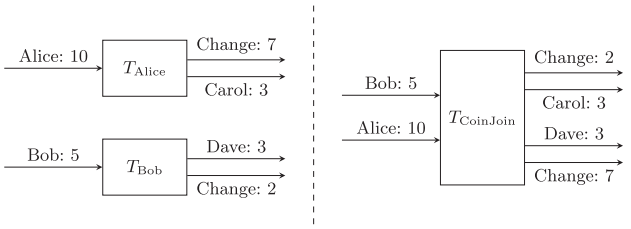

Figure 1. Two or more individual payments (left) can be combined in a single CoinJoin transaction (right). The spending amounts need not be identical, but in JoinMarket they are.

Transactions in Bitcoin specify the origin and the destination of the value transferred. They include a list of inputs, which are references to existing funds, and a list of outputs that specify its new owners. Whenever a user transfers bitcoins, she has to spend the full value of the input and therefore returns any surplus as "change" to an address under her control. A common deanonymization technique is to cluster addresses that are associated with inputs combined in one transaction, as this behavior suggests common knowledge of all the associated private keys [16, 17, 19].

CoinJoin, a special convention for transactions, intentionally breaks this heuristic [20]. Multiple senders and recipients of funds combine their payments in a single joint transaction (cf. Fig. 1). This is possible and secure because valid Bitcoin transactions require individual signatures for each public key associated with the funds used. If CoinJoin was default in Bitcoin, it would increase users' privacy by rendering the aforementioned multi-input heuristic more difficult to apply. A limitation remains: individual values may still leak sufficient information to derive matching subsets [21, 22].

JoinMarket

A major barrier toward the adoption of CoinJoin is to match users who are interested in creating a transaction. This opens up an opportunity for intermediaries. JoinMarket [23] is a platform for Bitcoin users wishing to participate in CoinJoin transactions. It has been operational since May 2015 [24]. In contrast to previous approaches [25, 26], JoinMarket does not aim at matching different users who all want to create a transaction at the same time. Instead, it divides users in two groups: (market) "makers" and "takers" of CoinJoin offers. Makers offer their bitcoins for use in takers' CoinJoin transactions. The advantage of this approach is that users do not have to wait for partners when creating CoinJoin transactions. Instead they can choose from a list of offers by the market makers. To incentivize participation, takers pay makers a small compensation (further referred to as "maker fee" to be distinguished from the general "miner fee" in Bitcoin).

The technical backend of JoinMarket is rather simple. It does not use mixing protocols such as Xim [27], CoinShuffle [28], CoinParty [29], or Mixcoin [30] that maintain decentralization, include defenses against Sybil attacks, or at least provide warranties. In the current implementation, takers and makers communicate via a centralized Internet Relay Chat (IRC) channel. Whenever a maker joins the channel she announces her offers to the channel. JoinMarket has no central server that stores a list of available offers; the offer announcements over time form a public order book that every member can maintain locally.[1] Whenever a maker announces or updates her offers, or leaves the channel, the local database is

1 Some websites allow to inspect the order book, e.g., https://joinmarket.me/ob/.

updated. A maker wishing to participate in CoinJoins must stay connected to the IRC server, typically by running an IRC bot that automatically announces offers and updates.

Each offer in the order book is uniquely identified by the combination of a username and an offer identifier (*uid*, *oid*). Every offer record further contains a maker fee f of type $t = \{\text{rel}, \text{abs}\}$, a minimum amount $v_{\min}$, a maximum amount $v_{\max}$, and a contribution to the miner fee c that has to be paid for the transaction to be included in the blockchain. Relative fees (with $t = \text{rel}$) are specified in percent of the value a taker intends to send, not the sum of her input values. The miner contribution is intended to reimburse the taker for the increase of the miner fee due to a larger transaction size. It gives more flexibility in specifying the total fees for the taker.

JoinMarket does not automatically match orders. A taker interested in creating a CoinJoin transaction will join the IRC channel, request the current list of offers and receive them from each individual maker in a private message. She will then choose a set of offers with one of the following methods:

1. Random draw from a distribution that weighs offers based on their fee;
2. deterministically choose offers with lowest fees; or
3. manual selection.

The first option is preferable over minimizing the fees because it avoids the risk that a single cheap offer dominates the market. After requesting a set of inputs and destination addresses from each participant, the taker constructs the CoinJoin transaction. Finally, she will publish the transaction on the Bitcoin network.

Two additional details are necessary to understand our analysis of JoinMarket. First, JoinMarket transactions are special cases of CoinJoin transactions with characteristics that allow us to identify them in the transaction graph generated from the public ledger. If n participants construct a transaction, it will have n outputs with the exact same value (we call this the *spend*) and usually also the same number of "change" addresses. Takers can also choose to sweep their wallets and send all of their funds to an address. In this case there will only be $n - 1$ change outputs. There must also be at least n inputs,[2] each associated with a different Bitcoin address.

Second, JoinMarket makers use a deterministic wallet as specified in BIP 32 [31]. Different wallet chains separate so-called mixing "depths". Each spend is sent to an address belonging to the next depth, while the change stays in the current depth. This prevents the reuse of a spend/change address pair as common inputs in a future JoinMarket transaction. In the early days of JoinMarket, each maker would simply pick the wallet chain with the highest amount of bitcoins available. Nowadays many makers publish offers for each of the chains, distinguished by the *oid*. This raises interesting questions related to the pricing of individual offers. Makers offering large amounts, e.g., could demand higher fees as long as there is sufficient demand. Depending on the distribution of the spending values, different mixing depths may also be priced differently to merge funds in such a way that they fit the distribution. These questions require a richer analytical model and answering them is beyond the scope of this article.

