Is There Power in Numbers?
Network Effects and Cryptocurrency Prices

Diego Lara de Andres*

March 2020

Abstract

Whether or not there is a relationship between the number of users and the price of cryptocurrencies is yet to be established by the economics literature. A positive relationship is assumed by theoretical papers, but its form isn't stated. Empirical work has tried to identify this relationship giving little attention to the econometric problems that arise as a consequence of the non-stationarity of the data. Furthermore, parameter interpretation has not been clear. This paper provides a network based framework to understand this relationship, providing clear parameter interpretability. Empirically we focus on the Ethereum network, the second most valuable cryptocurrency, estimating this relationship taking into account non-stationarity. A long-run cointegrating relationship is established, which shows prices are consistent with each user being connected to about 30 others, at current network sizes. An error correction model is also estimated, showing that any discrepancy between observed prices and those implied by the long-run relationship is corrected on average by 10% the following day. Finally, inference is obtained using a range of bootstrapping methods allowing to reject the null of a full network, known as "Metcalfe's Law" as well as the absence of network effects.

Keywords: Cryptocurrencies, Networks, Cointegration and Inference.

*Department of Economics, University of Bristol, The Priory Road Complex, 12 Priory Rd, Bristol, BS8 1TU, United Kingdom. Email: d.laradeandres@bristol.ac.uk.

1 Introduction

Over a decade ago Nakamoto (2008) introduced a peer-to-peer system of payments, giving rise to decentralised electronic money in the form of cryptocurrencies. These have attracted the attention of governments, corporations and the general public. Traded in online exchange platforms, the prices of these digital currencies and the drivers behind their movement, has been of interest to the economics literature. In particular there is an open question of whether or not there exists a positive relationship between the number of users of a given cryptocurrency and its prices. This is an underlying foundation of much of the theoretical work. Even if this relationship is present, its exact form and the implications in terms of network connectedness are not clear.

Existing empirical evaluations focus on a relationship where prices are proportional to the number of users squared, commonly referred to as "Metcalfe's Law". Some generalise this to an exponential function of prices. Estimations of this relationship have failed to address the econometric challenges that exist. At a basic level, the non-stationarity of the data opens the question of spurious results, as regular estimation techniques can be inconsistent. Stationarity of the errors emerging from any model is essential, otherwise the estimated relationship will not be stable across time. Even if cointegration is established, this will be a long-run relationship which may not hold at any given point in time. Obtaining inference on the cointegrating parameters is non-standard. Furthermore, interpretation of this parameter has not been well studied. Within this empirical literature, the underlying modelling choices made to arrive at the estimated relationship, where prices are an exponential function of users, are not clear.

In this paper, we address these econometric issues as well as providing a network framework interpretation of the relationship empirically tested in the literature. We estimate the relationship between prices and number of users for the Ethereum network, the second most valuable cryptocurrency. Our empirical strategy starts by establishing the existence of a cointegrating relationship between users and prices. Then an error correction model is estimated, which allows us to look at both the long-run relationship and how it affects the short-run dynamics of these two variables. Inference challenges are addressed by looking at the variability of the errors of the cointegrating relationship, using a pairwise Continuous Block Bootstrap and a Hansen-Racine unit root test. Finally, inference is also achieved using a distributed lags model.

We show that a relationship like "Metcalfe's Law" emerges in a network context if prices reflect the sum of user utility, in particular when user utility depends exclusively on how many other users there are. This means that the parameter estimated in the literature is a measure of network density in a regular graph, where all users are the same. This imposes a constraint on network structure. In the Ethereum network, prices are cointegrated with active addresses, our measure of users. The obtained parameter implies that prices are consistent with each user being connected to at around 30 others, at the current network size. These long-run dynamics have significant effects in the short-run movements of prices and activity levels when these deviate away from the cointegrating relationship. All methods used to obtain inference point towards the same cointegrating parameter and confirm the presence of network effects whereby prices have an exponential relationship with users. "Metcalfe's Law" and the hypothesis of a fully connected network is rejected. This is analogous to rejecting a user utility function that depends exclusively on how many other users there are. However, confidence intervals for the parameter of interest are vast. So is the difference in the implications on network density at either end of the possible range. At the lower bound users are connected to just one other, while at the upper bound users connect with over 150 other Ethereum addresses.

The structure of the paper is the following. Section 2 covers some of the related literature. Section 3 provides the network framework for this relationship. The Ethereum network and the data used are discussed in Section 4. The empirical work is carried out in Section 5. First the long-run cointegrating relationship is established and then the error correction model is presented. Inference is obtained in Section 6 by applying several bootstrapping techniques. Section 7 concludes.

2 Related Literature

The economics literature has explored the factors and mechanisms influencing digital currency prices. Cheah and Fry (2015) test a stochastic bubble model concluding that the fundamental price of bitcoin is zero and exhibits speculative bubbles. More theoretically, Bjerg (2016) asks *"How is Bitcoin money?"* His answer, it is *"commodity money without gold, fiat money without a state, and credit money without debt"*. Arguing there is no intrinsic value in cryptocurrencies, Corbet et al. (2018) examines both Bitcoin and Ethereum, finding that there are *"periods of clear bubble behaviour."* However, Blau (2017) uses a model that relates volume and speculative trading, finding no evidence of speculative bubbles causing the rise and crash of bitcoin in 2013 or its volatility.

Supporting the idea that cryptocurrency prices are driven by more than speculation, Mitchnick and Athey (2018) suggest that demand for cryptoassets is driven by transactional and storage motives. Empirically they look at Bitcoin and Ripple finding support to their claims. Assuming adoption rates they find vast potential valuations for these digital currencies. Within these explanations for price dynamics a relationship with user numbers, although implicit, is already present. The speculation bubble mechanism requires of future buyers. Transactional motives will depend on liquidity, which is provided by users willing to buy and sell. In both cases, larger user numbers can be seen to make it easier to find someone to buy or sell at a given price in the future.

Pagnotta (2018) provides a theoretical framework that is based on two main pillars, lack of censorship and security provided by the mining process. The decentralised nature of the Bitcoin protocol provides protection against censorship, as there is no individual agent with the power to block any given transaction. This may be appealing to some users. Security of funds is achieved as there is a high computational cost required to succeed in the mining process, which gives the ability to include new transactions without the holder's secret key. Pagnotta (2018) also provides a link to modern monetary economics literature and how cryptocurrencies fit in. In an earlier version, Pagnotta and Buraschi (2018), a short theoretical outline of how user utility may be affected by other users is given. Our framework captures the same features while being more general.

Biais et al. (2018) provide a theoretical reasoning for investments in cryptocurrencies based on an overlapping generations model. In their set up, returns are driven by transactional costs and benefits. They also relate their approach to the monetary literature in some depth. While estimating their model, they provide an index of how easy it is to purchase goods and services using bitcoin. This affects prices and depends on the number of users.

Throughout these two influential papers, the number of users a cryptocurrency has affects positively its price. As a clear example, Pagnotta (2018) states that *"everything else being constant, an increase in the network size raises the bitcoin valuation for any given security level."* Along the same lines, crucial to their new index, Biais et al. (2018) state that *"as more firms started accepting bitcoins to buy goods and services, transactional benefits became larger."*

Empirically the literature has focused on a specific formulation of this relationship, known as "Metcalfe's Law", whereby the price is proportional of the number of users squared. Shapiro and Varian (1999) provides an early reference, exploring product usage and valuations. They simplify network effects in a reduced-form relationship where each user cares about all other users, and the value of the network is the sum of all users valuation of the network. Zhang et al. (2015) empirically test this relationship and find support for it using Facebook and Tencent data.

In the field of digital currencies, Alabi (2017) and Van Vliet (2018) have findings that support a relationship between prices and the number of users that follows "Metcalfe's Law". However, they apply methods that do not address non-stationarity challenges. Wheatley et al. (2018) estimates a slightly more general version, where a complete network isn't imposed, on Bitcoin data. This is similar to the one we estimate, but they fail to account for non-stationarity in their estimation and have a parameter interpretation that differs drastically to the one we obtain from the network framework.

Some literature has looked at the time series nature of the data, but has failed to interpret what their results mean for user connectedness. Civitarese (2018) finds no cointegration relationship between the number of Bitcoin users and prices.[1] Finally, Bhambhwani et al. (2019) perform a cointegration analysis on a group of cryptocurrencies, finding cointegration between prices and the number of users. They also introduce other factors, such as security, which limits parameter interpretability in terms of network density.

