Can Individual Human Financial Behaviour Be Mathematically Modelled? A Case Study of Elon Musk's Dogecoin Tweets

Juraj Medzihorský[1] | *Matej Bel University, Banská Bystrica, Slovakia*
Peter Krištofík | *Matej Bel University, Banská Bystrica, Slovakia*

Received 11.2.2022, **Accepted** (reviewed) 23.3.2022, **Published** 17.6.2022

Abstract

The price of Dogecoin has been influenced by Elon Musk's tweets on several occasions. Moreover, there are repeating patterns in the Dogecoin prices. However, is there also a pattern to the timing of the tweets? Applying linear regression, we have been able to make the reverse analysis – to use hard financial data (prices) to analyse the human behaviour (tweets) that preceded and influenced the financial data. Selected tweets could be paired thanks to the projections of their timing on the regression line that had been created over the prices. Our model exhibits inaccuracies only in the order of the days. That is surprising, as pump schemes do not usually require such a high level of long-term deterministic timing.

Keywords	***DOI***	***JEL code***
Behavioral economics, cryptocurrencies, pump-and-dump scheme, linear regression, time series analysis	*https://doi.org/10.54694/stat.2022.9*	*G17, G41*

INTRODUCTION

Human behaviour has become an important component of financial market research, in addition to hard financial data. The proximate determinants of stock prices are supply and demand. That is, human activity is affected by sentiment as well as by firms` results. A clear example was the rise in the value of GameStop, caused by Reddit users (Morgia et al., 2021). Cryptocurrencies, and especially memecoins are more influenced by sentiment than are stocks, because they lack an agreed valuation standard. Social media amplifies these effects.

[1] Matej Bel University, Faculty of Economics, Department of Finance and Accounting, Tajovského 10, 975 90 Banská Bystrica, Slovakia. Corresponding author: e-mail: juraj.medzihorsky@umb.sk, phone: (+421)484466111.

Many studies have focussed on modelling human behaviour. For example, Pentland (2006), Aipperspach et al. (2006), Lieder and Griffiths (2019). It turns out that some features can apply to both economic and non-economic behaviour. For example, Aipperspach et al. (2006) showed that the highly-skewed power-law distribution could model human movements in a house. The same distributions are also suitable for modelling company size (Lyócsa and Výrost, 2018), and also transactions between crypto wallets. Anomalies from these expected distributions can be caused by non-human activity – in the case of crypto wallets – by trading-bots (Zwang et al., 2018).

The use of trading-bots is not new. But they can now be much more efficient thanks to pump-and-dump schemes. These are common in cryptocurrency markets (Kamps and Kleinberg, 2018; Xu and Livshits, 2019). And so is the impact of influencers like Elon Musk. The relations between his tweets and the price of bitcoin was confirmed by Tandon et al. (2021), and for Dogecoin by Cary (2021). His tweets can have an almost immediate major impact on the price of any altcoin. Occasionally, such impact can happen regardless of the intention. A quite bizarre recent example was Elon Musk's tweet that mentioned J. R. R. Tolkien's idea about free public ducks (Twitter, 2022). This caused a significant price rise in the homonymous cryptocurrency.

The aim of our paper is to model the timing of Elon Musk's Dogecoin tweets. Our hypotheses are:

H1: Behaviour of a market-influencer can be reversely derived from asset price movements.

H2: There is a pattern to the timing of Elon Musk's tweets.

