Department of Political Economy

King's College London

Dissertation cover sheet

Programme: History and Political Economy BSc

Module code: 6SSPP352

Candidate number: AE13330

Title of your dissertation: Unraveling the Interplay: Exploring the Nexus between COVID-19 and Herding Behavior in Cryptocurrency Markets

Supervisor name: Dr. Teresa Esteban-Casanelles

Word count*: 8453

Table of Contents

Abstract

Cryptocurrency markets are unregulated, highly volatile environments. This quality often causes them to behave in ways that normative models of markets cannot explain. As a result, examining this volatility from psychological and sociological axioms becomes necessary. I aim to examine the impact of a single cognitive bias, herd behavior, and its relationship with a black swan event, the COVID-19 pandemic. I employ the robust cross-sectional absolute dispersion measure (CSAD) as a proxy for herding to elucidate relationships with daily cases to understand how herding fluctuates in response to heightened volatility. I initially expected herding to increase in response to these events. However, I discovered that in the two markets I examine, a significant decrease in herding in the USA market and no significant change in herding in the European market is observed. This prompted me to posit that behavioral phenomena such as psychological reinforcement and an artificial mechanism called a ‘nudge’ increase the risk threshold or risk reference point of investors consistent with prospect theory, which builds resilience towards exogenous events that prevents them from succumbing to the herd mentality.

Acknowledgements:

A special thanks to my supervisor, Dr. Teresa Esteban-Casanelles, for helping me with my research objectives. I would also like to thank Dr. Andrew McFaull for providing guidance on cognitive biases.

Introduction:

1.1 The Setting:

The financial world is a complex and intricate sphere, one where normative models and theories have informed much of our understanding of its function and composition. However, these methods have been shown to be imperfect, and in many instances normative models fail to adequately model certain phenomena. Recent studies have shown that certain behavioral habits are the culprit of these unpredicted dynamics. These behavioral habits or 'cognitive biases' are an instinctive component of our decision-making process, in many instances overriding reason and rationality. Their manifestation can often lead to inefficiencies in places ranging from financial markets to grocery stores. For instance, market bubbles are a phenomenon that occur quite frequently in financial markets. Although bubbles initially lead to high growth rates, the eventual bursting of these bubbles, such as due to overvaluation or a loss of confidence, can lead to large-scale defaults and downturns. Thomas Lux asserts that normative models such as the efficient market hypothesis (EMH) cannot adequately explain the existence of phenomena such as bubbles and how they cause prices to deviate from their fundamental values. As a result, consulting sociological and psychological models to explain the behaviors that lead to these trends becomes necessary[1]. Scholars like Richard Thaler have also stated that although economics has for a long time been researched as a natural science, it was always inherently a human science due to the psychological factors at play such as

[1] Lux, Thomas, 'Herd Behavior, Bubbles, and Crashes', *The Economic Journal*, 105 (1995), 881

cognitive biases[2]. This paper examines one important bias called herd behavior or the herd mentality. It occurs when investors succumb to the fear of missing out (FOMO) or panic and follow the actions of other investors instead of consulting their own independent advice. My paper will seek to explore herd behavior's relationship with exogenous shocks, particularly through the lens of cryptocurrency markets during the COVID-19 pandemic. Cryptocurrency markets and the pandemic represent an opportunity for accurately analyzing herd behavior for several reasons. Firstly, cryptocurrency markets are unregulated, allowing for an analysis of herding without regulatory interference. Secondly, the pandemic is an exogenous event whose impact on financial markets and the economy is comparable to the 2008 financial crisis and the Great Depression[3]. These two components, coupled together, allow for the effects of herding on financial markets to be observed when it is in its most volatile, unabated state. Understanding the dynamics of cryptocurrency markets during these events is imperative to learning how to address any potential complications that may arise in the future.

1.2 The Pandemic:

The exact mechanisms through which the pandemic has affected herding have not been clearly identified. It is generally agreed that since the transmission of information and news about cryptocurrency markets relies heavily on social media, social contagion is a probable medium through which herding can be influenced by the pandemic. This is specifically the

[2] Thaler, Richard, in *Misbehaving: The Making of Behavioral Economics,* (New York 2015)
[3] https://www.imf.org/en/Blogs/Articles/2020/04/14/blog-weo-the-great-lockdown-worst-economic-downturn-since-the-great-depression, accessed 22 April 2024

case with case data and deaths, but more importantly the speed at which both are transmitted. Case data is far quicker in its transmission than deaths simply because the process of reporting is more streamlined. The reason for this is so that individuals can isolate themselves as quickly as possible to prevent further infection. Social media relies on the conciseness and speed at which narratives can be communicated, and thus many platforms such as Twitter (X) or Instagram often features posts from accounts detailing increases or peaks in daily cases. The way this links to herding is that the way this information is communicated is meant to capture the attention of users by capitalizing on their fears. As a result, this exacerbates social contagion, causing people to monitor and follow the actions of others in similar situations. This is corroborated by Avery and Zemsky (1999) who show that herding proliferates when uncertainty in markets persists[4]. COVID cases may create the same fears and tensions, making it an important variable for analysis.

1.5 Rationale and Research Questions:

What exactly is the importance of examining herding in cryptocurrency markets? Firstly, cryptocurrency markets are an emerging financial powerhouse, with a market cap of over 2.2 trillion USD in 2024, a 7.77% increase from the previous year[5]. Cryptocurrencies will continue to remain relevant in an increasingly digital world, as it streamlines transactions while still maintaining the same level of security that government-regulated assets do.

---

[4] Marco Cipriani and Antonio Guarino, ‘Introduction’, in *Herd Behavior in Financial Markets: An Experiment with Financial Market Professionals*, (2008), 4
[5] https://www.mordorintelligence.com/industry-reports/cryptocurrency-market, accessed 14 April 2024

Their increasingly relevant status in the financial world provides me with an incentive to understand how they behave. Cryptocurrencies do not follow the usual patterns of more regulated markets such as bond and equity exchanges, and thus require a more unique, behavioral approach to understanding how they fluctuate and what causes said fluctuations. The increasing importance of social media as well as the predominance of inexperienced investors shapes cryptocurrency markets into a highly volatile, rapidly changing landscape for trading. As a result, I want to examine what the relationship between herding and exogenous events like the pandemic is to get a better understanding of its impact on these increasingly relevant markets. I aim to answer the following research questions:

1. Did the COVID-19 pandemic influence herding significantly during the pandemic period (PP)?
2. How does the relationship between COVID-19 and herding evolve over time?

Literature Review:

This study makes several assumptions. By attempting to model herding during the pandemic, I not only assume that herding exists in markets, but that it contributes to market inefficiencies, leading to anomalies such as bubbles, crashes, and insider trading. The decision to position the paper in this way is informed by several seminal works. The first is Daniel Kahnemann and Amos Tversky's "Prospect Theory: An Analysis of Decision Under Risk". Their experiments showed that in different scenarios, depending on the framing of the scenario or the presence of risk, they observed preference reversal and non-linear risk aversion, which violated the von Neumann-Morgenstern Expected Utility Theory's assumptions of transitivity and linear risk aversion[6] [7]. They were able to successfully rationalize the behavior in their experiments through prospect theory, showing that individuals make these decisions due to factors such as biases, experiences, and psychological reference points. Their ability to reveal discrepancies between descriptive data and normative models highlights the importance of examining economic behavior through psychological axioms. This study uses this revelation as the foundation for its analysis on herd behavior.

By situating the study in behavioral theory, this paper goes against the assumptions of a normative model called the Efficient Market Hypothesis. It was first introduced by Eugene Fama in his seminal work, “Efficient Capital Markets: A Review of Theory and Empirical Work” (1970). He argued that markets are inherently efficient and fluctuate between three different forms. The first is weak-form efficiency, where asset prices reflect all past trading information, such as historical prices and trading volumes. This implies that future price movements cannot be predicted because current prices do not experience delays in incorporating past information. The second is semi-strong form, which builds upon weak-form efficiency by stating that not only past but also public information is incorporated into prices. The third and most efficient is strong form, where all public and private past information is incorporated[8]. These forms of efficiency posited by Fama highlight an

---

[6] Kahnemann, Daniel, and Tversky, Amos, ‘Prospect Theory: An Analysis of Decision Under Risk’, 47 (1979), 263

[7] John von Neumann and Oskar Morgenstern, in *Theory of Games and Economic Behavior*, (Princeton 1944), 38

[8] Fama, Eugene F., ‘Efficient Capital Markets: A Review of Theory and Empirical Work.’ *Journal of Finance*, 25 (1970), 383

important attribute of markets. As more investors attempt to exploit price inefficiencies, competition drives prices towards their fair value, making abnormal returns difficult to achieve on a consistent basis.[9] In markets described by the EMH, inefficiencies do not exist for long and anomalies that cause price deviations do not cause major changes to market dynamics. As a result, this paper's acknowledgement of the existence of anomalies like herding inherently challenges the EMH.

