NETWORK SCIENCE

From code to market: Network of developers and correlated returns of cryptocurrencies

Lorenzo Lucchini[1,2], Laura Alessandretti[3,4], Bruno Lepri[1], Angela Gallo[5], Andrea Baronchelli[6,7,8]*

"Code is law" is the founding principle of cryptocurrencies. The security, transferability, availability, and other properties of crypto-assets are determined by the code through which they are created. If code is open source, as is customary for cryptocurrencies, this would prevent manipulations and grant transparency to users and traders. However, this approach considers cryptocurrencies as isolated entities, neglecting possible connections between them. Here, we show that 4% of developers contribute to the code of more than one cryptocurrency and that the market reflects these cross-asset dependencies. In particular, we reveal that the first coding event linking two cryptocurrencies through a common developer leads to the synchronization of their returns. Our results identify a clear link between the collaborative development of cryptocurrencies and their market behavior. More broadly, they reveal a so-far overlooked systemic dimension for the transparency of code-based ecosystems that will be of interest for researchers, investors, and regulators.

INTRODUCTION

A cryptocurrency is a digital asset designed to work as a medium of exchange. The underlying Blockchain technology allows transactions to be validated in a decentralized way, without the need for any intermediary (1). Every cryptocurrency is entirely defined and governed by its code, which determines its security, functionality, availability, transferability, and general malleability (2). This "code is law" architecture immediately puts developers under the spotlight (3). Lack of transparency in the coding process might damage users and other stakeholders of the code (4).

"Open code" is identified as the antidote to lack of transparency (3). Even if the code is accessible to only a small fraction of users, the reasoning goes, it would protect the asset and stakeholders from manipulations (5). For this reason, the code of the vast majority of cryptocurrencies is stored in public repositories. GitHub alone currently stores the code of more than 1600 cryptocurrencies (6).

Cryptocurrencies are nowadays used both as originally intended, i.e., media of exchange for daily payments and, to a larger extent, for speculation (7, 8). The market value of a cryptocurrency is not based on any tangible asset, resulting in an extremely volatile, and largely unregulated, market (9–11). However, the cryptocurrency market has attracted private and institutional investors (12, 13). At the moment of writing, more than 3000 cryptocurrencies are traded, capitalizing together more than 200 billion dollars (14, 15).

Here, we challenge the view that open code grants transparency to cryptocurrencies, even accepting that literate users do check it carefully (which is, of course, far from obvious). We do so by analyzing 298 cryptocurrencies (i) whose code is stored in GitHub and (ii) whose daily trading volume has been, on average, larger than

10^5 U.S. dollars (USD) (16) during their lifetime. We show the following:

1) A substantial fraction of developers (4%) contributes to the code of two or more cryptocurrencies. Hence, cryptocurrencies are not isolated entities but rather form a network of interconnected codes.

2) The temporal evolution of the network of co-coded cryptocurrencies anticipates market behavior. In particular, the first time two independent codes get connected via the activity of one shared developer marks, on average, a period of increased correlation between the returns of the corresponding cryptocurrencies.

Thus, the temporal dynamics of co-coding of cryptocurrencies provides insights on market behaviors that could not be deduced on the basis of the combined knowledge of the code of single currencies and the present state of the market itself. In other words, transparency, i.e., the availability of relevant market information to market participants, is a systemic property. The whole network of cryptocurrencies should be considered both by regulators and by professional investors aiming to maximize portfolio diversification. From this point of view, our work contributes a new dimension to the literature focused on the properties of the cryptocurrency market, which has, so far, adopted approaches ranging from financial (17–21) to behavioral (22) and from evolutionary (21, 23, 24) to technological (25, 26) perspectives.

RESULTS

GitHub activity and the network of cryptocurrencies

We are interested in the coding and market activity concerning actively traded cryptocurrencies (see Methods). The 298 cryptocurrencies with trading volume larger than 100,000 USD whose code is stored on GitHub (298 projects) include 63 of the top 100 cryptocurrencies, ranked by average market capitalization during October 2019. A total of 6341 developers contributed to these GitHub projects, totaling 879,742 edits (see section S1.2 for more details). The number of developers working on a cryptocurrency project correlates positively with its market capitalization (Spearman correlation coefficient, 0.48, with $P < 0.0001$; see fig. S2A), as previously noted (6).