1 MATERIAL AND METHODS

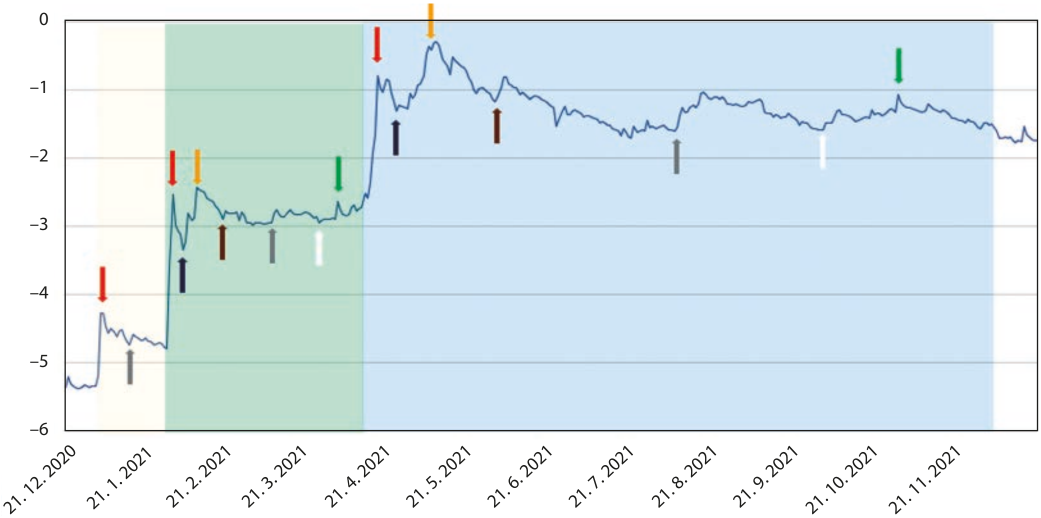
The methods used in the analysis are linear regression, and exploratory analysis. There are two datasets (Coindesk, 2022; Twitter, 2022): the time series of Dogecoin prices (daily maximums of DOGE/USD) since the large pump event of 28/01/2021, and the time series of Elon Musk's Dogecoin-related tweets during the same period. Daily maxima are used, as we intend to analyse pumps. During modelling all dates are converted to simple numbers. For example, 28/01/2021 = 1, and so on. We begin by setting out the time model of Dogecoin prices. As Figure 1 shows, there are similarities in price developments during Period 2 (blue area), compared to Period 1 (green area), and to Period 0 (white area). There are similarities in the patterns of peaks and troughs, though the length of the periods and the size of changes both increase over time. The similarities highlighted in alternative way are showed in Figure A1 in the Annex. This replication was also confirmed in our previous research, with 87% accuracy on a 3-month test set (Medzihorský, 2021). The accuracy was defined as 1 – γ, where γ is a mean error of prediction, calculated as follows:

$$\gamma = \frac{1}{n}\sum_{i=1}^{n} \frac{\left| Predicted\ Price_i - Actual\ Price_i \right|}{Actual\ Price_i} .$$

As the predictions on the test set were calculated in that research only from historical prices by a simple equation, such high accuracy supports the assumption about replication. This is consistent with the results of Uras et al. (2020), who confirm the existence of time regimes for selected cryptocurrencies. The regimes differ from a random walk: for the best forecasts they recommend taking 200-day sequences. In addition, Ozdamar et al. (2021) confirmed high correlations between expected returns on crypto-currencies and daily maxima during the previous month. So crypto-markets are not purely stochastic. They exhibit some level of a determinism.

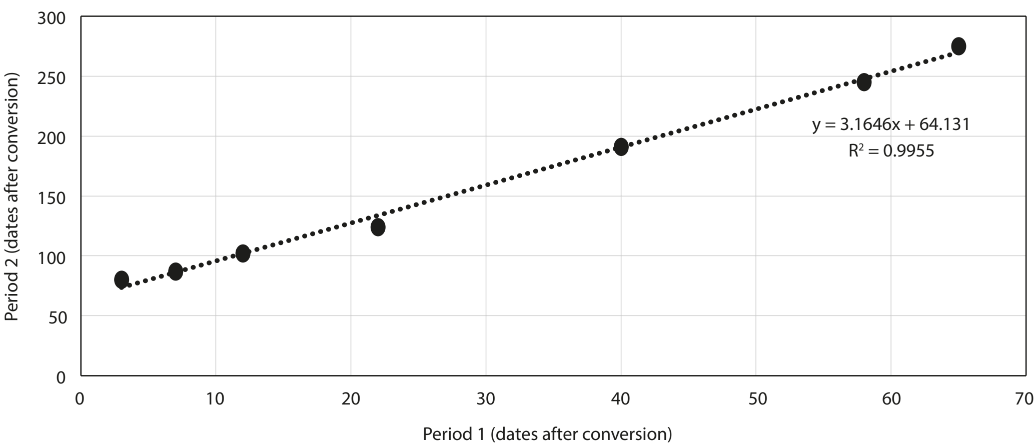
If we assume that price developments during different periods follow the same general pattern (see arrows on Figure 1), then combining points from Periods 1 and 2 we can produce the time series regression in Figure 2. The selected points represent local minimums, maximums, or high daily yields. This model will help predict the timing of tweets. The final step of the analysis is a simple projection of the dates of tweets on the same regression line that was created in the time series model of price development.