Several studies, such as by Morris and Shin (1999) and Persuad (2000), have shown that herding not only exists in equity markets, but that it also has the tendency to amplify volatility.[10] Thomas Lux also showed that it tends to have cyclical relationships with anomalies such as bubbles[11]. Additionally, Yarovaya et al. utilize the cross-sectional absolute dispersion (CSAD) measure in cryptocurrency markets to successfully identify herding during the pandemic[12]. The CSAD measure represents a highly useful proxy for measuring herd behavior, having also been used in Christie and Huang's study and Chang et. al's work in equity markets[13]. The ability of this measure to identify herding in a robust fashion makes it a useful tool for the analysis. In addition, other studies focusing on herding, such as Jirasakuldech et al., show that herding occurs during extreme market movements, in

---

[9] Ibid. 414
[10] Yarovaya, Larisa et al., 'The Effects of a "Black Swan" Event (COVID-19) on Herding Behaviour in Cryptocurrency Markets,' *Journal of International Financial Markets, Institutions & Money*, 75 (2021), 3
[11] Lux, Thomas, 'Herd Behavior, Bubbles, and Crashes', *The Economic Journal*, 105 (1995), 889
[12] Ibid. 4
[13] Ibid. 5

declining market environments, and on days with high trading volume[14]. These studies, to name a few, show concrete evidence of herd behavior in equity, bond, and cryptocurrency markets. This paper builds on these findings by acknowledging concrete evidence of herding and focusing primarily on modelling its behavior with exogenous events.

The final point where our paper distinguishes itself from other studies is in forecasting. Existing studies like Yarovaya et al. primarily focus on the identification of herding without predicting its evolution over time. This paper consequently introduces a level of foresight in the evolution of herd behavior that is absent from other studies.

Research Methodology and Design:

3.1 Data:

I narrowed the sample by deciding to examine two markets where herding was identified unconditionally by Yarovaya et al., the USA and Europe[15]. Additionally, analyzing two separate markets enables the study to identify and control for any regional dynamics that may exist in either market. The dataset used in this study is hourly price data taken from the Bitstamp cryptocurrency exchange.

Due to the unique nature of the pandemic as a black swan event, unique market dynamics would be difficult to identify without a comparative approach[16]. In this case, two periods of

---

[14] Jirasakuldech, B., and Emekter, Riza, 'Empirical Analysis of Investors' Herding Behaviours during the Market Structural Changes and Crisis Events: Evidence from Thailand,' *Global Economic Review*, 50 (2021), 1
[15] Yarovaya, 'Black Swan Event', 13
[16] Yarovaya, 'Black Swan Event', 1

analysis have been chosen. A pandemic period (PP) from November 1st, 2019, to January 1st, 2022, and a post-pandemic period (PPP) from January 30th, 2022, to January 30th, 2024. This was determined after taking several considerations. The COVID-19 disease was first reported in November 2019 in Wuhan, China[17]. However, the exact start date is unknown, which led me to choose the first day of the month where the pandemic began to spread. This allows previous market activity to be captured as well as short-run volatility from the news breaking around the world. The end date of the pandemic period is more difficult to determine. However, I employ simple visualization tactics using the WHO pandemic database to show when covid-19 cases began to dwindle[18]. I find that around January 1st, 2022, covid cases peaked, but then steadily began to decline[19]. It is important to clarify that new daily cases were still being reported after this date in the United States and Europe, and the point at which they began to dwindle differs between the two regions. Nonetheless, January 1st, 2022, represents an approximate start of decline in both regions.

As for the post-pandemic period (PPP), determining the gap between the pandemic and the period of recovery is important. The pandemic still affects the global economy in significant ways in 2024, but its impact is less pronounced. As a result, January 30th, 2022, is chosen as the start of the PPP so that there is not a large gap between the start of the decline in cases in the PP and the PPP. The end date is determined to be January 30th, 2024.

---

[17] Davidson, Helen, 'First Covid-19 case happened in November, China government records show - report,' *The Guardian,* 2020, Available at: https://shorturl.at/dkntZ Accessed 30 March 2024

[18] WHO COVID-19 Dashboard (https://data.who.int/dashboards/covid19/cases?n=c, accessed 23 March 2024)

[19] Ibid.

I chose to analyze two periods so that descriptive data analysis can be used to discover any differences in mean herding between the PP and the PPP. Additionally, I took several cross-sections of the data and analyzed them using an event studies approach. This is especially important for cryptocurrencies because their volatility depends significantly on external, administrative events rather than regulatory events. Controlling for these exogenous variables using a binary variable will improve the accuracy of the models. Subsequently, the analysis will be taken further by introducing a forecasting element. This will allow me to utilize my findings to predict the evolution of herding after the pandemic period. I predict that herding will increase substantially in response to an increase in cases, since this is in line with what many other studies found about herding's relationship with exogenous events. As a result, the analysis will be framed through the following hypotheses, one for our identification component:

> H_1: Herding behavior increases significantly in response to the uncertainty created by changes in daily COVID-19 cases.

One for our forecasting component:

> H_2: The forecast period is expected to see a decrease in herding compared to the PP because of a lower level of uncertainty created by the absence of daily COVID-19 cases.

Another important exogenous event that has significantly influenced markets in the PPP is the Russia-Ukraine War. Its impact on the economy and financial markets continues to persist in 2024. I examine close prices for each cryptocurrency and find only one period where prices experience a significant spike on days where important events in the war occur. This date is the outbreak of the war, February 24th, 2022.

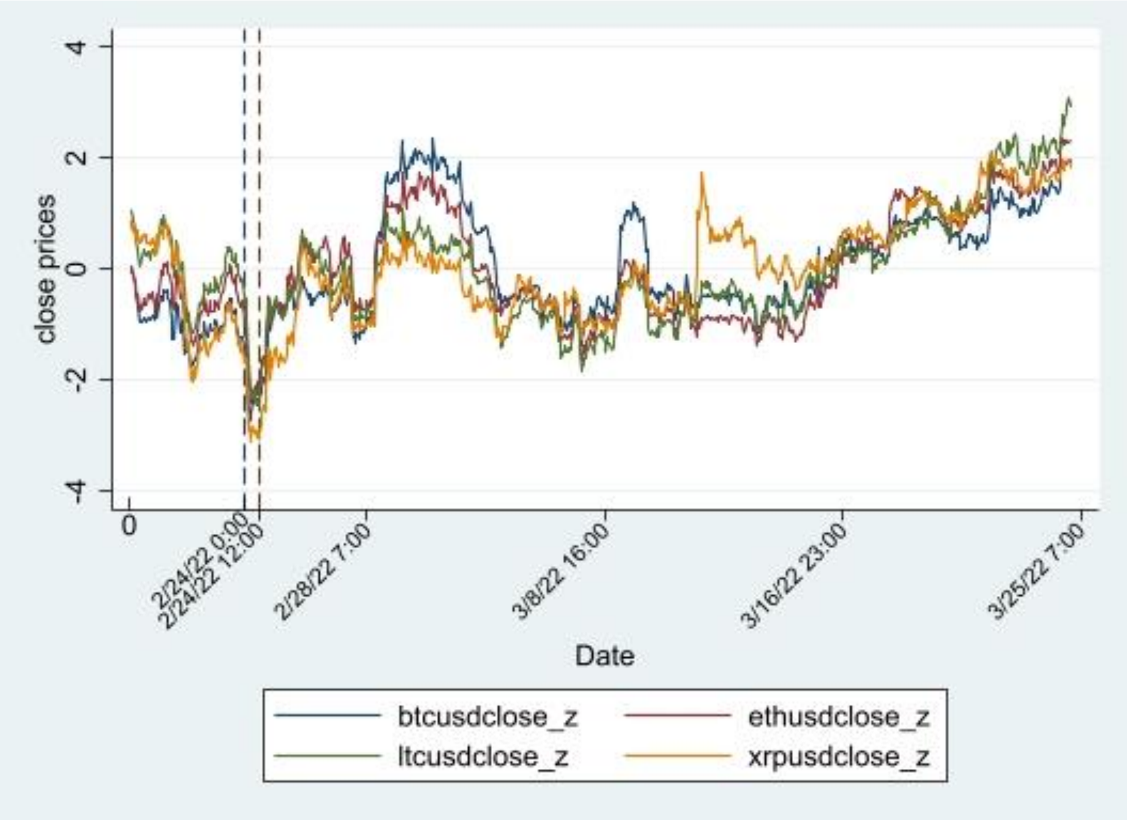

Figure 1. Time-series line plot of cryptocurrency close prices in the USA market between February 20th, 2022, and March 25th, 2022. The first vertical line marks the date where prices began to decline, and the second line shows when prices began to return to previous levels.