[1]Fondazione Bruno Kessler, Trento 38123, Italy. [2]Department of Information Engineering and Computer Science, University of Trento, Trento 38123, Italy. [3]Copenhagen Center for Social Data Science, University of Copenhagen, Øster Farimagsgade 5, Copenhagen K 1353, Denmark. [4]Technical University of Denmark, Kgs. Lyngby DK-2800, Denmark. [5]Department of Finance, Cass Business School, City University of London, London EC1Y 8TZ, UK. [6]Department of Mathematics, City University of London, London EC1Y 8TZ, UK. [7]UCL Centre for Blockchain Technologies, University College London, London, UK. [8]The Alan Turing Institute, British Library, 96 Euston Road, London NW12DB, UK.
*Corresponding author. Email: andrea.baronchelli.1@city.ac.uk

The activity of the developers is heterogeneous. Twenty-eight percent of developers focused only on the top 10 cryptocurrencies, producing 20% of the edits, while only 15% of the developers worked only on projects with a capitalization lower than the median capitalization of the market, producing only 11% of the developing events. The Ethereum community soars above the others in terms of editing activity (109,527 development events), while Bitcoin has the largest number of developers, 832 (fig. S1). In general, the number of developers and the number of edits for a given project strongly correlate (Spearman correlation coefficient, 0.92, with $P <0.0001$; see fig. S2B).

We find that 4% of developers contributed to more than one cryptocurrency and are responsible for 10% of all edits. We further investigate their role by representing the GitHub data as a bipartite network, where developers and cryptocurrencies (the nodes) are connected by edit events (the links; Fig. 1A). We then project the bipartite network and obtain the network of connected cryptocurrencies where cryptocurrencies are nodes, and a link exists between them if they share at least one developer (Fig. 1B). We find that this network has 204 links, activated first by 147 different developers, and 123 nonisolated nodes, of which 115 form a giant component. Bitcoin has the largest number of connections, 53, followed by Ethereum with 43. The remaining 175 projects do not share any developer (Fig. 1C). The presence of a small fraction of developers who contributed to more than two cryptocurrencies (22 of 147) makes the network rich in cliques (see section S1.3 for more analyses on the network).

Market synchronization of GitHub-linked cryptocurrencies

We now consider the temporal evolution of the cryptocurrency network over 5 years of coding activity (from 5 March 2014 to 30 May 2019). A link between two cryptocurrencies is created the first time that a developer of one of the two edits the other (Fig. 2A), referred in the following as the GitHub connection time. What happens to the market behavior of the two cryptocurrencies that have just been linked in the GitHub network?

We focus on the correlation between asset returns (40, 41). We rescale time so that the connection time corresponds to $d = 0$ for each pair of GitHub-linked currencies, and we measure the Spearman correlation over a backward rolling window of size $s = 4$ months [see Figs. 2 (B and C) and Fig. 3A and Methods for definitions; results are robust with respect to variations of this definition; see section S1.4]. To limit the effect of overall changes in market evolution, we standardize the value of the Spearman correlation, for a given pair of linked currencies and at a given time, by subtracting the average correlation across all possible pairs of currencies at that time and dividing by the corresponding SD (see Methods).

Figure 3A shows that the average standardized Spearman correlation between the returns of two linked cryptocurrencies, averaged over the set of 204 linked pairs, increases at the turn of the GitHub connection time, rising from 0.31 ± 0.01, on average (±SEM), in the 4 months before the connection time, to 0.66 ± 0.01, in the period included between 2.5 and 6.5 months after the connection time [Fig. 3A, significant under Welch test (42, 43), with $P = 0.02$]. This corresponds to a relative increase of almost 130% after the synchronization occurred (see section S1.9.2 for details about the synchronization period). This result is robust to major perturbations of the network, including the removal of Bitcoin or/and Ethereum from it (fig. S9).

We test that the observed behavior is specific to linked pairs by measuring the synchronization of a random sample of 10^4 cryptocurrency pairs, selected from the entire market excluding linked pairs. Their connection time is chosen at random from the list of actual GitHub connection times (see section S1.4.1 for different randomization approaches). We find that the standardized correlation of these pairs remains constant across the connection time, ruling out the possibility of ecology effects induced by the specific distribution of connection times (fig. S5). We note also that, on average, the standardized Spearman correlation is higher for linked pairs compared to random pairs.

The increase in correlation observed for linked pairs could (i) be driven by few outliers or (ii) reflect the behavior of the majority of them. Figure 3B shows the distributions of the increase in standardized correlation between the 4 months preceding and the 4 months included between 2.5 and 6.5 months after the connection time. The distribution of linked pairs is centered at positive values of change (i.e., increase in correlation) and shows a significantly higher average synchronization compared to the distribution of random pairs, e.g., under Welch test (for more statistical tests, see section S1.4.1). In particular, approximately 65% of linked couples increased their correlation after GitHub connection time, a percentage significantly higher than random (fig. S13). These observations confirm that the observed change in correlation is not simply driven by outliers, hence supporting hypothesis (ii).

The market behavior of cryptocurrencies is also characterized by other properties. We repeated the analyses reported above to study the correlations between the time series describing daily changes in trading volume and market capitalization. We found no significant effects of the connection time on those measures (see results in section S1.8) under a Welch test at a significance level of 0.05.