This paper formulates an explicit link between the more theoretical work on the drivers of cryptocurrency prices and the empirical work on the reduced form relationships between prices and the number of users. Introducing the network structure beyond the number of users opens the door to future research where characteristics such as centrality, the number and size of clusters as well as other network statistics may play a role in identifying which mechanisms are more important in the valuation process. The empirical application looks at the Ethereum network, the second most valuable cryptocurrency, which hasn't received as much attention as its older counterpart, Bitcoin. We estimate the relationship between prices and the number of users taking into account the challenges presented by the non-stationarity of the data, establishing cointegration. In a novel approach in this context, results for an error correction model are presented, where the long-run and short-run dynamics are estimated. Inference of this cointegrating parameter is obtained using several bootstrapping techniques, which focus on the variation of the errors. This highlights the limitations on parameter interpretability. Even though we can disregard complete and no connectedness, the confidence interval of how connected any user is remains large.

[1] We currently have work in process that will provide a justification for these results. We argue that their relationship appears unstable due to the sample period that is selected. With a larger sample and taking into account particularities of the Bitcoin network, the relationship is indeed stationary.

3 Price, Users and a Network Framework

The aim of this section is to introduce a new framework, based on a network approach to study the way price and users may be related. This can help better understand the origin and implications of relationships in the empirical literature. It will also make explicit possible relationship mechanisms implicit in the theoretical literature. We will first analyse the condition on supply needed to use "Metcalfe's Law", which is about network valuations, in the empirical literature, which looks at prices. We then develop the network based framework, using transferable utility and certain restrictions on network driven user utility. From this framework, we can see the implications that emerge from empirically tested relationships. A summarised version of the conditions imposed is given in Appendix A.1.

The empirical literature on this topic uses as a base to look at prices what is known as "Metcalfe's Law". This is based on a network valuation story, presented in Shapiro and Varian (1999). A reconciliation between prices and valuations is needed in order to use these two terms interchangeably. We define valuation by market capitalisation, the unit price multiplied by the total circulating supply. The necessary restriction is therefore on the supply. It is necessary to treat it as fixed.

Most cryptocurrency protocols, like Ethereum and Bitcoin, have an embedded supply function, which is public. For Ethereum every new block, designed to happen every 15 seconds, 2 Ether are created in the form of a mining reward. Another reward is also included for small side chains called uncles, but this is several orders of magnitude smaller than mining rewards. With this in mind, it is possible to treat the supply as a known function of time. And therefore treat it as fixed for our purposes, as there is perfect foresight. In the case of Bitcoin, not only is this function known, it also reduces block reward exponentially, so a maximum supply will be reached. After this the monetary base will be fixed. Treating prices and valuations as equivalent is supported empirically, our results are robust to using prices or market capitalisation.

Cryptocurrencies are usually the unit of account within a blockchain protocol. As such, their central intended purpose is to be the denomination for transactions between users. With this in mind, a network or graph can be defined by these transactions. Users, which hold cryptocurrencies, can be thought of as the nodes. Populating a set of vertices, V. Transactions of funds between users will generate links between them. These can be collected in a set of edges, E.

Mathematically, this network of a cryptocurrency is represented as a graph $\mathcal{G} := \mathcal{G}(V, E)$ where $m := |E|$ edges (transactions) are placed between $n := |V|$ vertices (users). Furthermore, we can define $a_{i,j}$ as the i, j entry of A, the adjacency matrix of $\mathcal{G}$. Transactions are directed and weighted edges. In order to simplify our set up, $a_{i,j} = 1$ if a transaction is made between i, j with $i \neq j$, 0 otherwise. Possible extensions introducing directed and weighted edges to this framework are clear avenues of future research.

Following Shapiro and Varian (1999), where network value is a sum of users' valuations and each individual cares about how many other users of the network there are, we can build our framework. More generally, the utility a user i gets from the network in a given day t is dependent on the transactions generated graph $\mathcal{G}_t$. We can relate prices and the network applying the concept of transferable utility, whereby users can transfer their utility using a common currency.

$$Price_t = \sum_{i=1}^{n_t} \mathcal{U}_i(\mathcal{G}_t). \tag{1}$$

From this point forward, focus lies on the user utility function and on the graph. Because of the level of anonymity that is usually present in these networks, and our data availability limitations, it isn't obvious what users care about. User utility may depend on a variety of network statistics, either about the user's position in the network or about the network in general. At the individual level, user valuation may be affected by measures such as centrality. In general terms, it may be important to users how distributed or decentralised the network is, measured by how many cliques there are and their size. Moreover, different users may value the network differently. It has been documented that there are different identifiable types of users within these networks. Tasca et al. (2018) find at least four different groups by activity, these include mining, gambling, exchange, the black market as well as others. In order to asimilate with this framework the theoretical and empirical literature that looks exclusively at the number of users, we impose further restictions. All users are homogeneous and their utility function is a stable weighted average of their direct links in that period t.

$$Price_t = \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} \beta a_{i,j,t}. \tag{2}$$

Due to user homogeneity and utility function stability all links are the same, hence there is no subscripts on the weight, β. Note that because of homogeneity there will also be graph regularity, all users have the same number of connections. Effectively, what (2) states is that the price is proportional to the total number of links in the network. This will depend on two factors, how big the network is and its density. As our focus is to evaluate and interpret the literature, that only focuses on number of users, the network density is restricted.[2] Introducing a network density parameter ϕ:

$$Price_t = \beta(n_t^{\phi} - n_t). \tag{3}$$

This covers relationships in the literature, such as in Wheatley et al. (2018), where price is an exponential function of the number of users. One special case is $\phi = 2$, "Metcalfe's Law":

$$Price_t = \beta n_t^2 - \beta n_t = \beta \sum_{i=1}^{n_t} (n_t - 1). \tag{4}$$

In this case, users care about how many other users there are, and price is a sum of users' valuations. Here the network is complete, with all users being connected. Note the network is large enough that the βn_t term becomes negligible.[3]

Using a relationship following (3) it is possible to test two key hypotheses. Whether there are any networks effects, $H_{01} : \phi = 1$ vs. $H_{A1} : \phi \neq 1$. Also whether "Metcalfe's Law" is present,

[2] This restriction is evaluated in Appendix A.4, where the total number of transactions is used. Our results show that this restriction on density may not be realistic. This highlights a limitation in the existing literature, as well as in the results presented in later sections. As a consequence of this, β may be underestimated.

[3] For estimation purposes, we drop this βn_t term, as our network has thousands of users. Estimation of both parameters will still be consistent, but a small negative bias will be present when n is small. This bias is smaller with larger values of ϕ.

H_{02} : $\phi = 2$ vs. H_{A1} : $\phi \neq 2$. Interpretation of this parameter can be divided into several regions.

If $\phi < 1$ user valuations would decrease when there are more users. This can be seen as a negative network effects scenario. Imagine a members clubs (the networks) which provide a nice room to be in (of the same size). Clubs with many members will have crowded rooms, where user utility is low. There will be two conflicting effects on the overall value of the clubs, which will make it unclear which club will be more valuable. Very populated clubs will have low utility per user but many users, while emptier clubs will have high user utility, but few users to sum across.

A network with $\phi = 1$ would be equivalent to what is called "Sornoff's Law". Users simply value the network at a constant level, β, regardless of the number of users. In this instance, there are no network effects, the number of users in the network doesn't affect user value. This can be consistent with a scenario in which users do not interact with each other and therefore there is no benefit from having more users in the network. Prices and users would follow a linear relationship.

In the case of digital currencies, and most social networks, the logical expectation is for $1 < \phi < 2$. This belief can be explained or mirrored in social network connections. Large social networks provide users with more utility than smaller counterparts, as more desired interactions become possible. However, this gain in utility is decreasing with size, tending to zero. Going from a small to a medium network delivers large gains in user utility, as users can find others that perhaps they wouldn't in a small network. Going from big to very big size may deliver very small user utility gains if users already had the desired number of links in the big network. Users want to be part of a social network where all their friends are present, but they don't need everyone they ever met to be in it. Note, the extra users don't reduce utility. This interpretation would translate to a scenario with saturation, where larger networks are sparser.

The case with $\phi = 2$, as covered above implies a fully connected network. With $\phi > 2$, network effects would be so extreme that larger networks not only allow more links between them and other users, but also links between other users that were previously not performed. For example, imagine friendship groups as networks, where links are formed when people play a sport together. First, a group of 10 friends can play basketball (5 links per game), but not football. Another group is of size 11, they can play basketball but also football. Adding this possibility, due to the size of the group, will be create links amongst those that would not have if the size was 10.

It is relevant to explore in more detail the scenario where $1 \leq \phi \leq 2$. To see the implications of this imposed structure on density across different network sizes, consider the fraction of links out of the total possible, $\kappa_{\mathcal{G}_\phi}$.

$$\kappa_{\mathcal{G}_\phi} = (n^\phi - n)/(n^2 - n). \tag{5}$$

From this measure we can retrieve the degree for each user, $d_{\mathcal{G}_\phi} = n \times \kappa_{\mathcal{G}_\phi}$, as well as the overall percentage of users in the network that are part of their neighbourhood (also $\kappa_{\mathcal{G}_\phi}$). Table 1 shows how the number of connections per user changes with ϕ and the number of users. Small variations in ϕ change drastically the number of connections. This strengthens the importance of inference, it matters how big the parameter confidence interval is. This is an issue overlooked in the prior literature, to which this paper devotes special attention.