Figure 1 Logarithmic price of Dogecoin with highlighted replication of the shapes



Note: Logarithmic scale (using natural logarithm) is used for clearer illustration of the shapes replication that is not obvious on a figure with the linear scale.
Source: Own processing from Coindesk (2022), and Medzihorský (2021)

Figure 2 Regression of selected time points



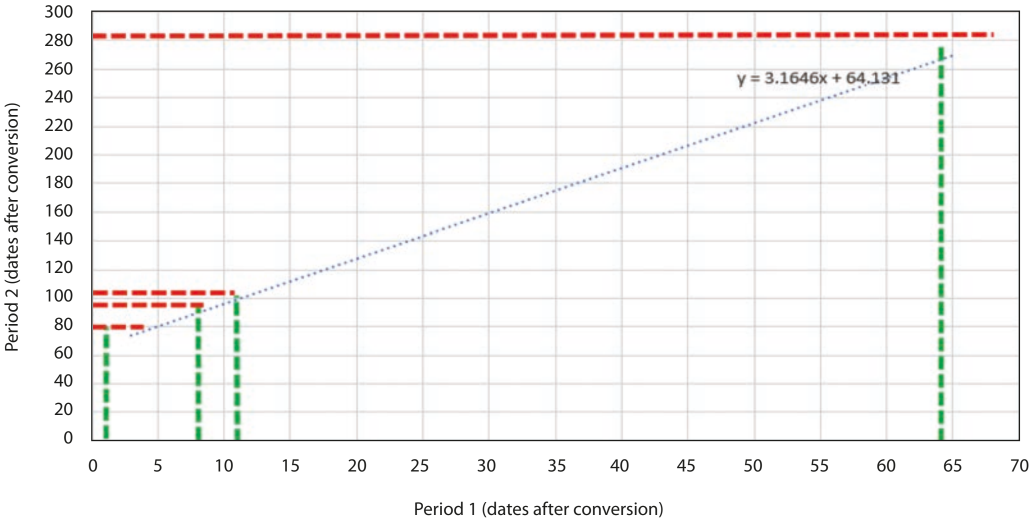
Note: Dates are converted to simple numbers. For example, 28/01/2021 = 1, and so on. The selected time points in Figure 2 are represented by the arrows in Figure 1.
Source: Own processing

2 RESULTS

There is almost a perfect correlation between the timing of selected points in Period 2 and Period 1 (see Figure 2). We can confirm that the timing of selected movements of Dogecoin price is significantly deterministic, and the lengthening of shapes is linear. However, this model is only auxiliary. We use it to find a suitable line to fit the timing of the tweets. Doing so, we also observe the determinism in the timing of the tweets (see Figure 3). As the intersections of the projections of tweet dates are approximately on the regression line, the hypotheses H1 and H2 are confirmed.

However, our approach has important limitations. Only selected tweets can be analyzed this way – there are more tweets in Period 2 than in Period 1, so some cannot be paired together. Only a limited period is studied. Finally, the exact timing cannot be calculated by a simple line – there are some inaccuracies in the order of the days – see Figure 3 and Table 1.

Figure 3 Timing of the tweets and their projection on the regression line of Dogecoin time points



Source: Own processing from Twitter (2022)

Applying the equation y = 3.1646x + 64.131 (Formula 1) to the converted dates of the tweets, produces the predictions shown in Table 1. The table shows some repeating inaccuracies. Predictions can be improved by incorporating this knowledge. We estimate the next tweet will be on 12 February 2022.