Market properties of GitHub-linked cryptocurrencies

We now consider the market properties of GitHub-linked cryptocurrencies across GitHub connection time. First, we focus on the difference in market capitalization and volume among pair constituents. We find that the absolute difference in market capitalization and volume between two linked cryptocurrencies is typically larger than that between randomly selected cryptocurrencies [see Fig. 4 (A and B) and section S1.9.2 for details; note also that the market capitalization and volume of currencies are highly correlated, as expected (fig. S17)].

Then, we shift our attention to differences in market age, defined as the difference in the amount of time since a currency appeared in the market. We find that the age difference of the two cryptocurrencies in a linked pair, measured at connection time, is significantly higher, on average, than the difference of market age observed for random pairs (Fig. 4C). In particular, we find that the second-edited currency is younger than the first-edited currency in 61% of the cases and has lower market capitalization in 65% of the cases.

Last, we investigate the factors responsible for the observed heterogeneity in synchronization across linked pairs (Fig. 3B). We find that, when a linked pair includes one of the top 10 linked cryptocurrencies in terms of market capitalization (evaluated in the period preceding connection time), the corresponding synchronization of returns following connection is significantly higher than average (fig. S21D). Other factors, including the type of development event (push or pull), the direction of the link (from younger to older or vice-versa), and the connection time, do not explain the observed differences in synchronization across pairs (fig. S21, A, B, E, and F).
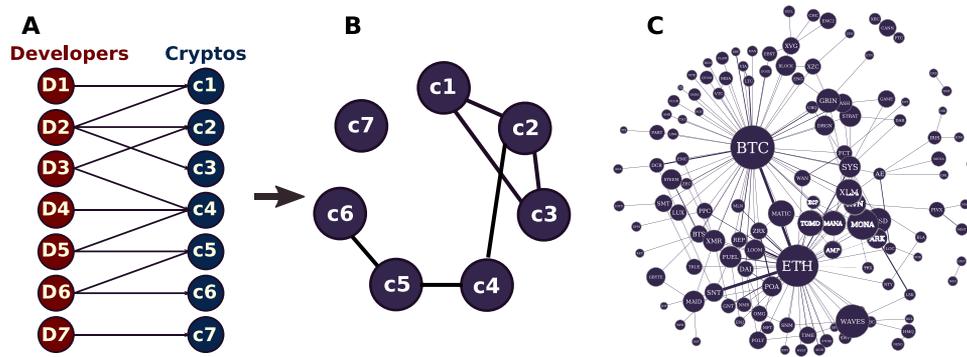
Fig. 1. The GitHub network of cryptocurrencies. (**A**) The GitHub dataset can be represented as a bipartite network, where developers (red circles) are linked to the cryptocurrencies (blue circles) that they have edited at least once. (**B**) Projection of the bipartite network; cryptocurrencies that have at least one common developer are connected. (**C**) The real network of 123 cryptocurrencies with at least one connection. Node size is proportional to the number of connections, and link width is proportional to the number of common developers between two cryptocurrencies. Bitcoin (BTC) and Ethereum (ETH) play a central role in the graph.
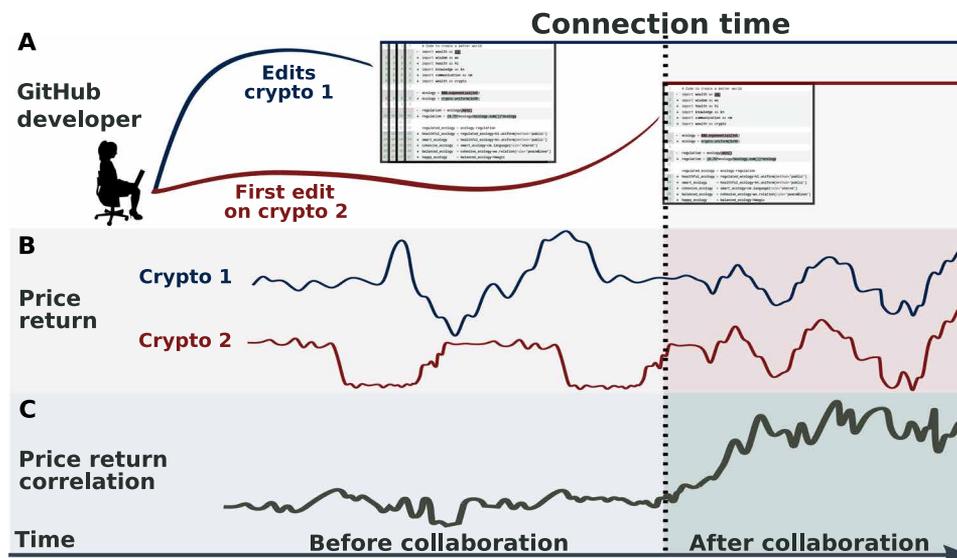
Fig. 2. GitHub co-development and cryptocurrency market synchronization. (**A**) A developer of cryptocurrency "crypto 1" publishes her/his first contribution to "crypto 2." If no other developer has worked on both currencies before, then this moment represents the GitHub connection time for the pair composed of "crypto 1" and "crypto 2." (**B**) The time series describing the asset returns of the two currencies synchronize after the connection time. (**C**) The Spearman correlation between the two time series increases when the asset returns synchronize.