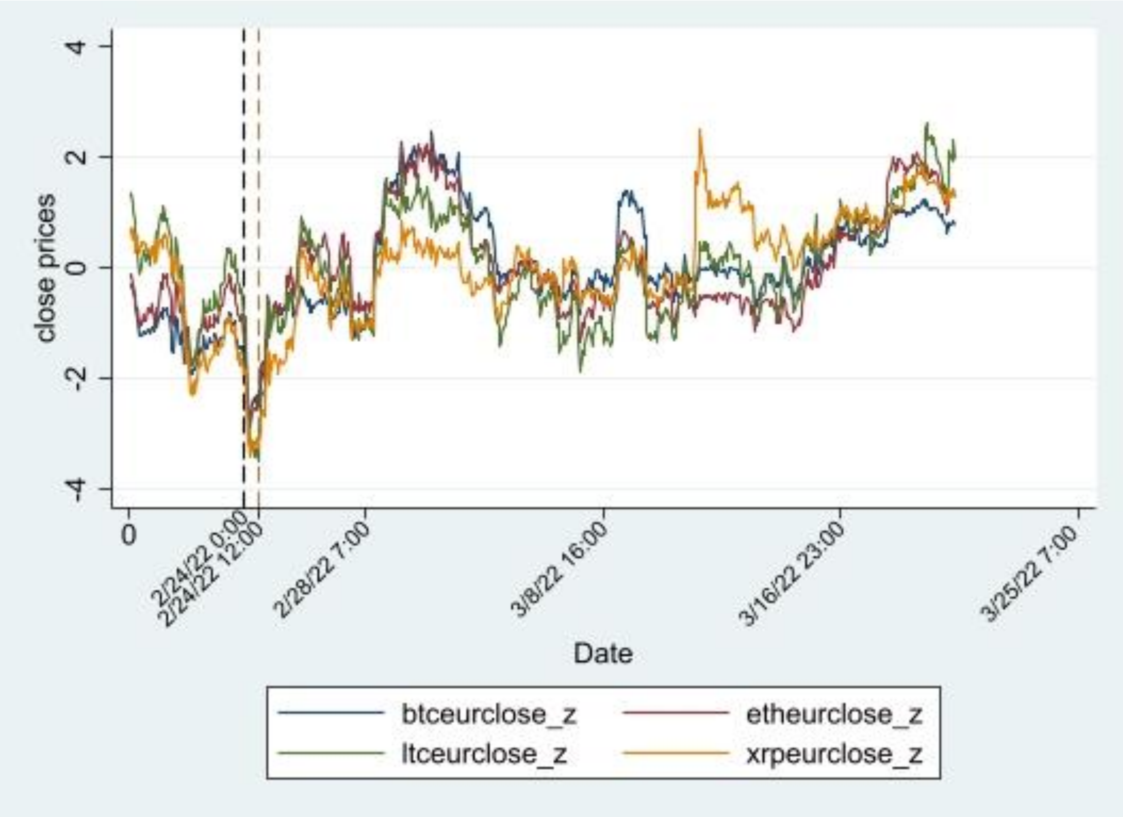

Figure 2. Time-series line plot of cryptocurrency close prices in the Europe market between February 20th, 2022, to March 20th, 2022.

Both markets show that the emergence of the Russia-Ukraine War caused a significant drop in close prices in all four cryptocurrencies on February 24th, 2022. The communication of the news through various channels, such as social media and news outlets, likely caused a panic among investors, especially those in Europe, where a slightly more pronounced decline in close prices is observed than in the USA. This is to be expected, considering that the war is taking place in eastern Europe. However, prices return to their normal values after several hours, indicating that the war causes short-run volatility spikes, but has little to no effect on long-term volatility. To address this, I remove the entire day of the 24th of February 2022. The war introduces complexities into our comparison of herding between the PP and the PPP, namely its short-run impact that may influence average herding during the PPP. The PPP is ideally meant to be compared with the PP as a period without any significant exogenous events that could cause an increase in volatility on the same level as the pandemic. As a result, I will remove the war's impact to reduce complexity and improve interpretability.

3.2 Data Organization

Our data is organized horizontally. I have one dataset for each period and each region:

1. PP (USA)
2. PPP (USA)
3. PP (Europe)
4. PPP (Europe)
5. Forecasts (USA)
6. Forecasts (Europe)

3.3 Cryptocurrency Selection

Cryptocurrencies are a relatively monopolized market at present, with only a few currencies dominating more than 70% of the market share. As a result, the selection of currencies aims to capture as much of the market as possible to maximize the statistical power of the analysis.

Bitcoin (BTC):

Bitcoin is the first and largest cryptocurrency, with a market cap of approximately 1.3 trillion USD. It dominated the market in 2019 and continues to do so, with a market share of approximately 49-54% during the PP and currently sits at 49% as of March 15th, 2024[20]. It is the main reason for the success of cryptocurrencies due to the level of investment from both retail and institutional investors. Nearly all cryptocurrencies fluctuate due to changes or events in the Bitcoin market due to the dominance it holds[21].

Ethereum (ETH):

Ethereum is based around a decentralized platform that enables any developer to build and deploy "smart" contracts as opposed to numerical currency volumes. Smart contracts are self-executing contracts with the terms of the agreement directly written into code. Ethereum was also the first cryptocurrency to introduce programmable blockchain. This allows for a wide range of use cases besides transactions such as tokenization[22]. Ethereum's market cap currently lies at 386 billion USD and has a market share of approximately 16%.

Litecoin (LTC):

Litecoin is often considered the silver to Bitcoin's gold, as its volatility patterns are mostly synonymous. It was designed to complement Bitcoin with faster transaction confirmation times and a different hashing algorithm (Scrypt). Litecoin aims to provide a more efficient

---

[20] Quarterly market share of selected cryptocurrencies, based on market cap 2013-2023 (https://www.statista.com/statistics/730782/cryptocurrencies-market-capitalization, accessed 30 March 2024)

[21] Adrian, Tobias, Iyer, Tara, and Qureshi, Mahvash S., 'Crypto Prices Move More in Sync with Stocks, Posing New Risks' *The International Monetary Fund* 2022, Available at: https://shorturl.at/abiyY, accessed 21 March 2024

[22] Ge Huang, Vicky, 'BlackRock Launches First Tokenized Fund on Ethereum Blockchain' *The Wall Street Journal* 2024, Available at: https://shorturl.at/bZ235, accessed 19 March 2024

and cost-effective means of transacting digital currency[23]. Its market cap is 6.3 billion and the market share is 0.25%.

Ripple (XRP):

The last cryptocurrency being examined in this study is Ripple. Ripple is both a payment protocol and a cryptocurrency token (XRP) developed by Ripple Labs. Unlike Bitcoin and Ethereum, Ripple is not based on blockchain technology but utilizes a distributed consensus ledger through a network of validating servers. Ripple aims to facilitate fast and low-cost cross-border payments for financial institutions[24]. Additionally, it is one of the only cryptocurrencies to face significant regulatory and legal challenges, largely due to the SEC lawsuit against Ripple regarding the unregistered selling of 1.3 billion in Ripple tokens[25]. The market cap is 33.3 billion and the market share is approximately 1.23%.

3.8 Variables:

Cross-Sectional Absolute Dispersion (CSAD):

The dependent variable that will be used for measuring herding is the cross-sectional absolute dispersion measure (CSAD).