Measurements

We now present what is to the best of our knowledge the first measurement study of JoinMarket.

2 Specifically, there must be n input subsets with a value greater or equal to the spend.

Data collection and preprocessing

We used JoinMarket's built-in order book watcher to monitor the available offers and stored a snapshot about every 5 min (288 snapshots per day) since the beginning of June 2015 until the end of January 2016, and every single minute (1440 snapshots per day) from February 2016 until the end of June 2016. From these, we extract all individual offers (44 million entries in total). Due to IRC disconnects and crashes of the order book watcher, we miss data for a few timestamps. In total, our dataset covers 99.48% of the whole timeframe.

One issue when analyzing the order book is that it is impossible to verify whether stated offers are indeed genuine – makers could easily overstate the amount of bitcoins they offer, or offer low fees and then fail to deliver. In principle, makers could even serve the market without ever stating offers on the public channel (i.e. only make private statements to takers requesting the order book). Still, we assume that the majority of offers is genuine and visible in the public order book. For data cleaning, we decided to remove offers with the *uid* "fakeorder", which claimed to offer up to 2.1 million bitcoins (i.e. 13% of all bitcoins in circulation). We also removed offers with a maximum amount lower than or equal to zero. To accommodate the risk of outliers due to short-lived exaggerated offers in the order book, we calculate the median over a time interval of one day whenever we report aggregated values.

Besides taking snapshots of the order book, we also ran our own maker bot for multiple weeks. Running a maker bot allowed us to analyze the characteristics of real CoinJoin transactions without significantly influencing what we aim to measure. Of course, we cannot rule out the possibility that the transactions we attribute to normal users are the result of other researchers' participation. Due to JoinMarket's protocol for constructing transactions, a maker does not collect information that would allow to directly associate inputs or outputs with specific users. We therefore consider participating on the maker side a reasonable trade-off between the takers' desire for anonymity and our interest in identifying the characteristics of this market. In total, we participated in 498 CoinJoin transactions with spends <0.5 BTC. This number was largely endogenous due to the behavior of other market participants. From the point of view of research ethics, we would have liked to keep this number lower. However, the market was bumpy at times and manual intervention would have compromised the reliability of the ground truth data. Note that participation as a maker (as opposed to being a taker) does not generate volume. If at all, it marginally increases the anonymity offered to market participants.

We then used one of the transactions we participated in as a starting point and traversed the transaction graph for other JoinMarket transactions between block heights 358 000 (end of May 2015) and 418 722 (end of June 2016). Our criteria for identifying JoinMarket transactions were n spending outputs (i.e. outputs with the same value) and n or $n - 1$ change outputs as well as at least n inputs with $n \geq 2$. We excluded some obvious false positives, such as transactions where all inputs belong to the same address or transactions where the largest input is necessary to create multiple spending outputs. With this technique, we identified 26 523 potential JoinMarket transactions between the beginning of June 2015 and the end of June 2016, which correctly include all ground truth transactions.

During our measurement period the Bitcoin exchange rate varied significantly, between 209 USD per BTC in August 2015 and 768 USD in June 2016 [32]. Whenever we report USD values, we use an exchange rate of 400 USD per BTC (unless stated otherwise), roughly the average exchange rate in early 2016.

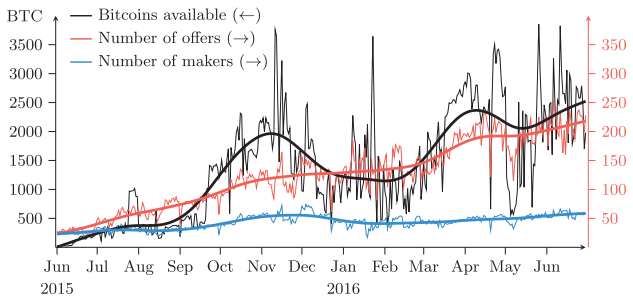

Figure 2. Key indicators of the JoinMarket order book.

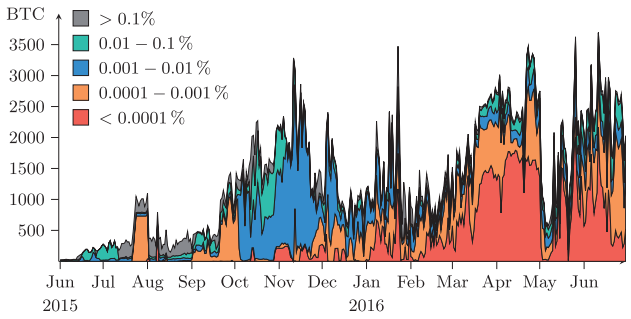

Figure 3. Amount of bitcoins available on JoinMarket at a certain maker fee.

Market overview

Figure 2 plots descriptive statistics of the JoinMarket order book over time. We report both the total amount of bitcoins available as well as the number of offers and makers. The thin lines connect daily medians. Thick lines are fitted smoothing splines with ten degrees of freedom. The total number of bitcoins available rose from initially a few hundred in mid-2015 to >2000 in November 2015, dropping to values between 1000 and 1500 in January 2016 and rising >2000 again in mid-2016. Overall there has been a steady increase in the number of offers, while the actual number of makers stays relatively stable. Our interpretation is that makers adopt more advanced bots which offer bitcoins at different mixing depths.