DISCUSSION

We analyzed the relationship between code and market for 298 GitHub-hosted cryptocurrencies whose trading volume was larger than 10^5 USD for the covered period. We showed that approximately 4% of developers contributed to the code of more than one cryptocurrency and that these developers are more active than the average, contributing together to 10% of all edits. We then defined the network of co-developed cryptocurrencies and showed that, for months after the GitHub connection time, the correlation between the return time series of two GitHub-linked cryptocurrencies increased, on average. We found that other market indicators, and in particular, volume, do not show the same behavior. Last, we showed that developers tend to work on an established currency first and that linked pairs containing at least one top cryptocurrency exhibited a larger correlation of returns following connection.

It is important to delimit the scope of our findings. First, we only considered projects developed on GitHub. While this is, by far, the largest repository of cryptocurrency open-source code (it hosts more than 99% of the project hosted on online repositories), alternatives exist, e.g., GitLab (*44*). Second, we selected cryptocurrencies on the basis of their average trading volume, possibly neglecting currencies with only a short history of significant trading volume. Third, we focused on the first connecting event and did not investigate the presence and consequence of a possibly increasing pool of shared developers between two cryptocurrencies and/or actions of the developer(s) in that pool. Fourth, we considered pairs of cryptocurrencies, neglecting other possible influences of the network built in the first part of the article. Last, we did not consider the structure of the code or the semantics of the coding that a developer of the first cryptocurrency performs on the second. All these are open directions for future work.

Of course, our analysis cannot identify the mechanisms that drive the observed market synchronization. Speculatively, at least
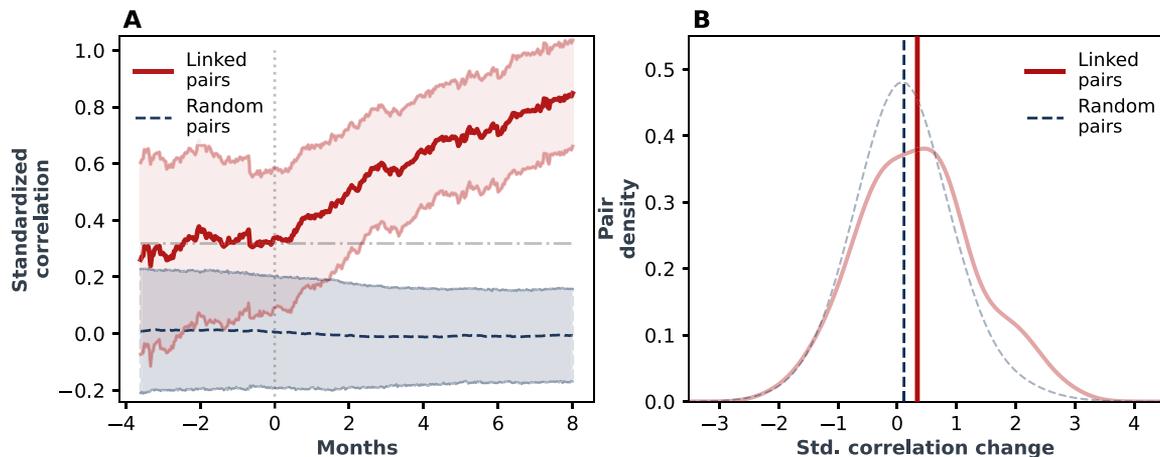
Fig. 3. Market synchronization following GitHub connection time. (**A**) Average standardized Spearman coefficients between return time series of linked pairs (red line) and a sample of random pairs of cryptocurrencies (dashed blue line). The size of the random samples is chosen to be the same as the number of existing linked pairs at each time. Its average size in the period reported in (A) is 124. Shaded areas represent 2 SDs of the mean and are determined via bootstrap (see Methods). The gray dot-dashed line corresponds to the average standardized correlation in the 3 months before the connection occurred. Time is shifted such that $d = 0$ corresponds to the GitHub connection time of each pair. Correlations are measured over a 4-month rolling window. (**B**) Distributions of the average correlation for linked and random pairs. Averages are computed over periods of 4 months: the 4 months before the connection time and the period between 2.5 and 6.5 months after the connection time. Vertical lines correspond to the average of each distribution. Pairs that synchronized after the connection time shift the distribution toward positive values. All the density distributions are computed using a Gaussian Kernel Density Estimation setting the bandwidth values to 0.39. For raw data histograms, see fig. S12.

two main dynamics might be at play. The first identifies code as an important "fundamental" for this market (*45, 46*). Traders would be aware of and operate (also) based on code and code development. The activity of developers would therefore represent a signal that, perceived by many traders, could result in the observed synchronization. The second dynamics, either complementary or alternative to the previous one, points to a greater role for developers, who could either directly own and trade large amounts of the cryptocurrencies that they edit or be hired by stakeholders who, in their turn, do the trade. At the systemic level, these interlocking directorates of developers/stakeholders would cast a shadow on the transparency of the market and potentially expose it to systemic risks due to hidden structural correlations between cryptocurrency prices.