Table 1 Tweet predictions based on Formula 1

Date of original tweet	Converted date *(x)*	Future tweet estimate *(y)*	Re-converted date of estimate	Real timing of future tweet	Inaccuracy
1/28/2021	1th	67	4/4/2021	4/15/2021	11
2/4/2021	8	89	4/26/2021	4/28/2021	2
2/7/2021	11	99	5/6/2021	5/7/2021	1
4/1/2021	64	267	10/21/2021	10/27/2021	6
4/15/2021	78	311	12/4/2021	12/14/2021	10
4/28/2021	91	352	1/14/2022	1/14/2022	0
5/7/2021	100	381	2/12/2022	N/A	N/A

Note: Twitter should show date and time in the user's time zone. In our case GMT+1 or 2 (depending on summer/winter time).
Source: Own processing from Twitter (2022)

Not only timing but also the quality and price impact of the tweets play a role (see Table 2). While the permanent impact of some earlier tweets – which led to a several-fold increase in price – could be valuable for Dogecoin holders, current tweets cause only pumps and dumps. There may be several reasons for this. First, the crypto market as whole is not currently achieving yields as high as it was, up to, say, May 2021, when Dogecoin recorded its all-time-high. Second, Dogecoin tweets are more common. They are not novel. Third, alternative dog-based memecoins have been recently created. Perhaps surprisingly, even though recent Dogecoin tweets have included more economic context, this has not changed market reactions. In fact, memecoins depend on market sentiment. So, a tweet that includes some genuinely informed economic analysis does not necessarily change prices more than one without it. Also, one has to wonder if the supply of new Dogecoin fans has been exhausted – and existing fans` demands are satiated. On the other hand, demand for Dogecoin or any cryptocurrency can be positively influenced by the growing inflation as the inflation negatively affects, especially, cash holdings (Pintér and Mešťan, 2020).

There has also been a change in the form of the tweets. Pictures are no longer used. Putting words like 'doge' or 'Dogecoin' directly in the text can be more profitable for traders using bots than for others.

Table 2 Qualitative value and impact of the tweets

Date	Quality of information	Form	Price impact	Duration of impact
1/28/2021	N-E	Text in picture	867%	Permanent
2/4/2021	N-E	Multiple tweets	90%	Permanent
2/7/2021	N-E	Picture	86%	Temporary
4/1/2021	D	Text	31%	Permanent
4/9/2021	N-E	Picture	8%	Permanent
4/15/2021	D	Text + picture	387%	Permanent
4/28/2021	N-E	Text	30%	Temporary
5/7/2021	N-E	Picture	15%	Pump-and-dump
5/11/2021	E	Text	-2%	Pump-and-dump
5/14/2021	E	Text	31%	Pump-and-dump
5/20/2021	N-E	Text + picture	0%	Pump-and-dump
6/2/2021	N-E	Picture	37%	Pump-and-dump
7/2/2021	N-E	Text in picture	4%	Nearly zero effect
7/25/2021	E	Text in picture	21%	Pump-and-dump
9/22/2021	E	Text	7%	Pump-and-dump
10/27/2021	D	Text	28%	Pump-and-dump
10/31/2021	E	Text	2%	Nearly zero effect
12/14/2021	E	Text	26%	Pump-and-dump
12/23/2021	E	Text	14%	Pump-and-dump
1/14/2022	E	Text	33%	Pump-and-dump

Note: N-E – non-economic information; E – tweets with serious economic information like an acceptance of Dogecoin by a merchant; D – economic value depends on wider context. For example, putting literal Dogecoin on the literal Moon would not be economic relevant. However, it actually is relevant, as lunar cargo for DOGE-1 mission is financed by Dogecoin. Permanent price impact means that price has not declined lower than it was before the tweet; Temporary impact means that the price remained higher than before the tweet for one or more weeks. Pump-and-dump represents a quick decline within a week of the pump. Price impact is a yield, calculated as follows: Yield = max {Daily high price $_t$, Daily high price $_{t+1}$, Daily high price $_{t+2}$} / Closing price $_{t-1}$ – 1; where t represents the date of the tweet.
Source: Own processing from Twitter (2022), and Coindesk (2022)

An important limitation of the calculation of price impact (see Table 2) – using a comparison with the closing price of the previous day – lies in price changes shortly before a tweet. An example is the tweet on 14 January 2022, when there was a significant rise in price one hour before the tweet. Contrary negative examples are the tweets on 11 and 20 May, when a decline of price before the tweets distorted the calculations. The actual effects of these tweets were, of course, positive.