While a four-digit figure of bitcoins available already suggests a large market (e.g. 2000 BTC correspond to roughly 800 000 USD), it is more instructive to look at the maker fees at which those bitcoins are available. Figure 3 breaks down the available bitcoins by maker fees in percent of the transaction amount. Absolute fees are converted to relative terms using the maximum available amount. Observe that the majority of bitcoins is available for a maker fee <0.01%. In comparison, centralized services offering coin anonymization often charge fees between 1% and 3% [33].

Transaction volume

It is not directly observable from the order book how many of the offers have been accepted. We present two estimates for the total number of JoinMarket transactions.

Our first approach is to identify transactions in the blockchain that are relatable to our ground truth JoinMarket transactions. This strategy is valid because the outputs of other makers are usually reused as inputs in other JoinMarket transactions. By traversing the transaction graph in both directions, following the inputs and outputs of transactions and identifying potential JoinMarket transactions, we obtain a set of 26 523 JoinMarket transactions. The total spending value of these transactions amounts to 66 288 BTC, which corresponds to almost 29.5 million USD (converted at daily exchange rates). While we cannot rule out that the Bitcoin transaction

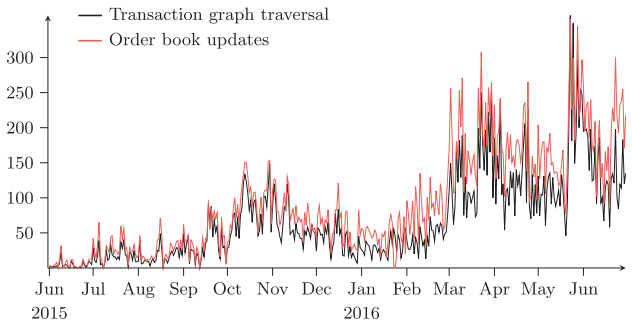

Figure 4. Estimated daily JoinMarket transaction volume based on transaction graph traversal and changes in the order book.

graph contains other subgraphs separate from our own transactions in the relevant timeframe, it is unlikely that we missed larger subgraphs due to the conjunctive nature of CoinJoin transactions.

Our second strategy is based on changes in the order book. Whenever a maker took part in a JoinMarket transaction, she updates her public offer(s) as the bitcoins available in her wallet chains have changed. We count 244 958 changes of offers in the order book between consecutive time stamps. We aggregate this data by the time stamp because most CoinJoin transactions involve multiple makers. This introduces a small probability of error if we aggregate offers belonging to different concurrent CoinJoin transactions. Given the transaction volume measured in the first approach, we are confident that the frequency of our timestamps is high enough to keep this error negligible. This approach yields an estimation of 36 393 JoinMarket transactions.

Figure 4 compares the estimated daily transaction volume for both methods over time. We see that in 2015 usage peaked between October and December at ∼150 transactions per day. Over the course of 2016 usage increased substantially, peaking at daily counts of >300 transactions. The substantial co-movement between both graphs ($r = 0.956$) suggests that our estimation heuristics are reliable. We conjecture that the larger estimates from the order book updates are due to manual interventions of makers and delayed updates of offers, which make them appear as separate transactions.

"Pay your attacker"

A major limitation inherent to JoinMarket is that an attacker can conduct a Sybil attack [34] to deanonymize takers. In such an attack, the attacker would impersonate a large number of makers to be the sole other participant in a CoinJoin transaction. As she knows which of the inputs and outputs belong to herself, she can attribute the remaining inputs and outputs to the taker. This renders the transaction as traceable as without CoinJoin. She could then sell this information to Bitcoin intelligence firms, for instance.

JoinMarket does not actively prevent Sybil attacks but relies on the market mechanism to make such an attack costly. Whenever a taker creates a CoinJoin transaction, she can choose one of the three offer selection methods mentioned in Section "JoinMarket". The default option, random draw with weighted probability function, selects offers over the course of multiple rounds depending on the number of makers chosen. First, it selects the cheapest matching offer from each maker. This limits the offers to choose from to one offer per maker so that makers with multiple offers gain no advantage in the selection process. Next, the offers are sorted by their total fee (i.e. adjusted maker fee minus the miner fee contribution) and each offer receives a probability based on an exponential function

parametrized to take into account the number of participants and the distribution of fees. Finally, an offer is chosen based on these selection probabilities. This process is repeated until the initially chosen number of makers is reached, in each round excluding all offers from maker(s) of previously chosen offers. This procedure bounds the success rate of a Sybil attacker by the number of coins at her disposal to outnumber all other offers for each potential spending amount.

We can estimate how many of the offers an attacker would have needed to control to be the sole participant of a user's CoinJoin. To this end, we first explore typical behavior by tabulating the choice variables from the JoinMarket transactions identified in the blockchain:

- the number of makers takers choose to join with and
- the distribution of spending values.

Figure 5 shows that most takers join with two makers. This was indeed the default until May 2016, when a new version of JoinMarket was released that by default would randomly select between 2 and 4 makers (we distinguish choices after this change in light blue). It is somewhat alarming to see that most users choose a low number of participants as this makes it easier to deanonymize the subsets [22]. Informed by this empirical distribution, we decided to calculate scenarios with 2, 3, and 5 makers. The spending values of the CoinJoin transactions follow a log-normal distribution (Fig. 6), from which we choose scenarios at the 25%, 50%, and 75% quantiles, which correspond to 0.038 BTC (15 USD), 0.211 BTC (84 USD), and 1.054 BTC (420 USD).