In this respect, it is worth noting that the lack of incentives for developers is a longstanding issue for cryptocurrencies. Some Bitcoin developers, for example, are paid by companies with an interest in Bitcoin (*47*); in the case of Ethereum, some are funded by the Ethereum Foundation itself, while bug-bounties, development grants, and visibility remain as other common incentives (*48*). In this context, our results could suggest that trading on the cryptocurrency market might play the role of incentive for developers to perform certain cross-currency actions. The lack of increase in synchronization for volumes suggests that the observed synchronization of returns is not due to an overall increase in trading interest toward the linked cryptocurrencies. Beyond these two mechanisms, more explanations may exist, and exhausting or testing them, if at all possible, is outside of the scope of this article.

Our results have broad implications. Code has become an important societal regulator that challenges traditional institutions, from national laws to financial markets (*5, 49, 50*). In particular, whether and how financial markets and technological code development interact is an open and debated question (*6, 25, 51, 52*). The case of cryptocurrencies is paradigmatic and still largely unexplored. Cryptocurrencies are open-source digital objects traded as financial
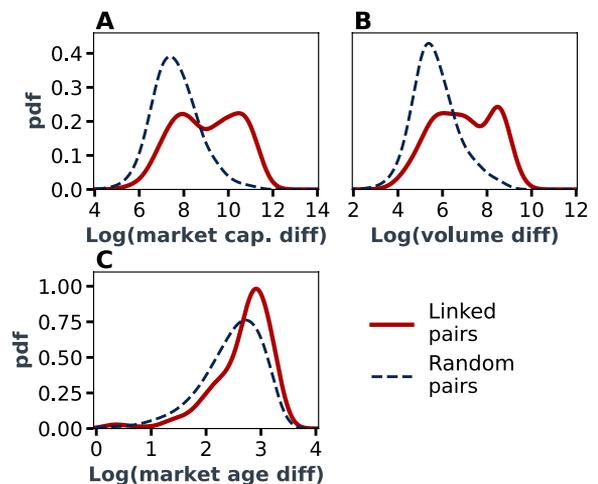

Fig. 4. Linked pair composition. (**A**) Probability density function (pdf) of the difference in market capitalization among cryptocurrencies forming linked pairs (continuous line) and random pairs (dashed line). (**B**) Probability density function of the difference in transaction volume among cryptocurrencies forming linked pairs (continuous line) and random pairs (dashed line). (**C**) Probability density function of the difference in market age at the connection time among cryptocurrencies forming linked pairs (continuous line) and random pairs (dashed line). All the density distributions are computed using a Gaussian Kernel Density Estimation setting the bandwidth values to 0.36.

assets that allow, at least theoretically, everyone to directly shape both an asset structure and its market behavior. Our study, identifying a simple event in the development space that anticipates a corresponding behavior in the market, establishes a first direct link between the realms of coding and trading. In this perspective, we anticipate that our results will be of interest to researchers investigating how code and algorithms may affect the nondigital realm (*53–55*) and spark further research in this direction.

METHODS
Data
The GitHub dataset
GitHub is a service providing a host for software development using Git version control system (27, 28) largely used in a variety of innovation fields, from science to technological development (29). Previous research on the platform focused on the understanding of collaborative structures and developer behavior, showing the importance of social characteristics in the selection of code modifications (30) and of socialization as a precursor of joining a project (31).

A project is stored on GitHub in a so-called "repository," and its production-ready code lives in the "master branch" of the repository (32) [called by default "main branch" starting from 1 October 2020 (33)]. Developers can modify the master branch in two ways, depending on their role. So-called "collaborators" are part of the core development team and can directly edit the code by triggering a "push event." In contrast, "contributors" are anyone who contributed some changes to a project, by submitting their suggestions through a "pull request" that was later accepted and merged by one of the collaborators. Thus, "push" and accepted pull requests are the core events in the development of cryptocurrency production-ready code (34).

We retrieved cryptocurrency GitHub repository names from CoinMarketCap (35). We find that 1668 of the 2225 cryptocurrencies listed in CoinMarketCap as of 9 June 2019 shared their source code on GitHub. Then, we queried the GitHub Archive dataset (36), which stores all events on public repositories from 2011, through Google BigQuery (37). This step provided us with all events related to the development of cryptocurrency GitHub projects. Specifically, we queried two types of events: "push events" and accepted "pull request events." Last, we removed all events triggered by GitHub apps (software designed to maintain and update the repositories), and we removed from our dataset GitHub profiles whose name included the term "bot" to not include noise from users that identified or were reported to be nonhuman.