---

[23] Napoletano, E., 'What is Litecoin? How Does it Work?', *Forbes* 2022, Available at: https://shorturl.at/gDMN9, accessed 30 March 2024.
[24] Naipal, Karan, 'Swift vs Ripple: 6 key areas of comparison for international business payments' *Medium* 2023, Available at: https://shorturl.at/oABE5, accessed 24 March 2024

[25] Godoy, Jody, 'US SEC Seeks $2 Billion from Ripple Labs, Chief Legal Officer Says' *Reuters* 2024, Available at: https://shorturl.at/hpuV3, accessed 29 February 2024

$$CSAD = \frac{1}{N}\sum_{i=1}^{N} = \frac{1}{N} \mid R_i - \bar{R} \mid$$

Figure 3. Formula for Cross-Sectional Absolute Dispersion (CSAD)

CSAD is a quantitative measure first and foremost, giving it an objective quality that is crucial to modelling psychological behavior in a robust and empirical fashion. This study only employs the CSAD measure instead of including the cross-sectional standard deviation (CSSD) measure for several reasons. Firstly, CSAD represents the real deviation between the position of investors and the average position in the market and it can capture the magnitude of this deviation without regards to direction. CSSD cannot accomplish these things and is thus less intuitive. What's most important however, is that CSAD is less sensitive to outliers and volatile patterns, which are quite common in cryptocurrency data:

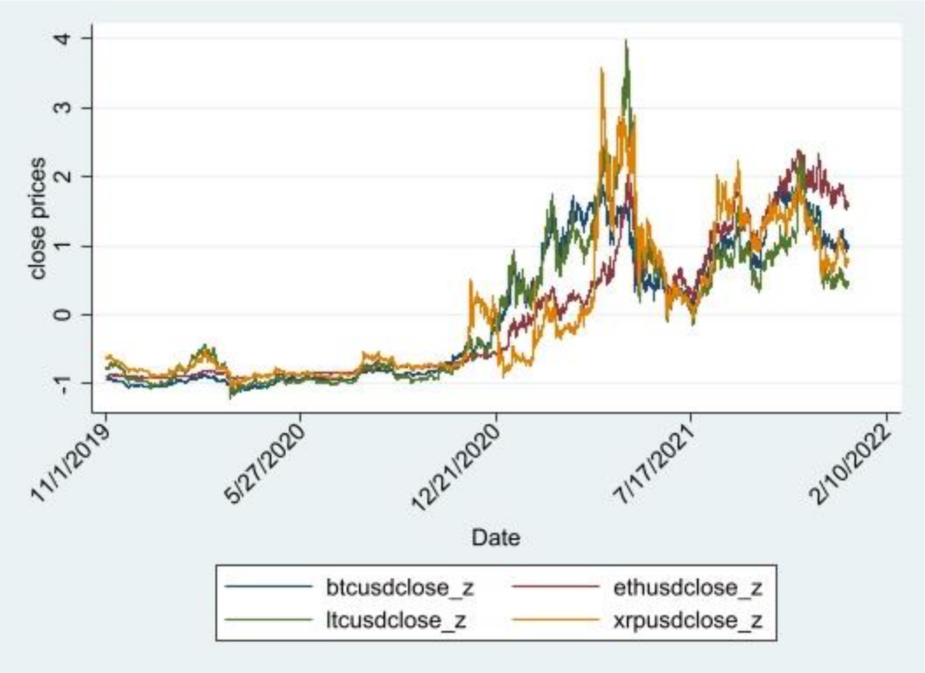

Figure 4. Time-series line plot for hourly close prices of cryptocurrency in the USA market during the PP.

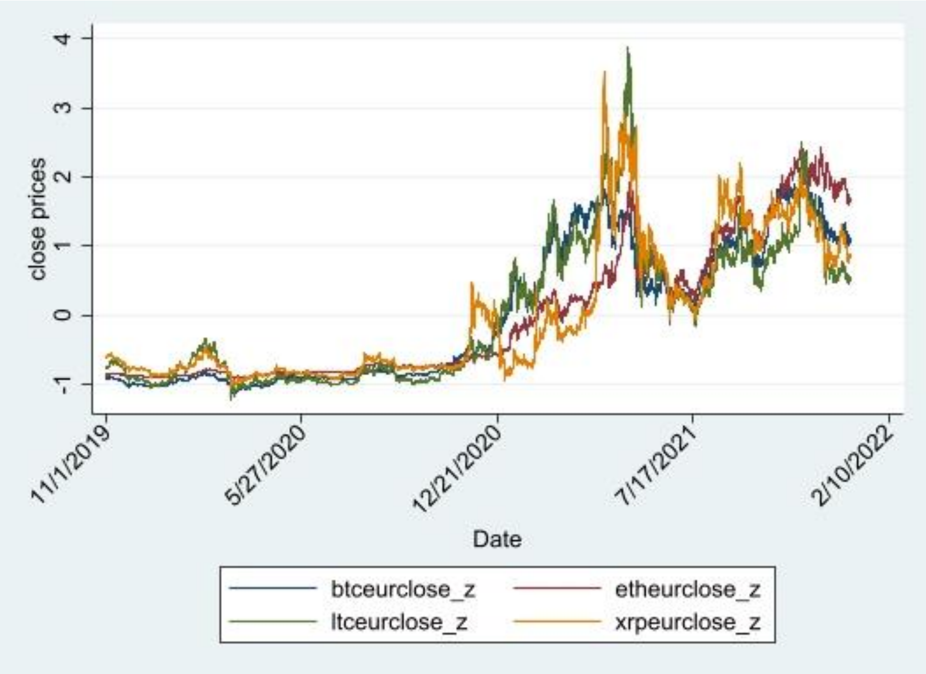

Figure 5. Time-series line plot for the European market during the PP.

The CSAD values were calculated on an aggregate basis, rather than an individual basis. This is safe to do so since the four cryptocurrencies have a relatively similar pattern of volatility. This makes it possible to create an aggregated measure without sacrificing granularity while also lowering the overall complexity of the time-series models. I have displayed this by examining different cross-sections of the data.

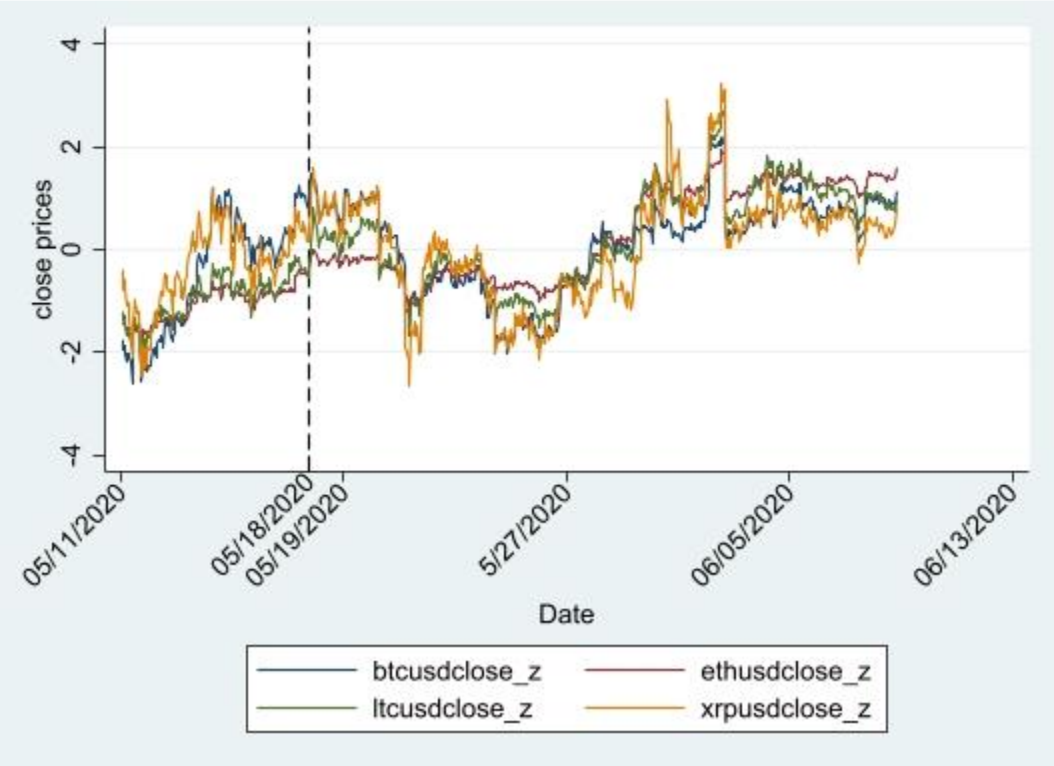

Figure 6. Time-series line plot for close prices of each cryptocurrency in the USA market between May 11th, 2020, and June 8th, 2020.

This event window is for the Bitcoin halving that took place on May 11th, 2020. The variables were normalized using z-scores for the sake of appropriate scaling. One can observe that there are initial spikes after May 11th up until May 18th because of the bitcoin halving. This shows how the currencies follow similar trends without major deviations.