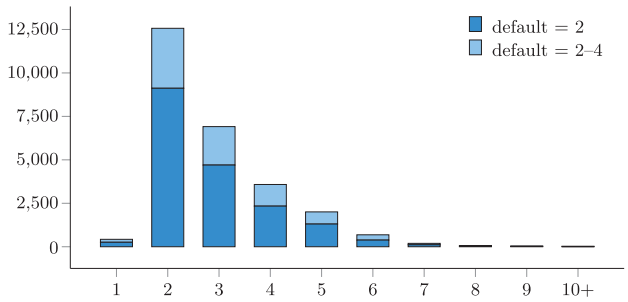

Figure 5. Number of makers chosen by the takers (the default was increased from 2 to 2–4 in May 2016).

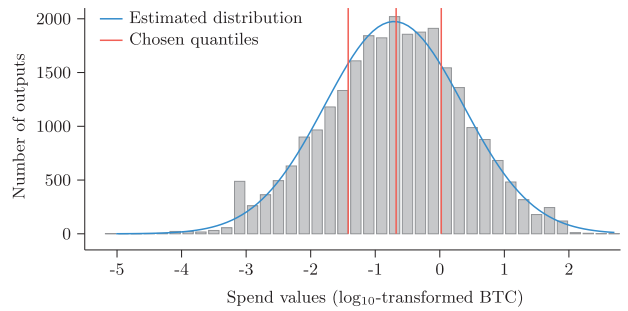

Figure 6. Spend values chosen by the takers follow a log-normal distribution.

We compute probabilities for all combinations of spending values and numbers of makers. To speed up the calculation, we only use a subset of our data, i.e. one snapshot per hour. Then we select the set of cheapest offers that, in total and across multiple rounds, yield a success rate of 90% for an imaginary attacker. To account for changing user behavior over time, we calculate these scenarios for three different periods that are motivated by shifts in the total number of bitcoins available (cf. Fig. 2) and the observed transaction volume (cf. Fig. 4): from June to September 2015, October 2015 to January 2016, and February to June 2016.

Table 1 reports the number of offers that are needed to reach a success probability of 90% in a given scenario defined by the taker's choice of spending value and number of participating makers m. For example, to be the sole other participant in a transaction that seeks three makers at a spending value of 1.054 BTC, the attacker would need to control (on average) between 12 and 17 of the cheapest offers. In general, we see that the number of offers needed grows with m. However, there is almost no difference between attacking transfers of 0.038 BTC, 0.211 BTC, or 1.054 BTC.

Next, we also extract the cumulative value of these offers, which gives us an indicator for the amount of bitcoins an attacker would have needed to control (cf. Table 2). For example, to reach a success probability of 90%, an attacker who wants to be the sole other participant in a CoinJoin transaction with $m = 3$ and a value of 0.211 BTC would require about 35 BTC (∼14 000 USD) in the first period, about 135 BTC (∼54 000 USD) in the second period and about 53 BTC (∼21 200 USD) in the third period (cf. Table 2). Note that this amount is not consumed but rather generates a profit during the attack. The actual cost of an attack is only the cost of capital. An attacker who operates in USD mainly faces exchange fees and possibly a risk premium for holding BTC instead of USD.

These figures are aggregates over our periods of observation. Figure 7 presents a longitudinal breakdown in relation to the overall value available on JoinMarket for the most common scenario with two makers and a success rate of 90%. Although the share of capital necessary to perform the attack has dropped significantly from the early days, it still remains between 4% and 20% in early 2016. At market volumes of 1500–2000 BTC, this yields a capital requirement in bitcoin equivalent to a medium five-figure to low six-figure USD amount.

Economics of JoinMarket

In the last section, we have described the properties of this market for anonymity and estimated the capital needed for an economic attack against the implied matching mechanisms. In this section, we draw on economic theory to explain why and under which

Table 1. Number of offers needed to be the sole other participant in a CoinJoin transaction with probability 0.9, depending on the number of participating makers m and the value transferred by the taker

Value (quantile)	m			m			m		
	2	3	5	2	3	5	2	3	5
0.038 BTC (25%)	9	13	17	11	16	23	11	18	23
0.211 BTC (50%)	10	14	19	11	16	25	11	15	26
1.054 BTC (75%)	9	12	15	11	16	22	11	17	23
Period	Jun–Sep'15			Oct'15–Jan'16			Feb–Jun'16		

Table 2. Cumulative value (in BTC) of the offers needed to be the sole participant in a CoinJoin transaction with probability 0.9, depending on the number of participating makers *m* and the value transferred by the taker

Value (quantile)	*m*			*m*			*m*		
	2	3	5	2	3	5	2	3	5
0.038 BTC (25%)	17.11	43.80	64.63	77.97	142.26	163.61	30.16	33.35	41.95
0.211 BTC (50%)	22.19	34.62	85.75	80.22	134.51	221.36	37.78	52.81	121.85
1.054 BTC (75%)	43.45	64.21	127.95	139.50	211.06	305.60	147.22	216.53	287.29
Period		Jun–Sep'15			Oct'15–Jan'16			Feb–Jun'16	

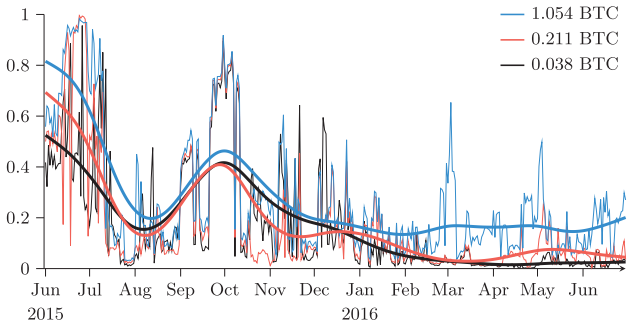

Figure 7. Share of volume present in 90% of all joins with two makers.