Our results raise several questions. Only selected tweets from Period 2 can be paired with the tweets from Period 1. These tweets from Period 2 can be paired with most recent tweets. However, do the rest of the tweets from Period 2 – which cannot be paired with the tweets in Period 1 – have the same predictive power? Will there ever again be a tweet that causes a yield of more than 100%, with a permanent price impact. Or will we only observe pump-and-dumps, with no more than 30% yields. Can past inaccuracies in our model be used to achieve more precise predictions, or to analyse a wider range of cases? What are the reasons, if any, for the price increase before the tweet on 14 January 2022? As these questions remain unanswered, the need for continuing research is clear.

CONCLUSION

It is obvious that Elon Musk influences the price of Dogecoin. So, we have been able to reverse analyse his behaviour – to use hard financial data (prices) to analyse the human behaviour (tweets) that preceded and influenced financial data. If there had been no repeating patterns in Dogecoin prices, or in the timing of the tweets, we would have been unable to model the timing of human behaviour by a simple line. However, what motivation, if any, might lie behind the timing of the tweets, remains hidden. Pumps do not require such a high level of long-term deterministic timing. Nor is it clear that the pattern of the timing of the tweets could partly determine the long-term price development of Dogecoin. Our contribution is only a first step in this analysis. Therefore, the paper is intentionally structured using a single-issue approach, as we expect wider discussion and further research of this issue in the future. The aim of more complex studies should include the analysis and prediction of the behaviour of ordinary traders from price movements, the analysis of other market influencers, and searching for any deterministic trends in such areas where stochasticity would be usually expected.

ACKNOWLEDGMENT

The authors wish to thank Dr. Colin W. Lawson from University of Bath for valuable comments and corrections of the manuscript.

References

AIPPERSPACH, R., COHEN, E., CANNY, J. (2006). Modelling Human Behaviour from Simple Sensors in the Home [online]. In: FISHKIN, K. P., SCHIELE, B., NIXON, P., QUIGLEY, A. (eds.) *Pervasive Computing*, Pervasive, Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, Vol. 3968. <https://doi.org/10.1007/11748625_21>.

CARY, M. (2021). Down with the #Dogefather: Evidence of a Cryptocurrency Responding in Real Time to a Crypto-Tastemaker [online]. *Journal of Theoretical and Applied Electronic Commerce Research*, 16: 2230–2240. <https://doi.org/10.3390/ jtaer16060123>.

COINDESK. (2022). *Dogecoin* [online]. <https://www.coindesk.com/price/dogecoin>.

KAMPS, J., KLEINBERG, B. (2018). To the moon: defining and detecting cryptocurrency pump-and-dumps [online]. *Crime Science*, 7(18): 1–18. <https://doi.org/10.1186/s40163-018-0093-5>.

LIEDER, F., GRIFFITHS, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources [online]. *Behavioral and Brain Sciences*, 43(1): 1–60. <https://doi.org/10.1017/S0140525X1900061X>.