One limitation of this framework is the possibility of endogeneity. The number of users or more

	ϕ						
	1.05	1.10	1.15	1.20	1.25	1.30	1.35
n							
100k	0.8	2.2	4.6	9	16.8	30.6	55.2
200k	0.8	2.4	5.2	10.5	20.1	37.9	70.7
300k	0.9	2.5	5.6	11.5	22.4	43.0	81.6
400k	0.9	2.6	5.9	12.2	24.1	46.9	90.4
500k	0.9	2.7	6.2	12.8	25.6	50.2	97.8
600k	0.9	2.8	6.4	13.3	26.8	53.1	104.3

Table 1: Number of connections per user ($d_{\mathcal{G}_\phi}$).

generally the network graph may be affected by the price. Therefore any causality arguments need to be handled with care. Our estimation will be mainly concerned with figuring out if prices and number of users behave in a way that is consistent with this framework. Giving a possible interpretation that has not been sufficiently explored in the literature. Appendix A.2 presents a vector error correction model estimation. This allows a preliminary exploration of how past user numbers (and prices) may affect present prices (and user numbers). Results support this network framework, but no definitive claims can be made on causality at this point. Finally, we take the network as given, this paper remains agnostic on the mechanisms and triggers for network formation. We do not aim to answer why users join the network or why the network is used by hundreds of thousands as supposed to only a handful.

4 Ethereum Network

One of many cryptocurrency blockchains is Ethereum, released in 2015. It fits the standard definition of a transactions ledger, that uses cryptography, distributed across many computers throughout the globe which add new transactions as well as verify historical data. A reward is provided for these computers for the computing power they provide via a process called mining. The unit of account, the cryptocurrency, is called Ether (ETH). It is transacted within the network, but also in external exchange platforms for other currencies.

The key aspect where the Ethereum network of users and transactions tries to differentiate itself from other blockchains is its so called "Virtual Machine". Within its blockchain, and accessible by computers across the globe, it is possible to store code that can be run by these computers and therefore acts as a remote computer. A user can activate or run this code, by sending a transaction to where this program is stored, a specific address, and reward the network for running it. In this sense, the Ethereum network can be thought of as having a pay per use computer within its blockchain network.

The Ethereum Virtual Machine, allows users to build and store their own cryptocurrency, within the network. Effectively becoming their own central bank for their own specific currency. This feature allows firms to create their own coin, tied to ownership, and sell it (in exchange for other currencies) in order to raise funds. Essentially like an Initial Public Offering (IPO), but on a specific exchange platform, virtually unregulated. These so called Initial Coin Offerings (ICOs) have raised over 20 billion dollars and captured a significant amount of attention by the media and regulators.

Ethereum has been for most of its history, the second most valuable cryptocurrency, following market capitalisation. Only surpassed by Bitcoin, the initial implementation of a blockchain and cryptocurrency. The importance of Ethereum in technical achievements, interest by developers, and its high valuation make it an interesting subject of study. Its relative longevity and data availability further the appeal as the research target of this paper.

4.1 Data

Empirically, data is obtained on both the number of active addresses (A_t) in the Ethereum blockchain and price (P_t) of Ether. These will be the empirical counterparts of the total number of users (n_t) and price ($Price_t$) respectively. These two time series cover the period from the 7th of August 2015 when this cryptocurrency started being listed and ends on the 1st of September 2018. Hence, $t = 1, ..., T$, with $T = 1121$. Daily active addresses are obtained using Bitinfocharts. This is the same source as Wheatley et al. (2018) cite for Bitcoin user activity. In order to be counted as active, an address has to either be a sender or a receiver in a transaction performed in the last 24 hours. It is important to note that there is one underlying assumption, that each address represents one unique user.[4] Price data is sourced form Etherscan, a block explorer and analysis platform that provides information on the network. It represents the trading price of one unit of Ether in several online exchanges.

Both series are made up of daily observations, with no interruptions except for 5 days on the Price (08/12/2017-12/12/2017) and 3 days on the Active Users (28/10/2017-30/10/2018). For both series, missing values were substituted with the previous available value. One further adjustment was necessary when taking the log of the price variable. In order to avoid taking logs of 0, the observation of the 10/08/2015 (4 days into the series) is modified to be 0.01. Dropping these conflicting observations does not affect the results presented.

4.2 Non-Stationarity

Figure I plots both time series of interest in levels and logs. As we can see both variables followed a very similar path which can be split into three rough stages. An initial period of slow growth until spring of 2017, followed by rapid growth until January 2018. Finally a downward trend.

Simple inspection of the evolution of these time series casts doubt over a key condition for most standard time series analysis, stationarity. More precisely, covariance stationarity is a requirement for most estimation techniques. This is the restriction that the mean and variance doesn't depend on time, and therefore is constant across the sample.

Formally we can test whether a series is stationary or not with a wide set of tests. Practically these check whether the process has a unit root. Recomended in Hamilton (1994) is the augmented Dickey-Fuller (ADF) test, which in this case fails to reject the null of non-stationarity for either series. The KPSS test by Kwiatkowski et al. (1992), points in the same direction, rejecting the null that the series are stationary. When logs are taken, both tests give the same results. Taking first differences of both series, in levels and in logs, reject the null of non-stationarity and fail to reject stationarity. This implies, following Granger and Newbold (1974), that the series we are dealing with are integrated of order one, $I(1)$. Together with non-stationarity this

[4]Like with emails, or bank accounts, any given user may own multiple addresses. This means that our estimations throughout are to be seen as an upper bound. Some work has been done to create decision rules to take account for this, such as Meiklejohn et al. (2013). Their implementation will be considered in future research.

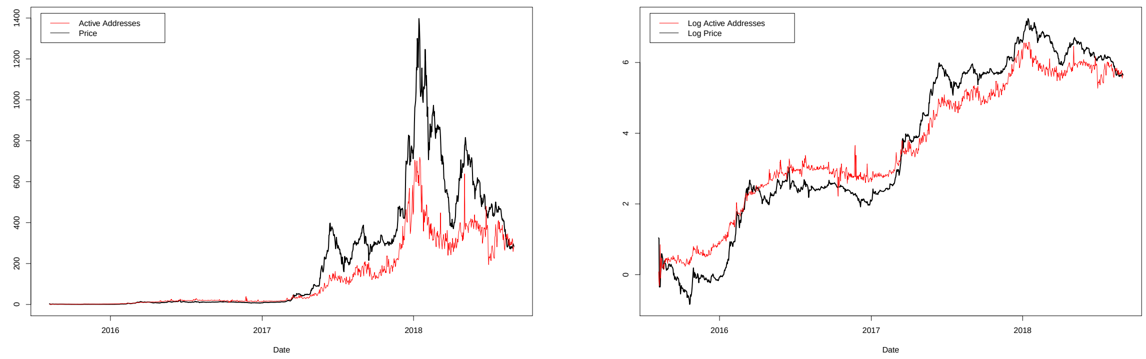

Figure I: Prices (P_t) in black and active addresses (A_t) in red. Levels on the left, logs on the right.

requires relevant econometric techniques which will affect the interpretation of results.

5 Parameter Estimation

5.1 Long-Run Relationship.

Following the textbook approach of Hamilton (1994), when non-stationary time series are regressed, the results may be spurious even if the estimation results appear significant. In order to rule out this spurious regression problem, a test for non-stationarity on the errors resulting from such regression must be performed. If an ADF test rejects the null hypothesis of non-stationarity we can be confident that the estimated relationship isn't spurious.

From the definition in Engle and Granger (1987) we can say that two series are cointegrated, if there exists a cointegrating vector that generates a relationship whose errors are of a smaller order of integration. In our case that would mean stationarity.

Our equation of interest for estimation emerges from (3), dropping the negligable βn_t term as A_t is large. We use the data discussed in Section 4. Following the approach of Wheatley et al. (2018) we log-linearise:

$$log(P_t) = log(\beta) + \phi * log(A_t) + \epsilon_t. \tag{6}$$

In order to rule out spurious results, it is required that the resulting linear combination, $\hat{\epsilon}_t$, is a stationary time series. If this is the case, there exists a cointegrating relationship.

This cointegrating relationship can be expected to hold in the long-run. We can then rely on Granger's Representation Theorem from Engle and Granger (1987), to build an error correction model that can also speak to the dynamics in the short-run. This model can also give information on the speed of convergence to the long-run relationship is. Reverse causality cannot be ruled out in cointegrating relationships, this is a clear limitation of this approach. Appendix A.2 uses a vector error correction model to assess this possibility.

Furthermore, although the least squares estimator of the cointegrating relationship is superconsistent, meaning it converges at rate T, inference isn't straightforward. The emerging standard errors, are misleadingly small, and therefore not reported here. Section 6 is devoted to obtaining inference using a variety of bootstrapping methods.