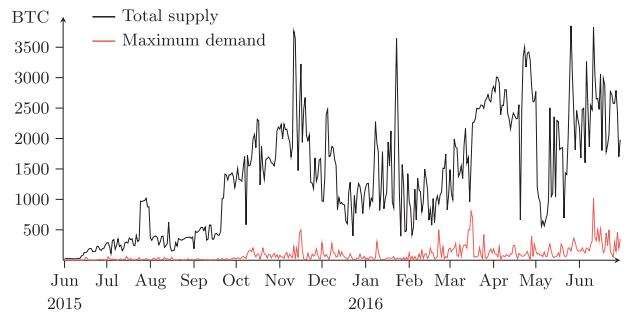

Figure 8. Supply and demand on JoinMarket.

conditions the market exists. We compare empirical facts against stylized models to validate our explanation attempts.

Excess supply

A market brings together demand and supply. Trades happen if agents on both sides differ in their preferences for bundles (q, v) composed of at least one tradable good (assumed homogeneous, divisible and with perfectly measurable quantity q) and monetary instruments of value v. Preferences change as trades happen, a relation typically expressed in demand and supply functions, which intersect at a market price p. Anonymity is a special good. Specifically, the service of participating in CoinJoin transactions increases q – think of the size of the anonymity set – for all involved parties. There is no simple explanation on why agents appear on both sides of a market. Why should some agents pay for the same service that others get paid for?

Indeed, the supply side of JoinMarket, that is where agents get paid, is more attractive. The red line in Fig. 8 visualizes the maximum concurrent demand per day. We compute the maximum demand based on the transactions extracted from the transaction graph. We aggregate the concurrent demand as the sum of spend values multiplied by the number of chosen makers per block, and then select the largest value for each day. The graph shows that at no point in time the demand came close to the total supply (the same series shown in Fig. 2 as bitcoins available). This confirms our initial doubt.

The excess supply corresponds well to the observable race to the bottom on maker fees (cf. Fig. 3). The following subsections model reasons why agents would appear on the demand side at all.

Time preferences

Agents may differ in their time preferences. In a stylized model, agents can be divided into two types. Type 1 wants to make an anonymous payment soon, e.g. in exchange for goods and services. Type 2 has a longer time horizon and wants to generate some return on her capital in Bitcoin. Type 1 is willing to pay a premium for immediate service, hence agents of this type appear on the demand side.

To underpin this theory with empirical evidence, we use a proxy for both types. We assume that agents of Type 1 use coins to fund a CoinJoin transaction faster than agents of Type 2, which rely on the market to request their funds for use in a transaction. Because takers first have to send their funds to JoinMarket's internal wallet to create a CoinJoin transaction, we compare the time difference between this funding and the spending transaction for a taker with the time gap to the previous transactions for the maker. To tell apart takers' inputs from makers' inputs, we use subset matching (cf. [21, 22]). We could unanimously identify the taker's subset of inputs and outputs for 18 218 transactions (67%). In these cases, the taker's subset is the combination that loses value. The makers' subsets usually increase as they collect a maker fee from the taker.

We perform a linear regression to evaluate the difference in the time gap between the previous transaction and the CoinJoin transaction dependent on funds belonging to a maker or a taker. The gap is measured in block intervals, Bitcoin's built-in time scale, where one unit corresponds to just <10 min on average. Because the overall distribution of time gaps is highly positively skewed, we log-transform the time gaps.

Table 3 presents the regression results. In the simple model (1) that only takes into account whether an input belongs to a maker or a taker, being a maker has a significant positive influence on the time gap, providing some early evidence for our explanation that takers are willing to pay for faster processing. To control for changing user behavior over time, we specify a second model that includes monthly fixed effects. Taking these effects into account (2) has only limited influence on the coefficient estimate, which gives us confidence that our model is robust.

In general, the importance of differences in time preferences should be inversely proportional to the liquidity. The higher the frequency of trades the more predictable is the time until a reasonably priced offer is taken, and the more attractive becomes participation on the supply side.

Table 3. Regression of the log-time gap (in blocks) to the previous transaction with (2) and without (1) monthly fixed effects (FE)

	without FE (1)	with FE (2)
(Intercept)	3.479***	3.340***
	(0.0128)	(0.0199)
Maker	0.604***	0.619***
	(0.0150)	(0.0148)
Monthly fixed effects	No	Yes
R^2	0.016	0.040
N	102 242	102 242

(Standard error in parentheses). Significance level code: ***p<0.001.

Qualitative differentiation

Time preference is not the only explanation. JoinMarket may indeed reflect the fact that bitcoins are not fungible. In simple terms, each individual bitcoin can be traced back to the unique block in which it was created ("mined" in jargon). What matters is that market participants can tell bitcoins apart. Arguably, bitcoins can have different value depending on their history (e.g. [35]). For example, increasing adoption of risk scoring may make it harder to spend funds that can be associated with a specific activity (e.g. criminal offenses) or whose transaction history contains patterns suggesting such use in the past [36]. Bitcoins are thus differentiable on a quality dimension, denoted in our model by $z \in [0, 1]$. Even if this quality is not directly observable, the perceived quality along with agents' expectations about the market valuation of coins of different quality can explain the existence of JoinMarket.