The market dataset
We collected cryptocurrency daily price, exchange volume, and market capitalization from three different web sources: CoinGecko (15), CryptoCompare (14), and CoinMarketCap (35) (the latter was used only until the end of July 2018 because of updates in the website regulations). We processed the data from CryptoCompare and CoinGecko following a standard procedure (38). We preferred the OpenHighLowClose (OHLC) data from the CryptoCompare Application Programming Interface (API). We adopted as a measure of the transaction volume the amount of USD traded for a crypto on the exchanges registered on CryptoCompare. Similarly, we retrieved the market capitalization of a cryptocurrency using the CoinGecko API and processed it to remove the structural biases found in (38), e.g., we shifted by 1 day all data starting from 30 January 2018.

The price of a cryptocurrency represents its exchange rate (with USD or Bitcoin, typically), which is determined by the market supply and demand dynamics. The exchange volume used is the total trading volume across exchange markets, from dollars to one crypto. The market capitalization is calculated as a product of a cryptocurrency's circulating supply (the number of coins available to users) and its price. We retrieved historical data for currently inactive currencies by querying all the 6000 and more cryptocurrencies recorded in the CoinGeko database (39). Our datasets include market indicators from 3 April 2013 (date by which all the webpages started collecting data) until 30 October 2019. Note that to study the effects

of GitHub development on market indicators, we collected market data for 6 months longer compared to the GitHub data.

In this work, we focus on cryptocurrencies that can be traded with sufficient ease. We, therefore, consider only cryptocurrencies whose trading volume is larger than 100,000 USD (16). We find that 521 cryptocurrencies meet this condition (see table S1 for full list), of which 298 share their code on GitHub.

Randomized pairs
We compare various quantities measured for GitHub-linked pairs to the corresponding values measured for random pairs. A random pair is obtained by (i) extracting 2 of the 521 cryptocurrencies that meet the condition of an average daily market volume larger than 100,000 USD and (ii) verifying that the two extracted cryptocurrencies do not form together a GitHub-linked pair. As for the average volume, days with zero transaction volume (days of market inactivity) were discarded and treated as missing values. The resulting set of 521 cryptocurrencies represents 27% of all the cryptocurrencies with a market history on both CryptoCompare and CoinGecko.

Time series analysis
A cryptocurrency asset return at time t is defined as $R(t) = \frac{P(t) - P(t-1)}{P(t-1)}$, where $P(t)$ is the price (56). The change in market capitalization at t is defined as $C_M(t) = \frac{M(t) - M(t-1)}{M(t-1)}$, where $M(t)$ is the market capitalization. The change in volume is defined as $C_V(t) = \frac{V(t) - V(t-1)}{V(t-1)}$, where $V(t)$ is the volume as time t.

Following a standard approach in time series analysis (57, 58), we measure correlation as the Spearman coefficient between two time series. To compare the correlation across pairs of currencies, following, e.g., Schruben (59), we compute the standardized correlation as

$$S C_k(t) = \frac{C_k(t) - \bar{C}(t)}{\sigma(t)},$$

where $C_k(t)$ is the the correlation time series, computed for a pair k by comparing the return time series $[R_i(t)$ and $R_j(t)]$ of paired assets i and j at time t, and $\bar{C}(t)$ and $\sigma(t)$ are the average correlation and corresponding SD across pairs. At each time step t, the set of pairs used to compute the standardized correlation consists of the pairs for which we had price data at time t.

Error estimate and bootstrapping
We compute the error associated with the average standardized correlation across pairs using bootstrapping (60). For each value of d, representing the number of days before/after the connection time (such that at the connection time $d = 0$): (i) We sample N_d pairs of currencies with replacement, where N_d is the number of existing linked pairs at time d; (ii) we compute the average standardized correlation $SC(d) = \sum_{k=1}^{N_d} S C_k(d) / N_d$ where k is running across the N_d pairs; (iii) we repeat steps (i) and (ii) 10^4 times; and (iv) we compute the mean and SD across the obtained values of $SC(d)$. These values provide an estimate of the average standardized correlation and associated error for the population of linked pairs d days after the connection. We follow the same procedure for random linked pairs.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/51/eabd2204/DC1