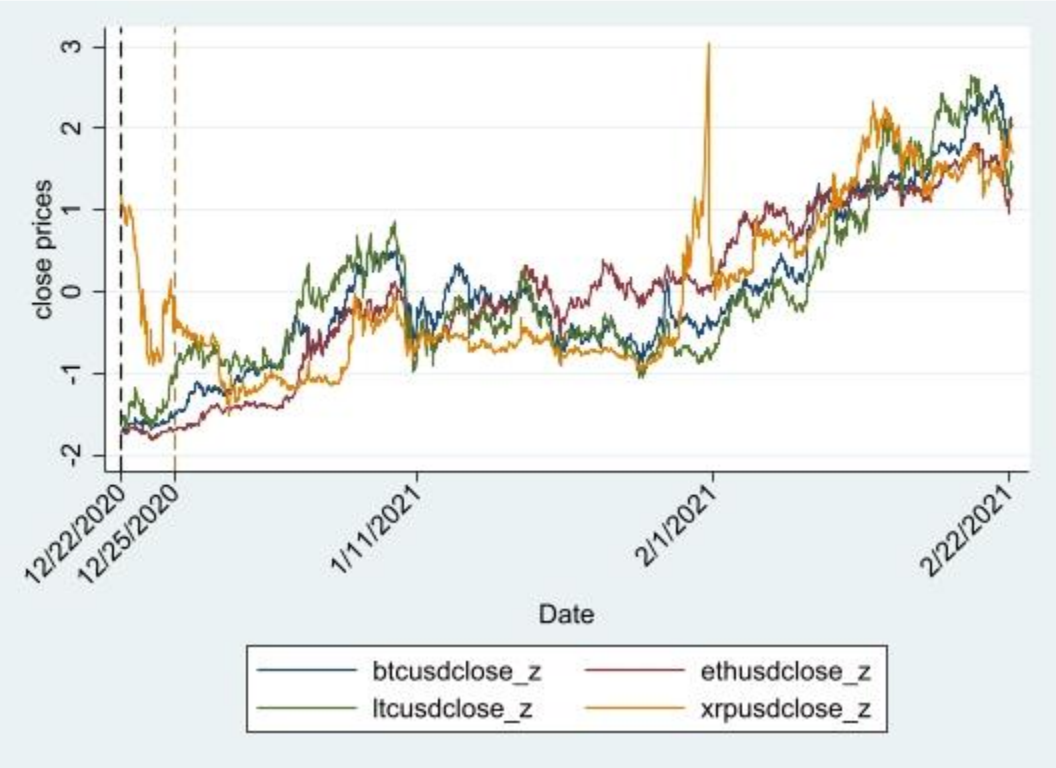

Figure 7. Time-series line plot for close prices in the USA market between December 22nd, 2020, and February 22nd, 2021.

This event also shows some short-lived spikes that do not cause major issues in the larger dataset's granularity. This is the largest deviation from the other cryptocurrencies in the entire dataset, and this is because this event is the SEC lawsuit filed against Ripple (XRP) in December of 2020[26]. The lawsuit plummeted its closing price between December 22nd and December 25th. Afterwards, it returns to relatively similar patterns to other currencies aside from one large spike around 2/1/2021.

[26] Godoy, 'SEC' *Forbes* 2024

Covid-19 cases:

I chose to incorporate a pandemic variable directly into the analysis as it reduces complexity. However, since COVID-19 cases are collected daily rather than hourly, I stretched the variable by 24 hours so that it fits the dataset's structure. A binary dummy variable was considered, but it is preferred to preserve temporal precision and granularity for the estimation model that would be used for forecasting later in the study.

Non-pandemic related events variable:

The other exogenous variable is our binary event variable, which shows a positive or negative value depending on whether certain dates fall within different event windows. It was decided to organize this variable on an aggregate basis, sacrificing granularity because it was shown by figures 6 and 7 that each cryptocurrency follows nearly the same price trends when each of the events occurs.

3.9 Descriptive Data Analysis

As stated earlier, using a frame of reference aids in identifying fluctuations in herding. I calculated the mean CSAD values for the PP and the PPP, and obtained the following results:

Period	**USA**	**Europe**	**Observations (USA and Europe)**	**Std. Dev.**	**Std. Dev.**
PP	0.0032	0.0031	19032	0.0035	0.0038
PPP	0.0023	0.0023	17545	0.0023	0.0023

Figure 8. Table for mean CSAD values

It can be seen here that the mean CSAD values for both regions are nearly the same and there is a slight decrease between the PP and the PPP. This implies that herding increased during the PPP, which is an interesting observation, since according to our hypotheses it is expected that herding increased during the PP in comparison to the PPP. Nonetheless, the difference is small, which may indicate potential market resilience in the face of exogenous events. The following time-series line plots also show a relatively weak positive correlation with CSAD values and daily cases.

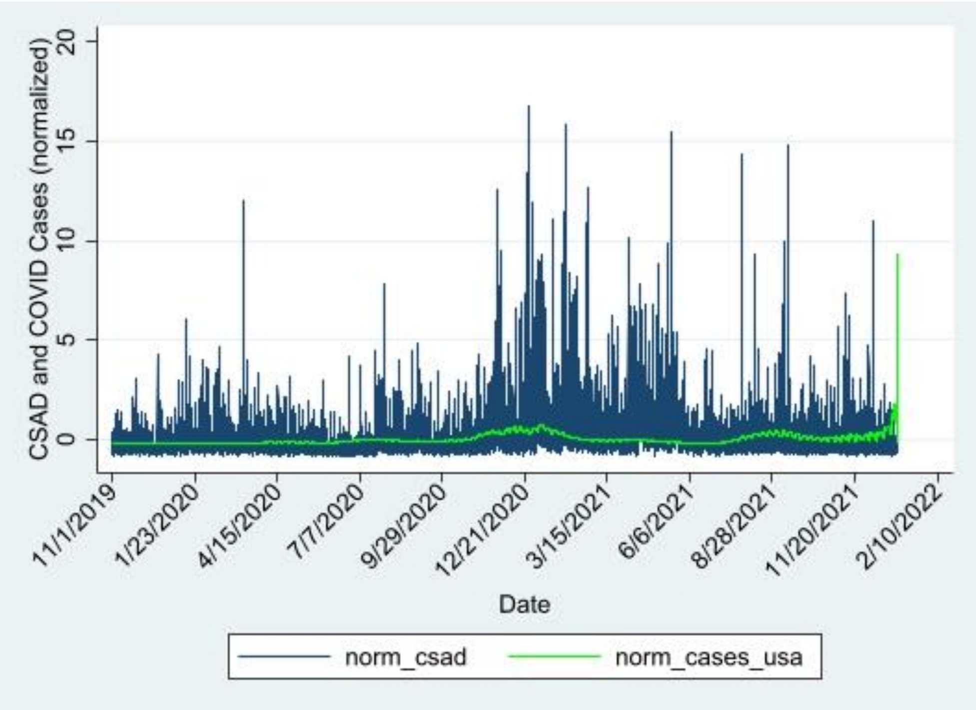

Figure 9. Time-series line plot for CSAD and Covid cases data (USA) during the PP

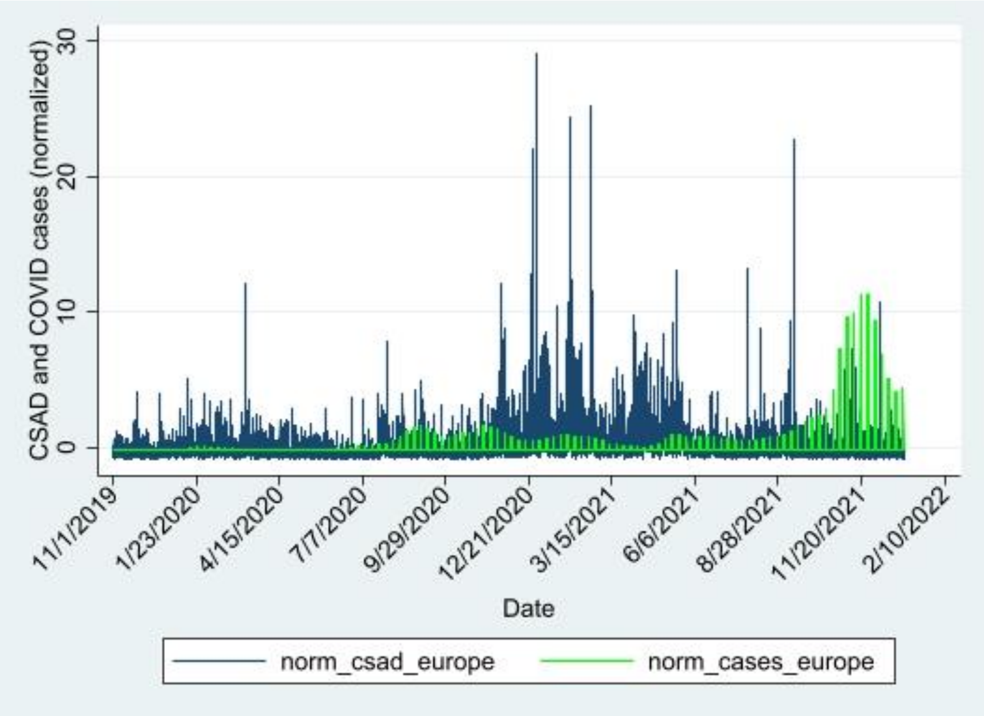

Figure 10. Time-series line plot for CSAD and Covid cases data (Europe) during the PP

From the time-series line plots there appears to be a weak positive correlation between daily covid cases and CSAD values, or rather that an increase in covid cases leads to a decrease in herding. This is because an increase in CSAD means that the cross-sectional dispersion increases as well, which implies lower price clustering and thus less herding. This is an interesting find because it further demonstrates resilience to shocks. Corbet et al. have shown in their study, ‘Any Port in the Storm: Cryptocurrency Safe-havens during the COVID-19 Pandemic’, that the resilience of cryptocurrency markets is not unprecedented[27]. However, to

[27] Corbet, Shaen, et al., ‘Any Port in the Storm: Cryptocurrency Safe-havens during the COVID-19 Pandemic’ *Economics Letters*, 194 2020, 1

concretely determine whether this is the case for the data, it is necessary to consult time-series models.