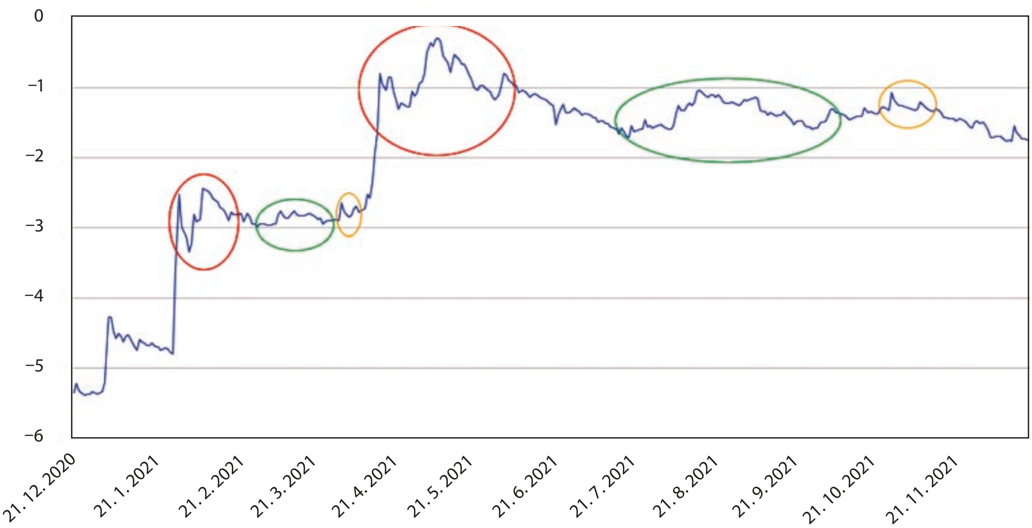
LYÓCSA, Š., VÝROST, T. (2018). Scale-free distribution of firm-size distribution in emerging economies [online]. *Physica A*, 508: 501–505. <https://doi.org/10.1016/j.physa.2018.05.088>.

MEDZIHORSKÝ, J. (2021). Dogecoin price prediction – can be a determinism supposed? [online]. *Journal of Economics and Social Research* 22(2): 67–81. <https://doi.org/10.24040/eas.2021.22.2.67-81>.

MORGIA, L., MEI, A., SASSI, F., STEFA, J. (2021). *The Doge of Wall Street: Analysis and Detection of Pump and Dump Cryptocurrency Manipulations* [online]. Ithaca, NY: Cornell University. <https://arxiv.org/abs/2105.00733v1>.
OZDAMAR, M., AKDENIZ, L., SENSOY, A. (2021). Lottery like preferences and the MAX effect in the cryptocurrency market [online]. *Financial Innovation,* 7(74): 1–27. <https://doi.org/10.1186/s40854-021-00291-9>.
PENTLAND, A. (2007). Automatic mapping and modeling of human network [online]. *Physica A*, 378: 59–67. <https://doi.org/10.1016/j.physa.2006.11.046>.
PINTÉR, Ľ., MEŠŤAN, M. (2020). *Kolektívne investovanie*. 1st Ed. Banská Bystrica: Belianum.
TANDON, C., REVANKAR, S., PALIVELA, H., PARIHAR, S. (2021). How can we predict the impact of the social media messages on the value of cryptocurrency? Insights from big data analytics [online]. *International Journal of Information Management Data Insights*, 1: 100035. <https://doi.org/10.1016/j.jjimei.2021.100035>.
TWITTER. (2022). *Elon Musk* [online]. <https://twitter.com/elonmusk>.
URAS, N., MARCHESI, L., MARCHESI, M., TONELLI, R. (2020). Forecasting Bitcoin closing price series using linear regression and neural networks models [online]. *Peer J. Computer Science*, 6: e27. <https://doi.org/10.7717/peerj-cs.279>.
XU, J., LIVSHITS, B. (2019). *The Anatomy of a Cryptocurrency Pump-and-Dump Scheme* [online]. Ithaca, NY: Cornell University. <https://arxiv.org/abs/1811.10109>.
ZWANG, M., SOMIN, S., PENTLAND, A., ALTSHULER, Y. (2018). *Detecting Bot Activity in the Ethereum Blockchain Network* [online]. Ithaca, NY: Cornell University. <https://arxiv.org/abs/1810.01591>.

ANNEX

Figure A1 Logarithmic price of Dogecoin with alternatively highlighted replication of the shapes



Note: Logarithmic scale (using natural logarithm) is used for clearer illustration of the shapes replication that is not obvious on a figure with linear scale.
Source: Own processing from Coindesk (2022)