In the presence of qualitative differentiation, CoinJoin transactions, in particular those matched on open platforms like JoinMarket, suffer from adverse selection. Agents with known "good" coins ($z = 1$) must expect that joining with k "random" coins in the system, the expected quality (over the randomization of the selection) of the outgoing funds z' is

$$z' = \frac{1 + \bar{z} \cdot k}{1 + k}, \tag{1}$$

where $\bar{z}$ is the average quality of all coins in circulation. This expectation is very optimistic, because it assumes that the average quality of the coins available in JoinMarket equals the average in the population. In practice, the decision to offer coins on JoinMarket is endogenous. Agents are more likely to offer known bad coins to bet on the chance of getting better ones. More precisely, an agent strictly improves his wealth if $z < z^*$, where z^* is the agents' expectation about the average quality of the subpopulation of coins on JoinMarket. As other agents anticipate this behavior, a race to the bottom of z^* is triggered, leading to a collapse of CoinJoin platforms for good coins [37]. The only way out for a platform is to enable signaling (and commitment) to a minimum coin quality as part of the offers. JoinMarket did not support this in the study period.

In fact, the absence of signaling may indeed explain why agents appear on both sides. Besides impairing traceability, agents on the demand side may be willing to pay a premium for raising the average (or expected) quality of their funds. Agents on the supply side offer better coins and have (or expect) fewer difficulties in exchanging coins with lower quality for goods and services. This difference may be enough to outweigh the transaction costs of trade and hence justify the existence of JoinMarket.

We considered underpinning this theory with empirical data. However, we did not find a good proxy in our data to elicit private information about the (expected) quality from the participants. An avenue for future research is to test this theory with data generated by Bitcoin risk scoring and intelligence firms, such as Elliptic,[3] Chainalysis,[4] or Scorechain.[5]

Other explanations

The true reasons for trade activity on JoinMarket may be a combination of factors, including mundane ones. For example, operating a maker bot on JoinMarket's supply side requires more technical sophistication than taking offers on the demand side. For some agents, this transaction cost may be higher than the fees paid for receiving the service as taker. Related to this, the risk of operating a "hot wallet"[6] may not be offset by the low maker fees. Moreover, information asymmetries, in particular a lack of understanding of the anomalies of information goods, may keep users on the demand side. To some extent, information disregarding the adverse selection problem discussed above adds to the information asymmetries:

> In JoinMarket, normal legitimate investors in bitcoin just want to earn their coinjoin fee. They probably bought their coins from Coinbase.com or bitstamp and only want to earn an extra few % over time. These are the people you're mixing with when you create a coinjoin transaction. This makes JoinMarket unique, unlike any other private enhancer (DarkWallet, BitcoinFog, etc.) where you only mix with others who also want to improve their privacy.
>
> Source: [38]

Discussion

JoinMarket offers an innovative way of solving the matching problem for CoinJoin transactions. The current implementation is suboptimal in many respects.

On the technical side, the software is still in a premature state. It provides little protection against Sybil attacks, Denial-of-Service, or privacy-invasive behavior of market participants. A malicious taker could, e.g., repeatedly ask for funds from maker bots without using them in CoinJoin transactions. Knowing which outputs belong to the same taker then enables her to deanonymize other takers' CoinJoin transactions by telling apart the funds of all participating makers [39]. Recent countermeasures [40] aim at deterring such behavior by incurring a cost for repeated requests, but cannot fully mitigate the issue. There is also little protection against makers unwilling to sign transactions. Even if all market participants follow the protocol, the resulting CoinJoin transactions are identifiable in the block chain with ease, as demonstrated in Section "Data Collection and Preprocessing".

The architecture is centralized around a single IRC server, with all known disadvantages, and quite in contrast to Bitcoin's principle of decentralization. As people entrust this market tens and hundreds of bitcoins, it may only be a matter of time until serious issues arise; possibly adding another data point to the list of incidents where Bitcoin users lost money (cf. [41]). It remains an open research question to design a robust, accountable, decentralized matching market

3 https://www.elliptic.co, retrieved on 2016-03-04
4 https://chainalysis.com, retrieved on 2016-03-04
5 https://scorechain.com, retrieved on 2016-03-04
6 Hot wallets are installations where the private key is entrusted to a device that is connected to the Internet 24/7. Hot wallets are worthwhile targets for cybercriminals.

for CoinJoin transactions which offers some principled privacy guarantees.

On the economics side, JoinMarket leaves us puzzled on why trades exist on a matching market where makers and takers improve their anonymity alike. Moreover, many details of the market mechanism seem to follow ad-hoc approximations, e.g. the default offer selection method. It would be interesting to study the special properties of markets for anonymity through the lens of mechanism design.

A relevant cost factor in anonymizing Bitcoin transactions are miner fees. They internalize the externalities of securing the public ledger and must be paid for each transaction to be included in the blockchain. Whether or not the system can maintain a fee level low enough for consumer transactions is a hot debate at the time of writing (cf. [42, 43]). On JoinMarket, takers bear all of the miner fees, which often greatly exceed the maker fees. If miner fees rise, JoinMarket's future is uncertain because anecdotal evidence suggests that takers often use multiple CoinJoin's, hoping to increase the size of the anonymity set.

Finally, we note that participation in decentralized systems is often motivated by political or altruistic beliefs. One user described their motives in JoinMarket's public IRC channel as: "I am not really trying to make a profit. I just want my coins to do some good work and maybe help the joinmarket network." These noneconomic factors are hard to capture in an economic framework, but might be necessary to explain otherwise seemingly irrational behavior.

Conclusion

This article documents the first study on markets for anonymity, an interesting and rather unexplored phenomenon, which we found in a niche (JoinMarket) of a niche (cryptographic currencies).

Our longitudinal measurement study reveals a growing market for anonymous bitcoin transfers, funded with multiple thousand bitcoins at an average fee <0.01%. We find evidence for JoinMarket transactions worth almost 29.5 million USD in 13 months. The interpretation of this number is not straightforward because takers may go through multiple JoinMarket transactions subsequently to increase their anonymity. The total amount anonymized is unknown, but likely only a fraction of the headline figure.