3.10 Model Estimation and Specification

To determine model fit, I consult the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). There are also several considerations and assumptions to be addressed. The important issue to address for this analysis is serial correlation or autocorrelation. Autocorrelation measures the degree to which the current value in a time-series dataset is influenced by its past values. Autocorrelation violates the assumption of independence among values in a time-series dataset, which can heavily bias parameter estimates and cause incorrect inferences and model misspecification. As a result, addressing autocorrelation becomes an important component of the models' specification.

I have determined several contenders for our estimation process, which involves an iterative approach by comparing each model's respective AIC and BIC values and their ability to handle autocorrelation. The lower the AIC/BIC value and the smoother the autocorrelation function plot, the better the model's fit. The following models chosen include:

1. Markov-Switching Models

A Markov-Switching model is a state-space and regime-switching model, which changes the relationship between the dependent variable and independent variables depending on the current regime. I identified two regimes in the data, an up-market condition, and a down-market condition. The equation below describes the base equation:

$$y_t = \beta_{0,r} + \beta_{1,r} x_{1,t} + \dots + \beta_{k,r} x_{k,t} + \varepsilon_t$$

Where y_t is the dependent variable at time t, x_1, x_2 up until $x_{k,t}$ are independent variables at time t, $\beta_{0,r}$, $\beta_{1,r}$, up until $\beta_{k,r}$ are the coefficients associated with the independent variables in regime r, and ε_t is the error term at time t[28].

4. Vector Auto-regression (VAR):

A Vector-Autoregression is a multivariate time-series model capable of capturing dynamic interdependencies among multiple variables over time. The model is specified using the following base equation:

$$Y_t = \alpha + \sum_{i=1}^{p} \theta_i Y_{t-i} + \beta_1 X_{t,1} + \beta_2 X_{t,2} + \varepsilon_t$$

Where Y_t is the dependent variable (CSAD) at time t, $X_{t,1}$ is the first exogenous variable (COVID cases), $X_{t,2}$, is the second exogenous variable (binary variable for non-pandemic events), and θ_i is the coefficient of the lagged values of the dependent variable[29]. To determine the correct lag-order, one can take several steps. Initially, one would consider prior knowledge of the dataset to set the model parameters. However, due to the size and complexity of the dataset, an iterative approach is more efficient. I used the *varsoc* command

---

[28] Bianchi, Francesco, 'Methods for Measuring Expectations and Uncertainty in Markov-Switching Models', *Journal of Econometrics* 190 2016, 81
[29] Abdulnasser, Hatemi-J, 'Multivariate Tests for Autocorrelation in the Stable and Unstable VAR Models', *Economic Modelling* 21 2004, 661-683

in stata to determine the optimal lag and found that the model accounting for the past 80 lags has the best fit. This produces the following equation[30]:

$$Y_t = c + \sum_{i=1}^{80} \theta_i Y_{t-i} + 0.0297212X_{t,1} + 0.0593143X_{t,2} + \varepsilon_t$$

Auto-regressive Integrated Moving Average with Seasonality (SARIMA):

A Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an extension of the Autoregressive Integrated Moving Average (ARIMA) model that incorporates seasonal components. It's designed to handle time series data that exhibit seasonality, meaning that the data exhibit patterns that repeat over fixed periods of time. The methods by which I determine the Autoregressive order (AR) and the Moving Average order (MA) also depend on an iterative approach. This is coupled with autocorrelation function plots (ACF) for the AR component and partial autocorrelation plots (PACF) for the MA component. I experimented with several ARIMA and SARIMA models, some of which are represented by the equations below:

Specification (1,0,1):

$$CSAD_{\text{market},t} = \beta_0 + \beta_1 cases_{\text{market},t} + \beta_2 event_{\text{market},t} + \phi_1 CSAD_{market,t-1}$$
$$+ \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

[30] The coefficients for the cases and non-pandemic related events variable can be found in figure 21.

Specification (1,0,1,12):

$$CSAD_{market,t} = \beta_0 + \beta_1 event_{market,t} + \beta_2 cases_{market,t} + \phi_1 CSAD_{market,t-1}$$

$$+ \theta_1 \varepsilon_{t-1} + \Phi_2 CSAD_{market,t-12} + \Theta_1 \varepsilon_{t-12} + \varepsilon_t$$

Where $CSAD_{market,t}$ is the dependent variable (CSAD) at time *t*, β_0 is the intercept β_1, and β_2 are the coefficients of the independent variables, $event_{market,t}$ is the binary variable for non-pandemic exogenous events, and $cases_{market,t}$ variable is for covid cases, ϕ_1 , ϕ_2 , and Θ_1 are the autoregressive terms, θ_1 and θ_2 are the moving average terms, and ε_{t-12} is the error term representing the difference between the observed value and the value predicted by the model at time *t*.

AIC and BIC Results:

The table below shows the different models and their specifications, as well as their AIC/BIC values:

Model and Specification	**Akaike Information Criterion (AIC)**	**Akaike Information Criterion (AIC)**	**Bayesian Information Criterion (BIC)**	**Bayesian Information Criterion (BIC)**
	USA	*Europe*	*USA*	*Europe*
Markov-Switching Model (2 Regimes)	47136.81	49167.28	47191.79	49222.25
Vector Autoregression (VAR) lags 1/4	51228.31	52627.69	51283.29	52682.66
Vector Autoregression (VAR) lags 1/50	50117.00	51476.55	50533.09	51892.66
Vector Autoregression (VAR) lags 1/80	50028.80	51372.98	50680.29	52024.50
Autoregressive Integrated Moving Average (ARIMA) (1,0,1)	50457.41	51885.11	50504.53	51932.23

Autoregressive Integrated Moving Average (ARIMA) (2,1,2)	50371.85	51576.52	50434.67	51631.49
Seasonal Autoregressive Integrated Moving Average (SARIMA) (1,0,1,12)	53426.52	54776.46	53473.64	54823.58

Figure 11. Table containing each model and its specification.

As can be seen in the table, the Markov-Switching Model has the lowest AIC and BIC values. However, there are still some concerns regarding autocorrelation. The Markov-Switching model has difficulties with adjusting to the many patterns in the data. This is likely because the volatile nature of the data makes it nearly impossible to effectively account for all the trends (the data being hourly likely contributes to this). However, the VAR model does this with great effect, as it is more flexible than the M-switch model. Although it has slightly higher AIC and BIC values, and thus a worse fit, there is no significant autocorrelation at the 95% level according to the ACF and PACF plots shown below:

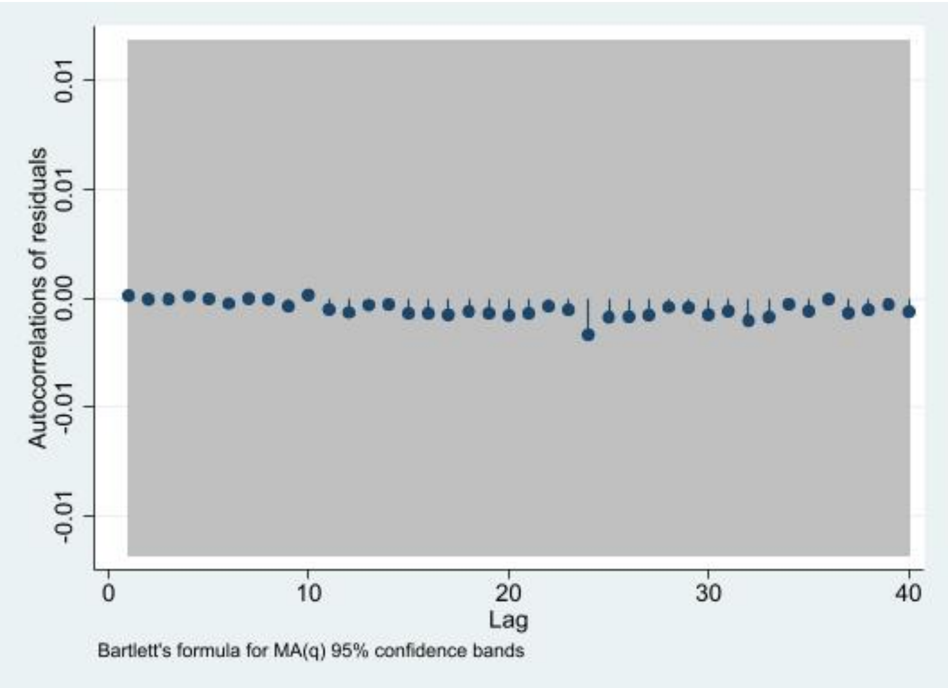

Figure 12. ACF Plot for the USA Vector Autoregression model (lags 1/80)