Performing least squares estimation, the following results are obtained:

$$log(P_t) = -1.36 + 1.2638 * log(A_t) + \hat{\epsilon}_t. \tag{7}$$

We perform the suggested tests for stationarity on the estimation errors $\hat{\epsilon}_t$. Both the ADF (p-value < 0.01) and KPSS (p-value > 0.1) tests support the presence of a cointegrating relationship. Confident that there is a cointegrating relationship, that curbs the threat of spurious results, attention can shift to parameter interpretation. At this point, we will rely on the superconsistency of the estimator to assume that indeed the parameters are correctly estimated and take them at face value. Understanding however that interpretation will be qualified once inference is achieved.

The results support a relationship between the number of users and the prices that is consistent with a network framework, as before, with saturation as $1 < \hat{\phi} < 2$. We can interpret these results in terms of connectedness between users. The degree of each user ($d_{\mathcal{G}_\phi}$) is easy to obtain with $\phi = 1.2638$, for several network sizes. This parameter implies that even for small network sizes, accounts are connected to at least more than a handful of others. This number grows, at a decreasing rate, when network size grows to much larger dimensions. Table 2 shows in more detail how the degree per user changes at different network sizes.

	Users (in thousands)										
	50	100	200	300	400	500	600	700	800	900	1,000
$d_{\mathcal{G}_{1.2638}}$	16.4	19.8	24	26.9	29	30.9	32.4	33.8	35.1	36.2	37.3

Table 2: Number of connections per user ($d_{\mathcal{G}_{1.2638}}$).

5.2 Error Correction Model

We have established that the two variables of interest are cointegrated. Hence, we can rely on Granger's Representation Theorem, which shows that for a set of cointegrated variables there exists a valid error correction representation.

In this two variable system, the error correction model (ECM) relates the change in one variable to past equilibrium errors, as well as to past changes in both variables.[5] That is, changes in prices today will depend on how far away from the long-run cointegration relationship prices are, as well as on some short-run dynamics of prices and activity. A nuisance parameter that is to be determined is the number of lags to be included. For simplicity, the model is presented with one lag and estimated as well with two. Robustness checks with a more fliexible lag structure are performed later, but does not drastically affect the results.

$$\Delta log(P_t) = \mu_0 + \alpha\epsilon_{t-1} + \mu_{p,1}\Delta log(P_{t-1}) + \mu_{a,1}\Delta log(A_{t-1}) + w_t. \tag{8}$$

Note that the ϵ_{t-1} here is the same from equation (6) and we can use $\hat{\epsilon}_{t-1}$ for estimation purposes. This will provide a measure of how far away from the variables of interest are from their cointegrating relationship. The coefficient on this variable (α) will measure the speed of convergence, or how tightly bound the variables are to their long-run relationship. When ϵ_{t-1} is

[5] As reverse causality cannot be ruled out, it is possible to write the ECM with changes in activity level as the dependent variable. Combining these two equations will give us a vector ECM (VECM), which is estimated in Appendix A.2.

positive, prices are above the long-run relation. It is necessary for the cointegrating relationship to hold that α is negative, to drive prices down and bring ϵ back to zero. Large absolute values of α will deliver a faster return to the long-run relationship.

This ECM can be used to re-estimate the parameter of interest ϕ. Substituting in (6) for ϵ_{t-1} and dividing the resulting coefficient estimated on past activity level by the convergence speed.

$$\Delta log(P_t) = \mu_0 + \alpha(log(P_{t-1}) - log(\beta) - \phi log(A_{t-1})) + \mu_{p,1}\Delta log(P_{t-1}) + \mu_{a,1}\Delta log(A_{t-1}) + w_t. \quad (9)$$

Table 3 shows the estimation results from the ECM above. Columns (I) and (Ia) with one lag of changes, (II) and (IIa) with two. Columns (Ia) and (IIa) show the estimates using the long-run relation, like in (9). This provides a new set of estimations of the parameter of interest. Standard errors are reported here in parenthesis. They are relevant for Columns (I) and (II), which involve only $I(0)$ variables. Furthermore, (Ia) and (IIa) have a cointegrating relationship.

	(I)	(Ia)	(II)	(IIa)
Intercept	0.0057	-0.146***	0.0055	-0.129***
	(0.0055)	(0.025)	(0.0056)	(0.026)
$log(P_{t-1})$		-0.11***		-0.099***
		(0.0156)		(0.016)
$log(A_{t-1})$		0.141***		0.125***
		(0.020)		(0.020)
ϵ_{t-1}	-0.11***		-0.099***	
	(0.015)		(0.016)	
Lags	1	1	2	2

Standard errors in parentheses (*** $p < 0.001$). $R^2 \approx 23\%$.

Table 3: ECM estimation results.

Focusing on columns (I) and (II), we obtain a small and insignificant intercept, which is consistent with price returns that are not deterministic. Furthermore, $\hat{\alpha}$ is negative and highly significant. Although not instant, this convergence speed shows that the long-run relationship can have very meaningful effects on short-run dynamics. For example, if prices yesterday were 5% above what is consistent with their long-run relationship with the number of users, then $\epsilon_{t-1} \approx 0.05$. The expected effect of this on returns today is of a 0.5% drop, a meaningful amount. About 10% of the distance between prices to their long-run relationship is expected to be corrected the following day.

In columns (Ia) and (IIa) the cointegration relationship has been substituted in, like in (9). In these cases, we do not have a convergence speed, as it is tangled up with the parameters of interest. What we can do is use the previously estimated convergence speed to get a new estimate of the parameter of interest, and vice-versa. Using the convergence speed in (I), the parameter of interest (ϕ) implicit in (Ia) is 1.2818. Similarly, using the convergence speed of (II), the parameter of interest emerging in (IIa) is 1.2626. These are very close to the estimated parameter in Section 5.1, supporting the robustness of our estimates. Using the parameter previously estimated, $\hat{\phi} = 1.2638$, we can do the reverse procedure and obtain new convergence speeds. The implicit convergence speed in (Ia) is 0.11 while for (IIa) it's 0.099, exactly the same as estimated in (I) and (II).

We have established that there is a long-run relationship relationship between prices and the number of users, which is consistent with each user being connected to about 30 others. This relationship is also relevant in the short-run price dynamics as when prices deviate from this, about 10% of the difference is corrected the following day. Using different estimation methods yields very similar results for our parameter of interest, ϕ, therefore providing some early indication that these results are robust. Having estimated and interpreted the parameters of interest, the rest of the paper will be concerned with inference within this non-stationary context.

6 Achieving Inference

Despite the superconsistency of the estimates of the cointegration relationship, the non-stationarity of the data does not allow for the use of standard errors as usual. Instead, inference can be achieved by using alternative methods focused on the variation in the errors resulting from the estimation, as suggested in Xiao (2012). Using different methods, this section will obtain a pseudo-distribution of the connectedness parameter. The three distinct approaches carried out, are all based on bootstrap techniques. They can be divided into two types, those applied to the cointegrating relationship and a distributed lags approach, in the spirit of an ECM. They support the hypothesis that the parameter of interest is contained within the interval consistent with positive network effects, $1 < \phi < 2$. Moreover, that this parameter is in the lower region of this range. This reinforces the interpretation that prices are consistent with each address being connected with more than a handful of others, a number in the tens rather than in the hundreds.

6.1 Bootstraps on Cointegrating Relation

One common method to obtain inference, when standard errors are not valid is to use bootstrap techniques. In this case however, we need the replicated data to still have the same non-stationarity properties as the original series. Paparoditis and Politis (2001) discuss these issues and introduce a new algorithm, the continuous block bootstrap (CBB). Mechanically, this algorithm fixes the initial level of the original series and takes first differences of the rest. Blocks of the differenced data are then combined, randomly, to generate a pseudo-series which is capable of replicating the behaviour of the original data. There are two related nuisance parameters to be selected in this method, the number of blocks and their length. We fix the replicated series to be the same length as the original so only block length is to be chosen. Selection of such parameter is however, crucial to the results that emerge from the CBB.

In order to be able to re-estimate the cointegrating relationship, the snipping and reshuffling of both time series is done simultaneously, at the same points, performing a paired bootstrap. Furthermore, in order to capture the long-run dynamics between these two series, we need the block size to be big enough. A small block will not replicate the interaction between the series and the underlying cointegration relationship. However, if the block is too big, there won't be enough reshuffling, and the different replications will be too similar to the original data. There is therefore a trade-off when selecting the block size.

We apply the CBB, with 1000 repetitions at various block sizes. The smallest block size being of one single observation, which would be the case if all observations were independent from each other. On the other hand, the largest block we report is of 150 observations, so that there are at least 10 blocks in each repetition.