REFERENCES AND NOTES

1. S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system, https://bitcoin. org/bitcoin. pdf (2008).
2. A. M. Antonopoulos, *Mastering Bitcoin: Unlocking Digital Cryptocurrencies* (O'Reilly Media Inc., 2014).
3. L. Lessig, *Code and Other Laws of Cyberspace* (Basic Books, 1999).
4. M. W. Brooks, J. Robichaud, The Crypto Paradox: Code is Law and Consensus Rules. *Natl. Law Rev.*, (2019).
5. L. Lessig, *Code: Version 2.0* (Basic Books, 2006).
6. A. Trockman, R. van Tonder, B. Vasilescu, *Proceedings of the 16th International Conference on Mining Software Repositories* (IEEE Press, 2019), pp. 181–185.
7. P. Ceruleo, Bitcoin: A rival to fiat money or a speculative financial asset?, http://tesi.luiss. it/13806/1/ceruleo-paola-tesi-2014.pdf (2014).
8. A. Rogojanu, L. Badea *et al.*, *Theor. Appl. Econ.* **21**, 103 (2014).
9. G. Herlean, Council post: The top 10 risks of bitcoin investing (and how to avoid them), https://forbes.com/sites/forbesfinancecouncil/2018/12/05/the-top-10-risks-of-bitcoin-investing-and-how-to-avoid-them/?sh=f709fc02407e [accessed 06 February 2020].
10. K. Gkillas, P. Katsiampa, An application of extreme value theory to cryptocurrencies. *Econ. Lett.* **164**, 109, 111 (2018).
11. S. Chan, J. Chu, S. Nadarajah, J. Osterrieder, A statistical analysis of cryptocurrencies. *J. Risk Financ. Manag.* **10**, 12 (2017).
12. N. Popper, *The New York Times, Technology* (2017).
13. P. Vigna, M. J. Casey, *The Age of Cryptocurrency: How Bitcoin and the Blockchain are Challenging the Global Economic Order* (Macmillan, 2016).
14. Cryptocompare.com: Live cryptocurrency prices, trades, volumes, forums, wallets, mining equipment, and reviews, http://cryptocompare.com [accessed 17 February 2020].
15. CoinGecko: 360 market overview of coins and cryptocurrencies, https://coingecko.com/en [accessed 6 February 2020].
16. L. Alessandretti, A. ElBahrawy, L. M. Aiello, A. Baronchelli, Anticipating cryptocurrency prices using machine learning. *Complexity* **2018**, 8983590 (2018).
17. E.-T. Cheah, J. Fry, Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Econ. Lett.* **130**, 32–36 (2015).
18. W. C. Wei, Liquidity and market efficiency in cryptocurrencies. *Econ. Lett.* **168**, 21–24 (2018).
19. S. Chan, J. Chu, S. Nadarajah, J. Osterrieder, A statistical analysis of cryptocurrencies. *J. Risk Financ. Manag.* **10**, 12 (2017).
20. L. H. White, The market for cryptocurrencies. *Cato J.* **35**, 383 (2015).
21. J. Fry, E.-T. Cheah, Negative bubbles and shocks in cryptocurrency markets. *Int. Rev. Financ. Anal.* **47**, 343–352 (2016).
22. V. Kallinterakis, Y. Wang, Do investors herd in cryptocurrencies–and why? *Res. Int. Bus. Financ.* **50**, 240–245 (2019).
23. A. ElBahrawy, L. Alessandretti, A. Kandler, R. Pastor-Satorras, A. Baronchelli, Evolutionary dynamics of the cryptocurrency market. *R. Soc. Open Sci.* **4**, 170623 (2017).
24. N. Gandal, H. Halaburda, Can we predict the winner in a market with network effects? Competition in cryptocurrency market. *Games* **7**, 16 (2016).
25. S. Wang, J.-P. Vergne, Buzz factor or innovation potential: What explains cryptocurrencies' returns? *PLOS ONE* **12**, e0169556 (2017).
26. R. van Tonder, A. Trockman, C. L. Goues, *Proceedings of the 16th International Conference on Mining Software Repositories* (IEEE Press, 2019), pp. 186–190.
27. P. Bell, B. Beer, *Introducing GitHub: A Non-Technical Guide* (O'Reilly Media, 2014).
28. E. Kalliamvakou, G. Gousios, K. Blincoe, L. Singer, D. M. German, D. Damian, *Proceedings of the 11th Working Conference on Mining Software Repositories*, MSR 2014 (ACM, 2014), pp. 92–101.
29. J. Perkel, Democratic databases: Science on GitHub. *Nat. News* **538**, 127–128 (2016).
30. J. Tsay, L. Dabbish, J. Herbsleb, *Proceedings of the 36th International Conference on Software Engineering* (ACM, 2014), pp. 356–366.
31. C. Casalnuovo, B. Vasilescu, P. Devanbu, V. Filkov, *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering* (ACM, 2015), pp. 817–828.
32. GitHub Guides, https://guides.github.com/ [accessed 20 December 2019].
33. GitHub, Renaming the default branch from master (2020), https://github.com/github/renaming [accessed 3 September 2020].
34. L. Dabbish, C. Stuart, J. Tsay, J. Herbsleb, *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (ACM, 2012), pp. 1277–1286.
35. Cryptocurrency market capitalizations|coinmarketcap (2020); https://coinmarketcap.com/ [accessed 6 February 2020].
36. Gh archive, https://gharchive.org/ [accessed 6 February 2020].