We propose and investigate several economic explanations for the existence of such a market for anonymity. In accordance with our model, we observe excess supply and willingness to pay for faster anonymous payments on the demand side. It remains unclear if the agents are aware of qualitative differences in coins being traded, and some public statements suggest the contrary [38].

Anonymity in society has a Janus face. Undoubtedly, cryptographic currencies facilitate some forms of criminal activity [44, 13]. In this sense, JoinMarket is the cheapest way to launder money in Bitcoin we are aware of. It inherits the security of CoinJoin transactions against fraudulent counterparts and intermediaries. At the same time it inherits the lack of trust of markets over hierarchies [45], which backfires if the anonymity of transaction flows becomes a design goal next to security. We have shown that it is possible and affordable for realistic attackers to "buy" themselves into the CoinJoin transactions matched on JoinMarket. If law enforcement agencies consider this a viable and legitimate option, they should be careful to coordinate among each other. The attack scenarios described in this paper apply to cases with a single attacker only. If multiple attackers compete on the market, they may thwart each other's effort pretty effectively. With the data at hand, we cannot exclude that the relatively high headline figure of available offers in the order of 800 000 USD is already inflated by such operations.

Acknowledgements

The authors thank Chris Belcher and Adam Gibson for their feedback on an earlier version of this article, Christian Rückert for useful discussions, and the reviewers and participants of the 15th Annual Workshop on the Economics of Information Security (WEIS 2016) for their helpful comments.
This work was supported by the German Bundesministerium für Bildung und Forschung (BMBF) [grant agreement No 13N13505] and Archimedes Privatstiftung, Innsbruck.

References

1. Villas-Boas JM. Dynamic competition with customer recognition. *RAND J Econ* 1999;**30**:604–31.
2. Acquisti A, Varian HR. Conditioning prices on purchase history. *Market Sci* 2005;**24**:367–81.
3. Pfitzmann A, Köhntopp M. Anonymity, unobservability, and pseudonymity – a proposal for terminology. In: Federrath H (ed.), *Designing Privacy Enhancing Technologies*, Vol. **2009**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2001, 1–9.
4. Dingledine R, Mathewson N. Anonymity loves company: usability and the network effect. In: Anderson R (ed.), *Proceedings of the Fifth Workshop on the Economics of Information Security (WEIS 2006)*. UK: Cambridge, 2006.
5. Posner RA. The economics of privacy. *Am Econ Rev* 1981;**71**:405–9.
6. Laudon KC. Markets and privacy. *Commun ACM* 1996;**39**:92–104.
7. Lesk M. The price of privacy. *IEEE Secur Priv* 2012;**10**:79–81. September
8. Acquisti A, Brandimarte L, Loewenstein G. Privacy and human behavior in the age of information. *Science* 2015;**347**:509–14.
9. Acquisti A, Dingledine R, Syverson P. On the economics of anonymity. In: Wright RN (ed.), *Financial Cryptography and Data Security*, Vol. **2742**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2003, 84–102.
10. Spiekermann S. The desire for privacy: insights into the views and nature of the early adopters of privacy services. *Int J Technol Hum Int* 2005;**1**:74–83.
11. Köpsell S. Low latency anonymous communication – how long are users willing to wait? In: Müller G (ed.), *Emerging Trends in Information and Communication Security (ETRICS 2006)*, Vol. **3995**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2006, 221–37.
12. Egelman S, Molnar D, Christin N *et al*. Please continue to hold: an empirical study on user tolerance of security delays. In: *Proceedings of the 9th Workshop on the Economics of Information Security (WEIS 2010)*. Cambridge, MA: American Economic Association (AEA), 2010.
13. Böhme R, Christin N, Edelman B, *et al*. Bitcoin: economics, technology, and governance. *J Econ Perspect* 2015;**29**:213–38.
14. Nakamoto S. *Bitcoin: A Peer-to-Peer Electronic Cash System*. 2008. https://bitcoin.org/bitcoin.pdf (4 March 2016, date last accessed).
15. Bonneau J, Miller A, Clark J *et al*. Research perspectives and challenges for Bitcoin and cryptocurrencies. In: *2015 IEEE Symposium on Security and Privacy*. San Francisco, CA, USA: IEEE, 2015.
16. Ron D, Shamir A. Quantitative analysis of the full Bitcoin transaction graph. In: Sadeghi AR (ed.), *Financial Cryptography and Data Security*, Vol. **7859**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2013, 6–24.
17. Reid F, Harrigan M. An analysis of anonymity in the Bitcoin system. In: Altshuler Y, Elovici Y, Cremers A, Aharony N and Pentland A, (eds.) *Security and Privacy in Social Networks*. New York: Springer, 2013, 197–223.
18. Meiklejohn S, Pomarole M, Jordan G *et al*. A fistful of bitcoins: characterizing payments among men with no names. In: *Proceedings of the 2013 Internet Measurement Conference (IMC), October 2013*, pp. 127–140, ACM, 2013.
19. Harrigan M, Fretter C. The unreasonable effectiveness of address clustering. In: *The 13th IEEE International Conference on Advanced and Trusted Computing*. Toulouse, France: IEEE, 2016.