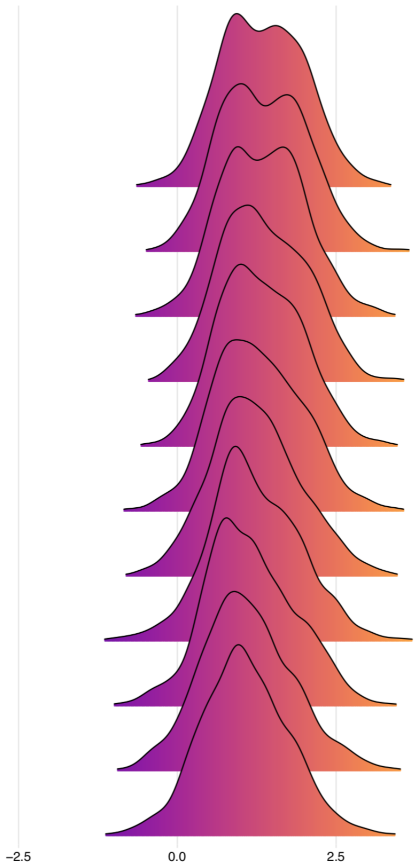

Figure II: Parameter distribution (ϕ) for different block sizes.
From 50, at the bottom, to 150 at the top in increments of 10.

Figure II plots the distribution that emerges for the parameter of interest, ϕ, at 10 different block sizes. The smallest block size, at the bottom, is of 50 observations, almost two months, and the largest is of 150, about 5 months. Results for block sizes below 50 are not reported. These block sizes are too small to capture any long-run relationship, the parameter is insignificant, with the mean below 1. For all block sizes reported, the mean is above one, and slowly increases with the block size until it stabilises around $\phi = 1.3$. Standard errors reduce with larger block sizes, becoming significant at the 10% significance level for block sizes above 70 observations. For block sizes above 100 it is significant at the 5% level. Despite becoming significant it is not possible to reject either the null of no network effects ($H_{01} : \phi = 1$) or of a full network ($H_{02} : \phi = 2$) with any meaningful significance.

More explored in the literature is the use of bootstrap techniques to test for unit roots. Using the Hansen and Racine (2018) test, it is possible to indirectly generate a pseudo-distribution of the cointegrating parameter. The Hansen-Racine test for unit root is a model averaging procedure, where no nuisance parameters are required (the block size is determined automatically). The null which is tested is that the time series variable of interest is non-stationary. In order to get a feel for the parameter distribution, we generate errors, $\tilde{z}_t$, and test for stationarity at different parameter levels, $\tilde{\phi} \in [0, 3]$. We fix the constant in the long-run relationship to the estimate in Section 5.

$$\tilde{z}_t = log(P_t) + 1.36 - \tilde{\phi} * log(A_t). \tag{10}$$

The true parameter of ϕ will generate a $\tilde{z}_t$ that will be able to reject the null of non-stationarity at much higher confidence levels than false parameters. Furthermore, the closer $\tilde{\phi}$ is to the true parameter the higher the conficence level of rejection will be.

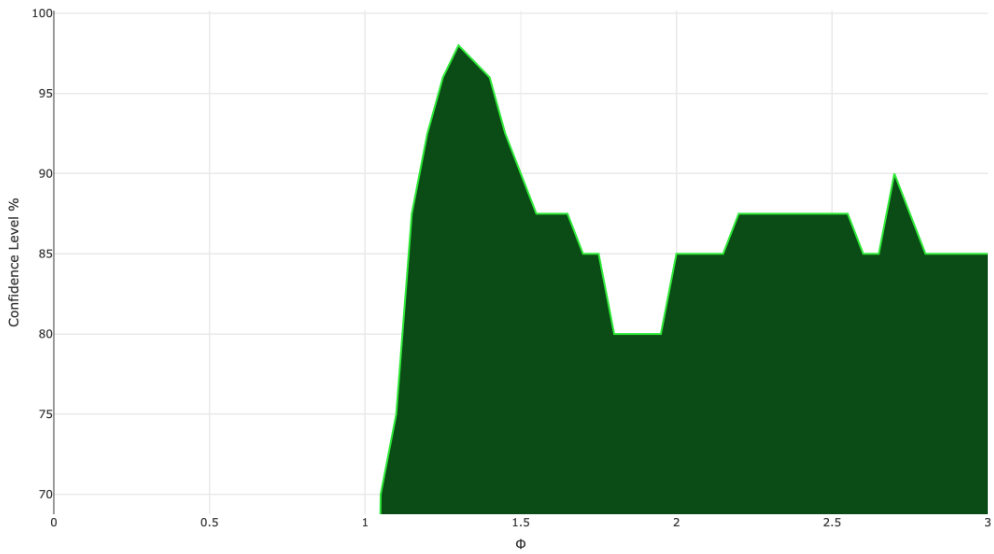

Figure III: Rejection confidence from Hansen-Racine on errors.

Figure III, shows the parameter pseudo-distribution that emerges from this exercise. We have a peak at $\tilde{\phi} = 1.2$, where it is possible to reject the non-stationarity of the errors generated with 99% confidence. With $\tilde{\phi} = 1$ it is not possible to reject the null of spurious relationship with any confidence, this means that with $\tilde{\phi} = 1$ the errors produced are non-stationary. This allows to reject the null of no network effects ($H_{01} : \phi = 1$). The parameter also fails to generate a stable cointegrating relationship as it approaches 2, weakening the hypothesis of a complete network ($H_{02} : \phi = 2$).

We observe a second peak, as the parameter goes beyond 2. The explanation for this is simple, multiples of the true cointegrating parameter will also deliver stationary errors. This second maximum point in the distribution occurs around $\tilde{\phi} = 2.4$, twice the parameter with the highest rejection confidence.

The Hansen-Racine procedure selects a block size of 95. This provides some further validation of the CBB results with similar block sizes, that found significance. It can be said that inference based on bootstraps on the cointegrating relationship place the point estimate at a similar level to previous estimations, around $\phi \approx 1.2 - 1.3$. Moreover using the Hansen-Racine test we can reject the null of no network effects and cast significant doubt on the validity of the complete network hypothesis.

6.2 ARDL

An alternative approach to obtain inference, but also to estimate the long-run relationship between the variables of interest, is to take an autoregressive distributed lags (ARDL) approach.

In an ARDL the variable of interest is explained by p lags of itself, the explanatory variables and q of its lags. An ARDL(p,q) can also include a constant and a time trend. The addition of a time trend isn't justified theoretically, and even when included does not vary the results here presented in any meaningful way. The specification of interest is therefore:

$$Log(P_t) = \gamma_0 + \sum_{j=1}^{p} \gamma_j Log(P_{t-j}) + \sum_{h=0}^{q} \psi_h Log(A_{t-h}) + v_t. \quad (11)$$

We can modify an ARDL(p,q), to have a modified-ARDL in first differences. The resulting modified-ARDL(p,q) can be written as:

$$\Delta Log(P_t) = \delta_0 + \delta_1 Log(P_{t-1}) + \delta_2 Log(A_{t-1}) + \sum_{j=1}^{p-1} \theta_j \Delta Log(P_{t-j}) + \sum_{h=0}^{q-1} \rho_h \Delta Log(A_{t-h}) + v_t. \quad (12)$$

In equilibrium, where $A_t = A_{t-1} = A^*$ and $P_t = P_{t-1} = P^*$:

$$0 = \delta_0 + \delta_1 Log(P^*) + \delta_2 Log(A^*). \quad (13)$$

A decision is necessary on how many lags are enough. It is hard to justify a decision on theoretical grounds. Cryptocurrency trading markets are open 24 hours, 7 days a week, 365 days a year. Therefore usual concerns of market timings and weekends in stock markets, where traders don't want to have positions open while the market is closed disappear. As a result, we base the lag selection purely on a statistical basis, using the Bayesian information criterion (BIC).

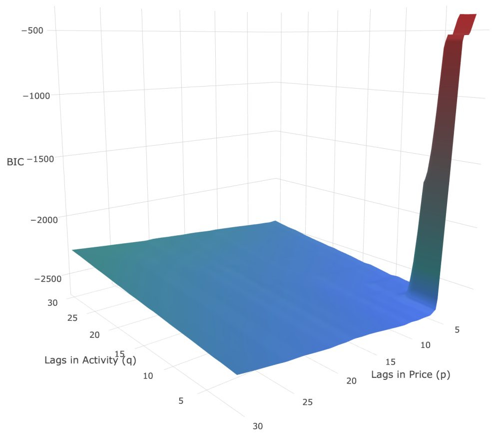

Figure IV: BIC Lag Selection.

From the ARDL(30,30) space, the resulting model using the BIC is an ARDL(6,1), one week of price data and two days of activity (t and $t-1$).[6] Empirically, the addition of exactly a week

[6]This ARDL method does not allow for holes in the lag structure, selection of later lags forces more recent lags into the model. As a robustness check, Appendix A.3 presents a similar model where the lag structure is selected using a Lasso. Results are robust to having a more flexible lag structure.

of price data in this model may hint that there is a theoretical mechanism where a week is a period that matters in cryptocurrency markets, we refrain from addressing this further in this paper. Furthermore, the reduced lagged effect of activity on prices, seems to hint towards the idea that information of activity changes filters through quickly to prices. Figure IV shows the BIC values for the different models in the grid. The minimum occurs on the bottom right of the surface.