37. Google BigQuery, https://console.cloud.google.com/bigquery [accessed 2 December 2020].
38. C. Alexander, M. Dakos, Available at SSRN 3382828 (2019).
39. CoinGecko, All cryptocurrencies, https://coingecko.com/en/coins/all [accessed 6 February 2020].
40. S. Miccichè, G. Bonanno, F. Lillo, R. N. Mantegna, Degree stability of a minimum spanning tree of price return and volatility. *Phys. A* **324**, 66–73 (2003).
41. G. Bonanno, G. Caldarelli, F. Lillo, S. Micciche, N. Vandewalle, R. N. Mantegna, Networks of equities in financial markets. *Eur. Phys. J. B* **38**, 363–371 (2004).
42. B. L. Welch, The generalisation of student's problems when several different population variances are involved. *Biometrika* **34**, 28–35 (1947).
43. G. D. Ruxton, The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behav. Ecol.* **17**, 688–690 (2006).
44. GitLab, The first single application for the entire DevOps lifecycle, https://about.gitlab.com/ [accessed 19 February 2020].
45. L. Kristoufek, What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis. *PLOS ONE* **10**, e0123923 (2015).
46. X. Li, C. A. Wang, The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin. *Decis. Support. Syst.* **95**, 49–60 (2017).
47. A. V. Wirdum, Who funds bitcoin core development? How the industry supports bitcoin's 'reference client,' 6 April 2016; https://bitcoinmagazine.com/articles/who-funds-bitcoin-core-development-how-the-industry-supports-bitcoin-s-reference-client-1459967859.
48. S. Voshmgir, *Token Economy: How Blockchains and Smart Contracts Revolutionize the Economy* (BlockchainHub, 2019).
49. P. D. F. De Filippi, *Blockchain and the Law: The Rule of Code* (Harvard Univ. Press, 2018).
50. R. Houben, A. Snyers, *EU publications; Directorate-General for Internal Policies of the Union* (2018).
51. D. G. Baur, D. Cahill, K. Godfrey, Z. F. Liu, Bitcoin time-of-day, day-of-week and month-of-year effects in returns and trading volume. *Financ. Res. Lett.* **31**, 78–92 (2019).
52. S. Corbet, B. Lucey, A. Urquhart, L. Yarovaya, Cryptocurrencies as a financial asset: A systematic analysis. *Int. Rev. Financ. Anal.* **62**, 182–199 (2019).
53. V. Mayer-Schönberger, Can we reinvent the internet? *Science* **325**, 396–397 (2009).
54. J. Yuan, T. Huang, P. Cheng, Y. Zou, H. Zhang, J. L. Yang, S.-Y. Chang, Z. Zhang, W. Huang, R. Wang, D. Meng, F. Gao, Y. Yang, Enabling low voltage losses and high photocurrent in fullerene-free organic photovoltaics. *Nat. Commun.* **10**, 570 (2019).
55. M. Ali, P. Sapiezynski, M. Bogen, A. Korolova, A. Mislove, A. Rieke, Discrimination through optimization: How Facebook's Ad delivery can lead to biased outcomes. *Proceed. ACM Hum. Comput. Interact.* **3**, 1 (2019).
56. F. Lillo, R. N. Mantegna, Variety and volatility in financial markets. *Phys. Rev. E* **62**, 6126–6134 (2000).
57. S. Miccichè, G. Bonanno, F. Lillo, R. N. Mantegna, Degree stability of a minimum spanning tree of price return and volatility. *Phys. A* **324**, 66–73 (2003).
58. W. H. Press, S. A. Teukolsky, B. P. Flannery, W. T. Vetterling, *Numerical Recipes in Fortran 77: Volume 1, Volume 1 of Fortran Numerical Recipes: The Art of Scientific Computing* (Cambridge Univ. press, 1992).
59. L. Schruben, Confidence interval estimation using standardized time series. *Oper. Res.* **31**, 1090 (1983).
60. M. R. Chernick, *Bootstrap Methods: A Guide for Practitioners and Researchers* (John Wiley & Sons, 2011), vol. 619.
61. S. Alizadeh, M. W. Brandt, F. X. Diebold, Range-based estimation of stochastic volatility models. *J. Finance* **57**, 1047–1091 (2002).
62. A. R. Gallant, C.-T. Hsu, G. Tauchen, Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Rev. Econ. Stat.* **81**, 617–631 (1999).

Citation: L. Lucchini, L. Alessandretti, B. Lepri, A. Gallo, A. Baronchelli, From code to market: Network of developers and correlated returns of cryptocurrencies. *Sci. Adv.* **6**, eabd2204 (2020).

Science Advances

From code to market: Network of developers and correlated returns of cryptocurrencies

Lorenzo Lucchini, Laura Alessandretti, Bruno Lepri, Angela Gallo and Andrea Baronchelli

ARTICLE TOOLS	http://advances.sciencemag.org/content/6/51/eabd2204
SUPPLEMENTARY MATERIALS	http://advances.sciencemag.org/content/suppl/2020/12/14/6.51.eabd2204.DC1
REFERENCES	This article cites 30 articles, 1 of which you can access for free http://advances.sciencemag.org/content/6/51/eabd2204#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service