20. Maxwell G. *CoinJoin: Bitcoin Privacy for the Real World*. 2013. https://bitcointalk.org/index.php?topic=279249.0 (4 March 2016, date last accessed).
21. Atlas K. *Weak Privacy Guarantees for SharedCoin Mixing Service*. 2014. http://www.coinjoinsudoku.com/advisory/ (4 March 2016, date last accessed).
22. Meiklejohn S, Orlandi C. Privacy-enhancing overlays in Bitcoin. In: Brenner M, Christin N, Johnson B, and Rohloff K (eds.), *Financial Cryptography and Data Security, 2nd Workshop on BITCOIN Research*, Vol. **8976**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2015, 127–41.
23. *JoinMarket-Org/joinmarket: CoinJoin Implementation with Incentive Structure to Convince People to Take Part*. https://github. com/JoinMarket-Org/joinmarket (1 March 2016, date last accessed).
24. Belcher C. *JoinMarket release on mainnet*. 2015. https://www. reddit. com/r/joinmarket/comments/358dlv/joinmarket_released_on_mainnet/ (4 March 2016, date last accessed).
25. Blockchain. *Shared Coin*. https://www.sharedcoin.com/ (4 March 2016, date last accessed).
26. *Darkwallet*. https://www.darkwallet.is/ (4 March 2016, date last accessed).
27. Bissias G, Ozisik AP, Levine BN, Liberatore M. Sybil-resistant mixing for Bitcoin. In: *Proceedings of the 13th ACM Workshop on Workshop on Privacy in the Electronic Society*. ACM, 2014.
28. Ruffing T, Moreno-Sanchez P, Kate A. CoinShuffle: practical decentralized coin mixing for Bitcoin. In: *ESORICS'14. Proceedings of the 19th European Symposium on Research in Computer Security*. Vol. 8713. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2014, 345–64.
29. Ziegeldorf JH, Grossmann F, Henze M, Inden N, Wehrle K. Coin-Party: secure multi-party mixing of bitcoins. In: *Proceedings of the 5th ACM Conference on Data and Application Security and Privacy*. San Antonio, TX, USA: ACM, 2015, 75–86.
30. Bonneau J, Narayanan A, Miller A *et al*. Mixcoin: anonymity for Bitcoin with accountable mixes. In: Christin N and Safavi-Naini R (eds.), *Financial Cryptography and Data Security*, Vol. **8437**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2014, 486–504.
31. Wuille P. *Hierarchical Deterministic Wallets*. 2012. https://github.com/bitcoin/bips/blob/master/bip - 0032. mediawiki (4 March 2016, date last accessed).
32. CoinDesk. *Bitcoin Price Index*. http://www.coindesk.com/price/ (10 October 2016, date last accessed).
33. Möser M, Böhme R, Breuker D. An inquiry into money laundering tools in the Bitcoin ecosystem. In: *Proceedings of the APWG E-Crime Researchers Summit*. San Francisco: IEEE, 2013, 1–14.
34. Douceur JR. The Sybil attack. In: Druschel P, Kaashoek F and Rowstron, A (eds.), *Peer-to-Peer Systems*, Vol. **2429**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2002, 251–60.
35. *Looking to buy an old 50 BTC block. Where to buy?* 2015. https://np.reddit.com/r/Bitcoin/comments/3sg8vm/looking_to_buy_an_old_50_btc_block_where_to_buy/ (4 March 2015, date last accessed).
36. Möser M, Böhme R, Breuker D. Towards risk scoring of Bitcoin transactions. In: Böhme R, Brenner M, Moore T and Smith M (eds.), *Financial Cryptography and Data Security, 1st Workshop on BITCOIN Research*, Vol. **8438**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2014, 16–32.
37. Akerlof GA. The market for "lemons": quality uncertainty and the market mechanism. *Quart J Econ* 1970;**84**:488–500.
38. *With JoinMarket, You Mix with Clean, Untainted Bitcoins*. 2015. https://www.reddit.com/r/joinmarket/comments/38fi5m/with_joinmarket_you_mix_with_clean_untainted/ (4 March 2015, date last accessed).
39. Belcher C. *Bad Faith Taker Spy Not Filling Orders So That It Learns Which UTXOs Belong to Which Maker, Allowing Future Unmixing*. 2015. https://github.com/JoinMarket-Org/joinmarket/issues/156 (22 February 2016, date last accessed).
40. Gibson A. *Sourcing Commitments for Joins*. 2016. https://github. com/JoinMarket-Org/joinmarket/wiki/Sourcing-commitments-for-joins/d33bc1c300672cbd038b6bb98567ad9d8a9fa842 (14 October 2016, date last accessed).
41. Moore T, Christin N. Beware the middleman: empirical analysis of Bitcoin-exchange risk. In: Sadeghi AR (ed.), *Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers*, Vol. 7859. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2013, 25–33.
42. Croman K, Decker C, Eyal I *et al*. On scaling decentralized blockchains. In: Clark J, Meiklejohn S, Ryan P, Wallach D, Brenner M, and Rohloff K (eds.), *3rd Workshop on Bitcoin and Blockchain Research*. Financial Cryptography and Data Security. FC 2016, vol 9604. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2016, 106–125
43. Möser M, Böhme R. Trends, tips, tolls: a longitudinal study of Bitcoin transaction fees. In: Brenner M, Christin N, Johnson B and Rohloff K (eds.), *Financial Cryptography and Data Security, 2nd Workshop on BITCOIN Research*, Vol. **8976**. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, 2015, 19–33.
44. Christin N. Traveling the silk road: a measurement analysis of a large anonymous online marketplace. In: *Proceedings of the 22nd International World Wide Web Conference*. Rio de Janeiro: ACM, 2013, 213–24.
45. Williamson OE. Markets and hierarchies: some elementary considerations. *Am Econ Rev* 1973;**63**:316–25.