We can estimate equation (12) using least squares and get valid standard errors. With these estimates, we can then get estimates for the parameters of interest in (13) using the delta method. These are reported in Table 4.

This allows us to get another point estimate, that is a little smaller than that found in the previous sections. But it is still very consistent with having small, but meaningful network effects. But the main advantage is to be able to build a confidence interval for these parameters. This 95% CI allows to reject the null of no network effects ($H_{01} : \phi = 1$) and the null of a complete network ($H_{02} : \phi = 2$) and thus "Metcalfe's Law".

	Point Estimate	2.5%	97.5%
$\hat{\phi}^{ARDL}$	1.218	1.052	1.384
$log(\hat{\beta})^{ARDL}$	-0.869	-1.707	-0.031

Table 4: ARDL estimation results.

Similar to previously, we can build a table of the implied degree with at the 95% CI, reported in Table 5. Despite being able to reject no network effect as well as a complete network, the disparity between the lower and the upper bound in terms of implications is vast. For the lower bound ($\phi = 1.052$), each address is estimated to have less than one connection for most of the historically observed network sizes. This is impossible, as all addresses must have at least one connection to be active and therefore included in the sample.

	Users (in thousands)									
ϕ	100	200	300	400	500	600	700	800	900	1,000
1.052 |	<1	<1	<1	<1	<1	<1	1	1	1	1.1
1.384 |	82.2	107.5	125.8	140.6	153.3	164.5	174.6	183.8	192.4	200.4

Table 5: Number of connections per user ($d_{\mathcal{G}_\phi}$).

On the other hand, the upper bound delivers a much larger degree than in the point estimate, with users being connected to about a hundred users, rising to a couple hundred when the network reaches a million users. Considering this relates to financial data, and the relatively low cost of creating an address, this doesn't seem too unrealistic. It is possible to imagine users connecting with numbers of sellers and buyers in the hundreds.

The delta method isn't the only way that can deliver inference from the ARDL. Using the errors that are generated from (12), which are stationary, a bootstrap approach can be carried out. A bootstrap is performed on the modified-ARDL(6,1) with 20,000 repetitions[7]. This allows

[7]Using the `boot()` function in `R`

to generate a pseudo-distribution for $\hat{\phi}^{ARDL}$, which is plotted in Figure V. Furthermore, we can get some descriptive statistics of this distribution, reported in Table 6.

Mean	Median	90% CI	95% CI
1.14	1.21	(0.886, 1.396)	(0.637, 1.488)

Table 6: Bootstrap distribution statistics for $\hat{\phi}^{ARDL}$.

Looking exclusively at confidence interval emerging from the distribution, we can reject the hypothesis of a fully connected network ($H_{02} : \phi = 2$). However, we cannot reject $H_{01} : \phi = 1$. Moreover, the mean and median values are slightly below prior estimates. The distribution of this parameter is given in Figure V. As most of the mass is located in $1 \leqslant \phi \leqslant 1.5$ there is further support to the view that network effects are present, but moderate, in line with the rest of our results.

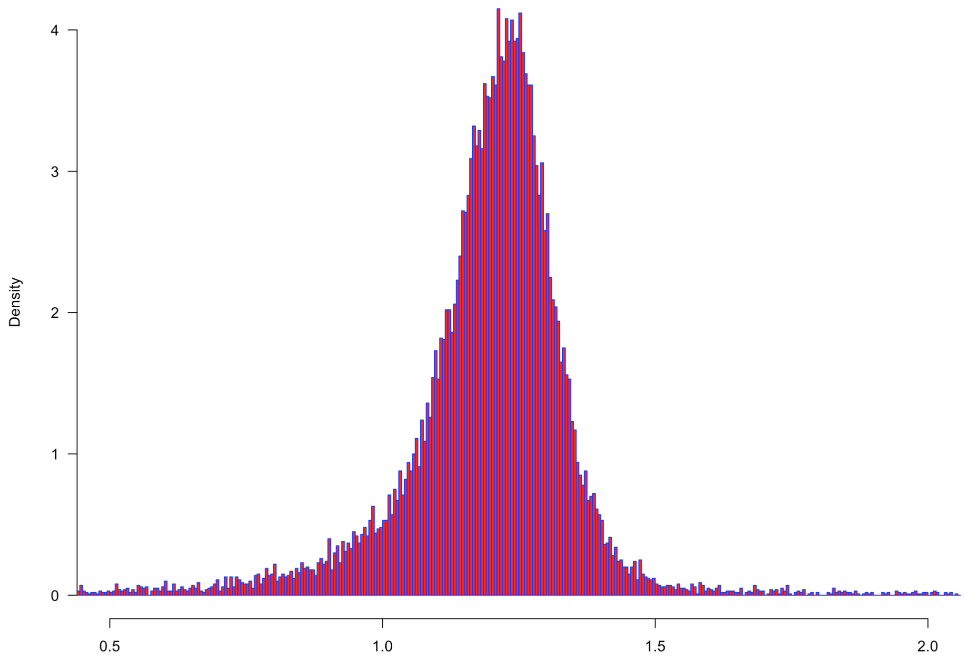

Figure V: Bootstrap distribution for $\hat{\phi}^{ARDL}$.

7 Conclusion

Within the economics literature a positive relationship between prices and the number of users of a cryptocurrency is usually underlying in the theoretical works. Empirically only basic relationships are tested, where econometric issues emerging from data non-stationarity are not addressed and parameter interpretation is given little attention. This papers presents a network framework, where restrictions on the relationship of interest are clearly stated. This allows to interpret the parameters in the empirical literature in this network context. More importantly, it opens the door to future research where not only the number of users is important to prices, but many more network characteristics such as centrality.

Empirically we estimate this relationship between users and prices in the Ethereum network, the second most valuable digital currency. In doing so we take into account the non-stationarity, establishing cointegration between the series, as well as estimating an ECM. From this we can say that prices are consistent with small network effects, where each user is connected to about 30 other users, at observed user numbers. This long-run relationship has strong implication for short-run dynamics. Using the speed of convergence from our ECM, any difference between observed prices and those implied by the long-run prices will be corrected by 10% the following day. Another contribution to the literature is the focus on inference, which is also affected by the non-stationarity of the data. Using a range of bootstrapping methods we are able to reject the null of a complete network, often assumed empirically. Although not as confidently, the null of no network effects is also rejected.

References

Alabi, K. (2017). Digital blockchain networks appear to be following metcalfe's law. *Electronic Commerce Research and Applications*, 24:23–29.

Bhambhwani, S., Delikouras, S., and Korniotis, G. M. (2019). Do fundamentals drive cryptocurrency prices? *Working Paper*, Available at SSRN: https://ssrn.com/abstract=3342842 or http://dx.doi.org/10.2139/ssrn.3342842.

Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., and Menkveld, A. J. (2018). Equilibrium bitcoin pricing. *Available at SSRN:*, http://dx.doi.org/10.2139/ssrn.3261063.

Bjerg, O. (2016). How is bitcoin money? *Theory, Culture & Society*, 33(1):53–72.

Blau, B. (2017). Price dynamics and speculative trading in bitcoin. *Research in International Business and Finance*, 41:493–499.

Bulligan, G., Marcellino, M., and Venditti, F. (2015). Forecasting economic activity with targeted predictors. *International Journal of Forecasting*, 31:188–206.

Cheah, E. and Fry, J. (2015). Speculative bubbles in bitcoin markets? an empirical investigation into the fundamental value of bitcoin. *Economics Letters*, 130:32–36.

Civitarese, J. (2018). Does metcalfe's law explain bitcoin prices? a time series analysis.

Corbet, S., Lucey, B., and Yarovaya, L. (2018). Datestamping the bitcoin and ethereum bubbles. *Finance Research Letters*, 26:81–88.

Engle, R. and Granger, C. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–276.

Granger, C. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2:111–120.

Hamilton, J. (1994). Time series analysis. *Princeton University Press.*

Hansen, B. and Racine, J. (2018). Bootstrap model averaging unit root inference. *Working Paper.*

Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159–178.

Li, J. and Chen, W. (2014). Forecasting macroeconomic time series: Lasso-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting*, 30:996–1015.

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G., and Savage, S. (2013). A fistful of bitcoins: Characterizing payments among men with no names. *Proceedings of the 2013 Conference on Internet Measurement Conference*, https://doi.org/10.1145/2504730.2504747:127–140.

Mitchnick, R. and Athey, S. (2018). A fundamental valuation framework for cryptoassets.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.

Pagnotta, E. (2018). Bitcoin as decentralized money: Prices, mining, and network security. *Working Paper*.

Pagnotta, E. and Buraschi, A. (2018). An equilibrium valuation of bitcoin and decentralized network assets. *Working Paper*, Available at SSRN: https://ssrn.com/abstract=3142022 or http://dx.doi.org/10.2139/ssrn.3142022.

Paparoditis, E. and Politis, D. (2001). The continuous-path block-bootstrap. *Asymptotics in Statistics and Probability.Papers in Honor of George Rouss*, pages 305–320.

Shapiro, C. and Varian, H. (1999). Information rules. *Harvard Business School Press*.

Tasca, P., Hayes, A., and Liu, S. (2018). The evolution of the bitcoin economy extracting and analyzing the network of payment relationships. *The Journal of Risk Finance*, 19(2):94–126.

Van Vliet, B. (2018). An alternative model of metcalfe's law for valuing bitcoin. *Economics Letters*, 165:70–72.

Wheatley, S., Sornette, D., Huber, T., Reppen, M., and Gantner, R. (2018). Are bitcoin bubbles predictable? combining a generalized metcalfe's law and the lppls model. *Working Paper*.

Xiao, Z. (2012). Robust inference in nonstationary time series models. *Journal of Econometrics*, 169:211–223.

Zhang, X., Liu, J., and Xu, Z. (2015). Tencent and facebook data validate metcalfe's law. *Journal of Computer Science and Technology*, 30(2):246–251.

A Appendix

A.1 Network Framework Summary

Here we summarise the network framework, stating clearly the assumptions we have made.

Assumption 1: Valuation and price can be used interchangeably. This emerges as the supply function for Ethereum and Bitcoin is known and for the latter reaches a maximum level. Therefore changes can be predicted.

Assumption 2: We can define a graph $\mathcal{G} := \mathcal{G}(V, E)$ where $m := |E|$ edges (transactions) are placed between $n := |V|$ vertices (users). A, the adjacency matrix of $\mathcal{G}$ has entries $a_{i,j} = 1$ if a transaction is made between i, j with $i \neq j$, 0 otherwise.

Assumption 3: The valuation of a network is given by the sum of the utility of its users (transferable utility):

$$Price = \sum_{i=1}^{n} \mathcal{U}_i. \tag{A3}$$

Assumption 4: The utility of a user in the network at time t can be explained by the network graph, $\mathcal{G}_t$.

$$Price_t = \sum_{i=1}^{n_t} \mathcal{U}_i(\mathcal{G}_t). \tag{A4}$$

Assumption 5: There is user homogeneity and stable utility functions.

Assumption 6: User utility is a weighted sum of direct links:

$$Price_t = \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} \beta a_{i,j,t}. \tag{A6}$$

Assumption 7: The total number of links can be captured by a density parameter ϕ.

$$Price_t = \beta(n_t^{\phi} - n_t) = \beta n_t^{\phi} - \beta n_t. \tag{A7}$$

Assumption 8: The network is large enough that the βn_t term is negligible.

$$Price_t = \beta n_t^{\phi}. \tag{A8}$$

A.2 Vector Error Correction Model

In order to look a little deeper at the relationship between these variables and the directions of influence, we look at a vector ECM (VECM). This is a model where in addition to the ECM presented in Section 5.2, equation (8), an analogous relationship is added, where the activity level is the dependent variable.

$$\Delta log(P_t) = \mu_{p,0} + \alpha_p \epsilon_{t-1} + \mu_{p,p} \Delta log(P_{t-1}) + \mu_{a,p} \Delta log(A_{t-1}) + w_{p,t}. \tag{A.14}$$

$$\Delta log(A_t) = \mu_{a,0} + \alpha_a \epsilon_{t-1} + \mu_{p,a} \Delta log(P_{t-1}) + \mu_{a,a} \Delta log(A_{t-1}) + w_{a,t}. \tag{A.15}$$

The errors from the cointegration relationship (ϵ_t) are the same for each equation as they represent how far away both time series are from the long-run relationship. Estimating both equations delivers the following (bold estimates are significant at the 99.99% level):

$$\Delta log(P_t) = 0.006 - \mathbf{0.111}\epsilon_{t-1} - \mathbf{0.495}\Delta log(P_{t-1}) + \mathbf{0.251}\Delta log(A_{t-1}). \tag{A.16}$$

$$\Delta log(A_t) = 0.007 - \mathbf{0.111}\epsilon_{t-1} - \mathbf{0.138}\Delta log(P_{t-1}) - \mathbf{0.272}\Delta log(A_{t-1}). \tag{A.17}$$

In both cases the constants, as expected, are insignificant and the convergence parameter is identical, as estimated in the main body of the paper. The differences emerge in the short-run dynamics.

Positive changes in activity levels have a negative effect on present activity level changes, a rebound effect. This is similar to the effect of past returns on present returns, although in prices this effect is more pronounced. As expected, and investigated in the main text, past activity levels have a positive effect on present returns. Which can be in line with the network framework developed and the idea that larger user numbers lead to higher prices and valuations. Unexpectedly however, positive past returns lead to reduced activity. Although not a very large effect, this strengthens the idea that the long-run relationship that we observe is indeed driven by the network framework story and not by an effect of prices on activity. If anything the effect of users on prices is underestimated as there is an effect in the opposing direction. One possible explanation for this phenomenon, by which past returns lead to less activity, is if users decide to hold off their activity when they see price increases. This would be consistent if they hold cryptocurrency as an investment and hope for these increases to continue. Further research into the mechanisms and causality within prices and user numbers is required.

A.3 Lasso Selected Lag Structure

Economic theory doesn't provide a clear basis for the determination of the lag structure in the ARDL model of Section 6.2. This allows the use of different lag selection techniques, such as BIC. The use of a Lasso approach provides completeness, allowing holes in the lag selection. This means, selection of the n^{th} lag does not impose selection of the $n-1$ previous lags, which is the case in the context of the BIC. The approach here followed is in the same vein as in papers such as Li and Chen (2014) and Bulligan et al. (2015). They use the Lasso selection criteria, in the context of time series, applied primarily for a forecasting purpose.

Starting from a modified-ARDL(10,10), the Lasso procedure is applied to select lags. The optimal weight (λ^*) by cross validation, with 10 folds. The use of this technique to select the weight, is valid as the model is stationary, and leaving out a certain fold does not affect estimation. Figure A1 shows the coefficient evolution at different penalties (λ).

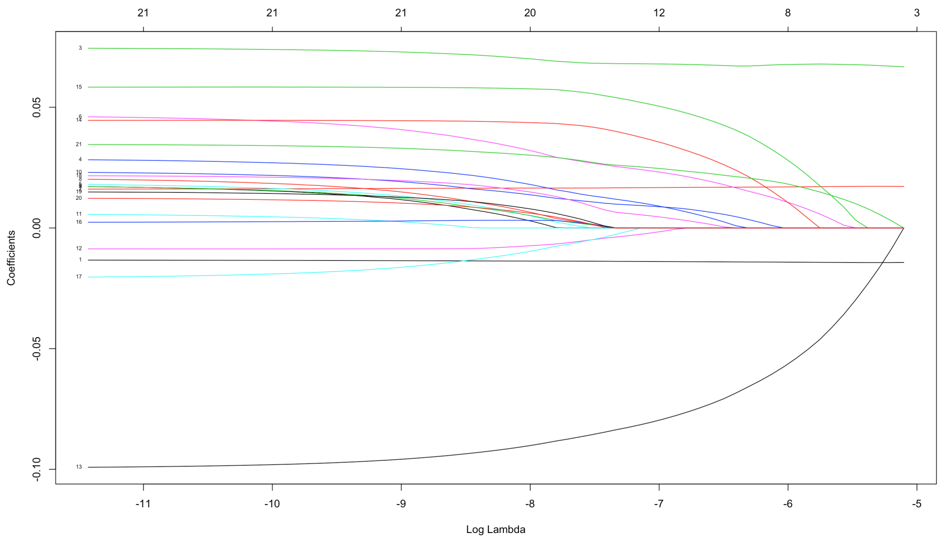

Figure A1: $log(\lambda^*) = -6.125$.

The emerging model has some holes, resulting in the following lag structure:

$$\Delta Log(P_t) = \delta_0 + \delta_1 Log(P_{t-1}) + \delta_2 Log(A_{t-1}) + \sum_{j=1}^{3} \theta_j \Delta Log(P_{t-j}) + \theta_9 \Delta Log(P_{t-9}) \\ + \rho_0 \Delta Log(A_t) + \rho_3 \Delta Log(A_{t-3}) + \rho_7 \Delta Log(A_{t-7}) + v_t. \tag{A.18}$$

Similar to the case where the lag structure was selected via BIC, we can look at the equilibrium case and build a estimate for the parameter of interest and use the delta method to obtain confidence intervals. These are very similar to the ones in Table 4.

Furthermore, we can run the bootstrap procedure also on this Lasso modified-ARDL and obtain a distribution for $\hat{\phi}^{Lasso}$. The results that emerge are in line with the main inference findings and consistent with the theory that the price reflects some small but significant network effects. We are able to reject both hypotheses of interest, $H_{01} : \phi = 1$ and $H_{02} : \phi = 2$.

	Point Estimate	2.5%	97.5%
$\hat{\phi}^{Lasso}$	1.203	1.025	1.382
$log(\hat{\beta})^{Lasso}$	-0.838	-1.731	0.055

Table A1: Estimation results from the Lasso selected modified-ARDL.

Mean	Mode	Median	90% CI	95% CI
1.22	1.25	1.21	(0.922, 1.384)	(0.749, 1.453)

Table A2: Bootstrap distribution statistics for $\hat{\phi}^{Lasso}$.

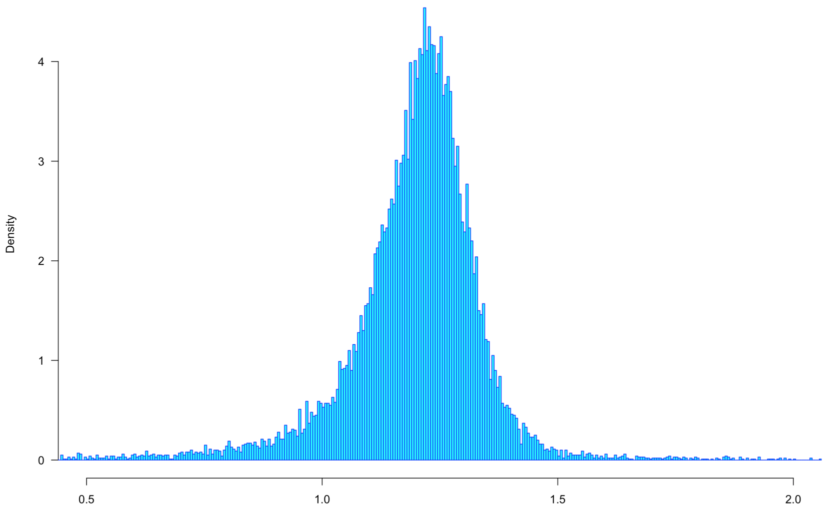

Figure A2: Bootstrap distribution for $\hat{\phi}^{Lasso}$.

A.4 From Number of Users to Number of Transactions

We can evaluate the validity of the assumption made on density structure, by using a different set of data. It is possible to use the total number of transactions, measured in thousands (X_t). This data covers the same sample as before. As can be seen by Figure A3, the total number of transactions follows a very similar path to the total number of active addresses.

$$Price_t = \beta \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j,t} = \beta X_t. \tag{A.19}$$

$$X_t = n_t^{\phi} - n_t. \tag{A.20}$$

In estimation, error stationarity will still be key. X_t is not exactly what is desired as it will count multiple times multiple transactions between i and j. However, it can give us a new estimate for β using (A.19). Using (A.20), it can also provide a new estimate of ϕ, that is not derived from the Prices.

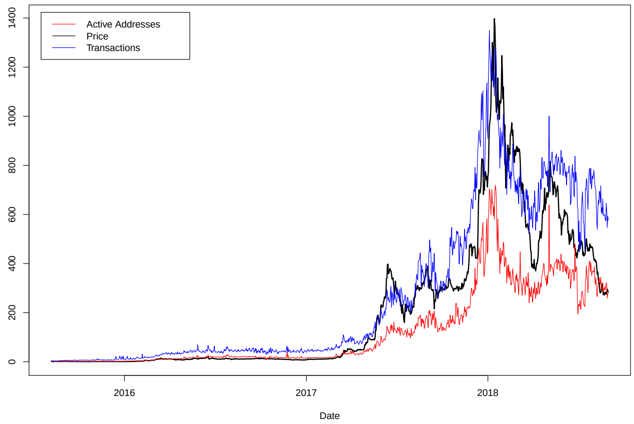

Figure A3: Price (P_t), Active Addresses (A_t) and Total Transactions (X_t).

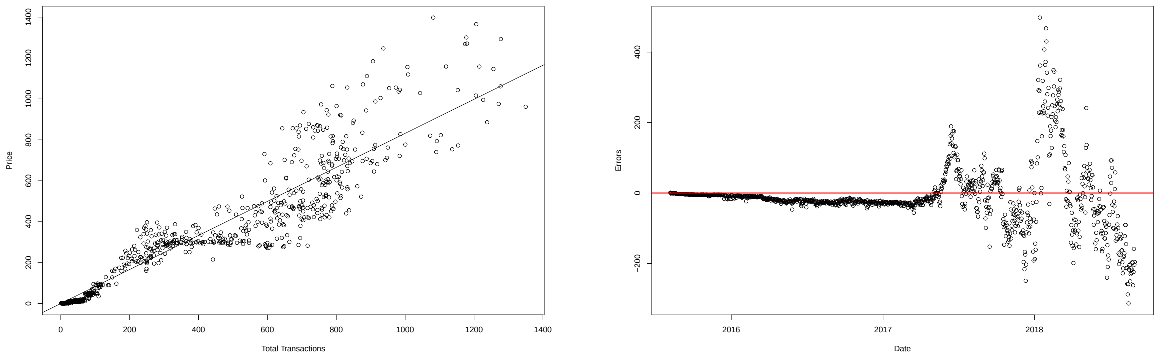

Figure A4: Relationship and errors between Price (P_t) and Total Transactions (X_t).

$$P_t = 0.832X_t + \hat{\epsilon}. \tag{A.21}$$

Equation (A.21) gives the estimation results for (A.19). As both variables are I(1), it is necessary to check the stability of the errors to confirm that this is indeed a cointegrating relationship. The ADF test yields a tentative rejection of non-stationarity (p-value = 0.049) while the KPSS test supports stationarity (p-value > 0.1). This is a hint that the variance may not be constant across the sample as well as heteroskedasticity may be present. This indeed appears when plotting the relationship and errors (Figure A4). We can compare this estimate of β with the one emerging in Section 5.1, $log(\hat{\beta}) = -1.36$. This corresponds to $\hat{\beta} = 0.257$. As such, using this measure leads to a much larger effect of transactions on the price.

Another possible exercise, that is perhaps more interesting is the use of (A.20) to obtain a new estimate of ϕ. In order to do so, we use a log-linearised version:

$$log(X_t + n_t) = 1.263log(n_t) + \hat{\epsilon}. \tag{A.22}$$

This result, would suggest that the parameter of interest ϕ implied by the observed prices is almost exactly the same as what is implied using the total number of transactions. However, these

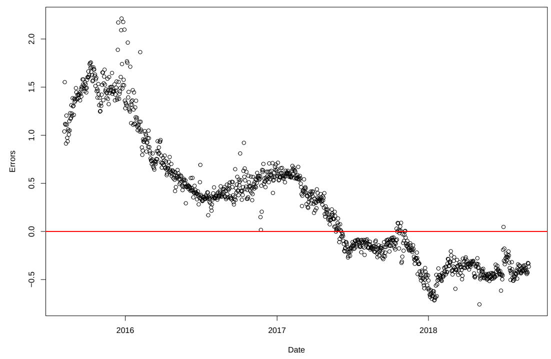

Figure A5: Errors from (A.22)

results may be spurious as the errors aren't stationary. An ADF test yields a p-value = 0.583, failing to reject non-stationarity, these are plotted in Figure A5. A stationary relationship is obtained when we allow for an intercept, errors show in Figure A6. However in this scenario the presence of a constant would change the parameter interpretation. With an intercept:

$$log(X_t + n_t) = 1.558 + 0.92log(n_t) + \hat{\epsilon}. \tag{A.23}$$

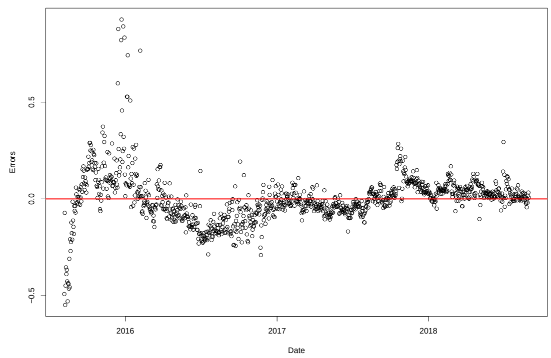

Figure A6: Errors from (A.23)

This will lead to a new interpretation, following (E3), whereby the average number of links users have decreases as the network size gets larger, tending towards one as the network grows to infinity. This casts doubt on the regularity imposition, as the results could be driven by an increasing number of users that perform only one transaction in larger network sizes. This would go against the interpretation we have given throughout, where users had more links when the network was larger. This case would be one of negative network effects as $\hat{\phi} = 0.92$.

$$Links/user = \frac{X_t}{n_t} = 4.75n_t^{-0.06} - 1. \tag{E3}$$