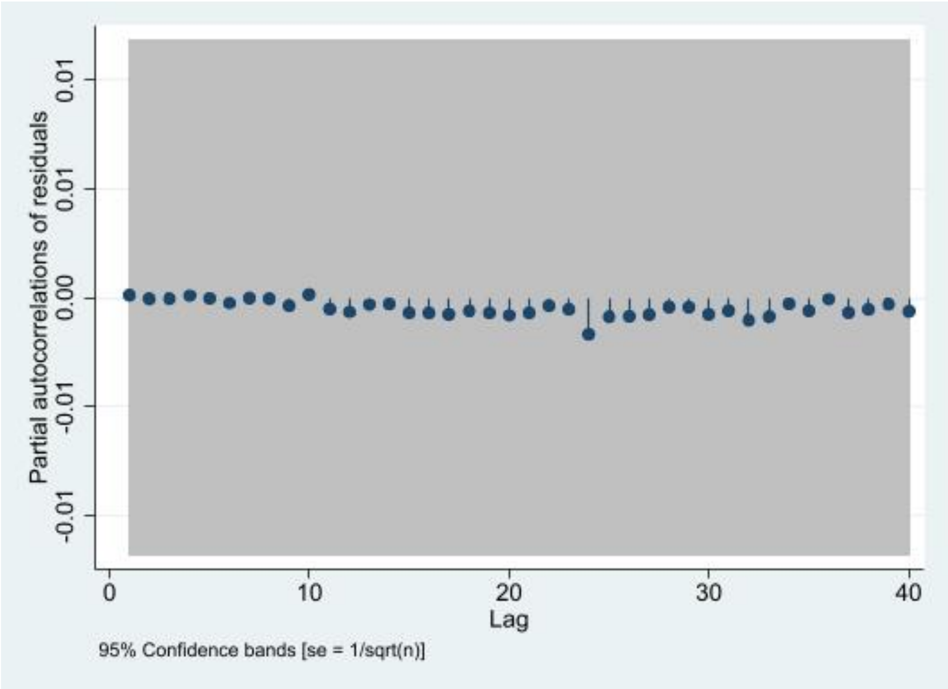

Figure 13. PACF plot for the USA Vector Autoregression model (lags 1/80)

I also conducted a Ljung-Box test to get a formal understanding of the patterns in the model and calculated a statistic of 7.65 and a p-value of 1.000 for the USA model and 0.5156 and 1.000 for the Europe model, neither of which are significant at the 95% level in the slightest. The ACF and PACF plots for the Europe model are also below:

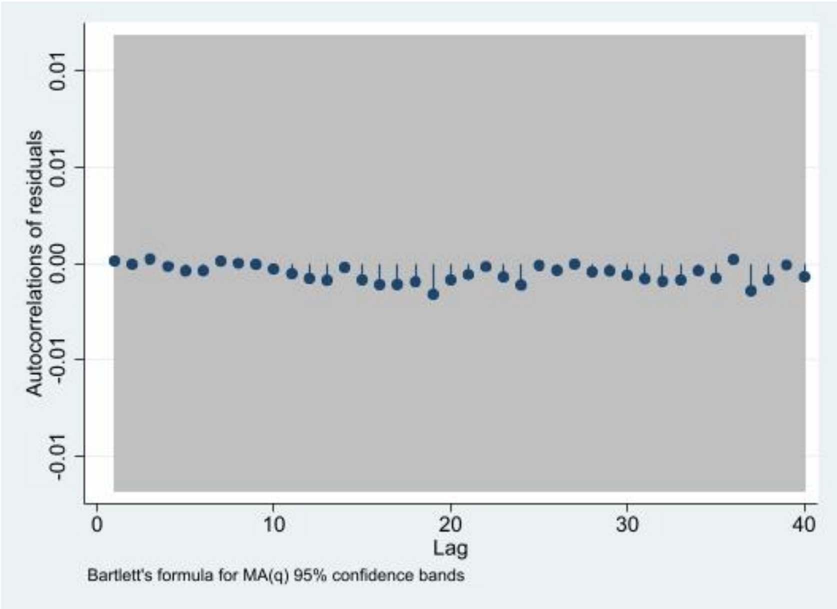

Figure 14. ACF Plot for Europe Vector Autoregression Model (lags 1/80).

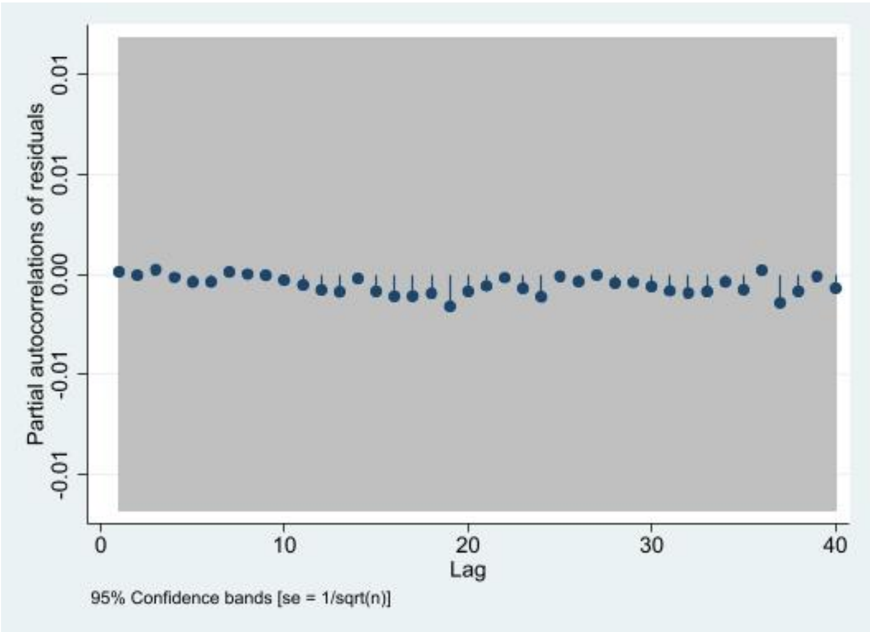

Figure 15. PACF plot for Vector Autoregression model (lags 1/80).

Now these results can be compared to that of the Markov-Switching model:

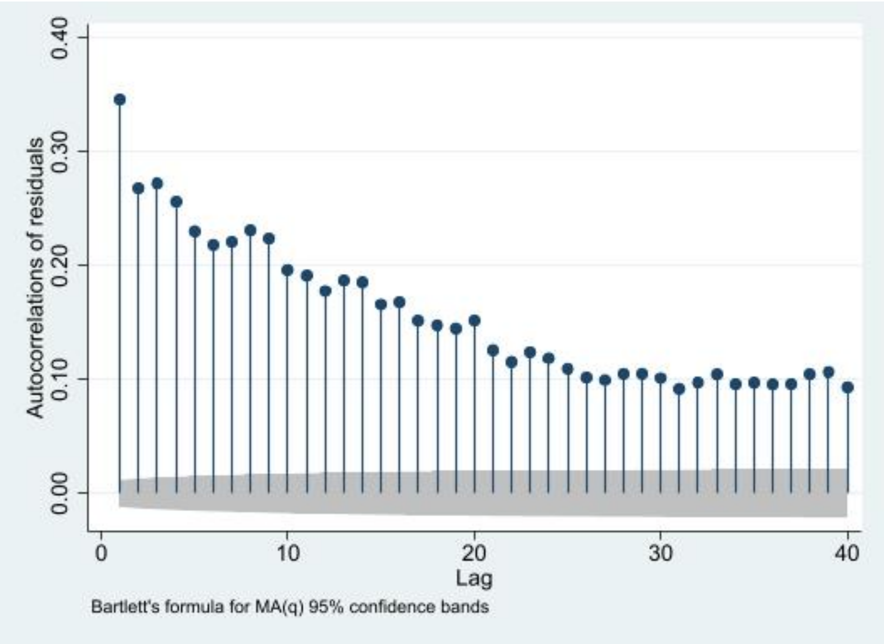

Figure 16. ACF plot for the USA Markov-Switching model residuals

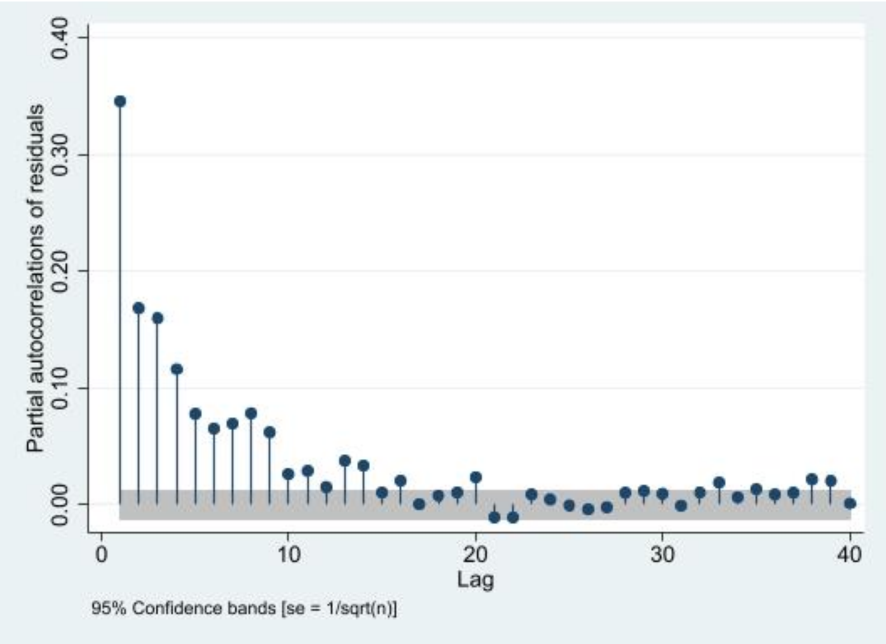

Figure 17. PACF plot for the USA Markov-Switching model residuals

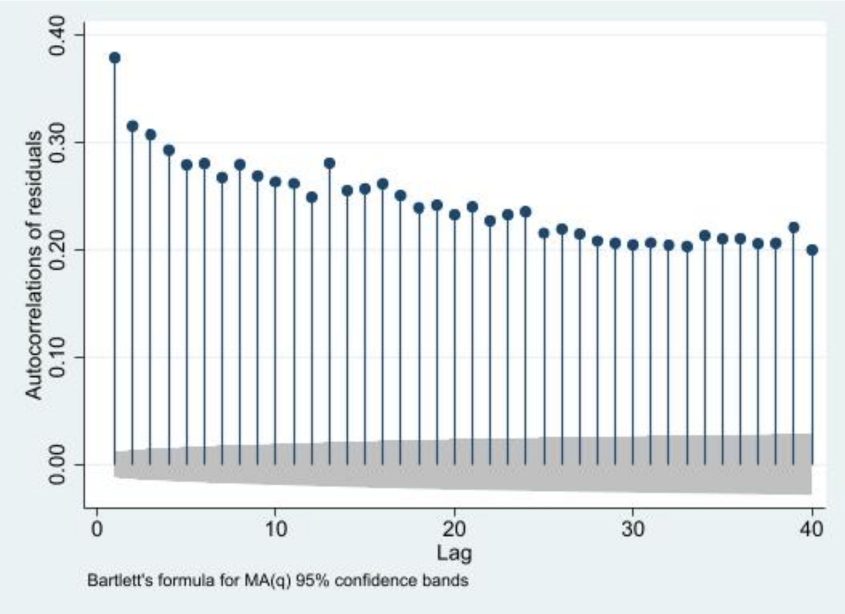

Figure 18. ACF Plot for Europe Markov-Switching model residuals

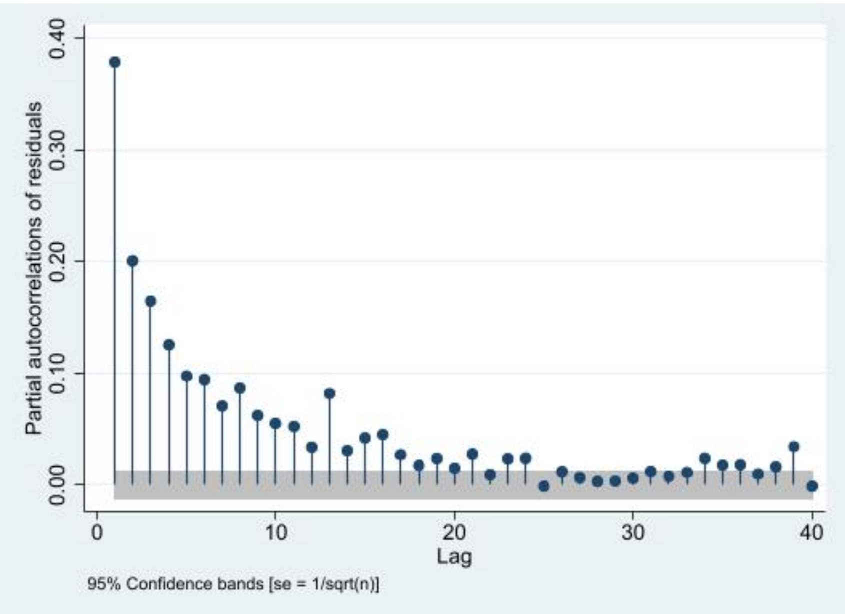

Figure 19. PACF plot for Europe Markov-Switching model residuals

From the plots, there is a highly significant level of autocorrelation at the 95% level across multiple lags. The Ljung-Box test statistic and p-values for the USA model are 31262.69 and 0.0000. The results for the Europe model are 67762.37 and 0.0000. As a result, I chose the model with the best fit and ability to address autocorrelation, the VAR model with 80 lags. The model with 50 lags was considered, but it also does not handle autocorrelation as well as the 80-lag model.

3.12 Forecasting Model

The forecasting models are chosen by consulting the AIC and BIC, but I also utilized other measures to determine the accuracy of the model, namely, the mean absolute error (MAE), the mean squared error (MSE), and the root mean squared error (RMSE). These measures allow for the measurement of the average deviation of each forecasted value from its real counterpart.

When predicting future values, I created a training and test set. The training set is what is used for the estimation model, which is used to subsequently forecast new data. The test set is used to calculate the various measures described above. They are displayed in the table below:

Model	**MAE**	**MSE**	**RMSE**
	USA		
Markov-Switching Model (2 Regimes)	0.0859	0.0192	0.5538
VAR (1/80)	0.0708	0.6895	0.6290
SARIMA (1,0,1,12)	0.1185	0.2275	0.6291
	Europe		
Markov-Switching Model	0.0552	0.3171	0.5538
VAR (1/80)	0.0012	0.3171	0.6291
SARIMA (1,0,1,12)	0.0446	0.3014	0.6291

Figure 20. Table containing MAE, MSE, and RMSE values for each estimation model

These metrics are calculated using unseen data, which provides the best possible diagnostic check for the accuracy of the forecasting model. Our training set includes data from the PP.

The forecasted data falls in the range of January 30th, 2022, to March 23rd, 2025. Our test data, or the data used to calculate the accuracy metrics, falls within the same range but stops on January 30th, 2024, the same range as the PPP. The predicted values and the test set are then used to calculate the metrics. The table shows that the m-switch model and the SARIMA model have the highest level of accuracy for forecasting. However, they do not address autocorrelation well, and thus lead to various inaccuracies with the measures above. This is exemplified by how changing the number of lags accounted for in a VAR model can affect these values:

Model	**MAE**	**MSE**	**RMSE**
VAR (lags 1/4)	0.0312	0.1253	0.3539
VAR (lags 1/80)	0.0708	0.6895	0.6290

Figure 21. Forecasting accuracy metrics for different VAR model specifications.

Although the metrics are more desirable in the model with lower lag specifications, the first VAR model addresses autocorrelation poorly in comparison to the second model, which makes its estimates unreliable. This leads us to choose the 80-lag VAR model for our estimation process.

Results and Analysis:

4.1 Estimation Results

Now that I have gone through the procedure of choosing and specifying our identification and forecasting model, we follow these steps:

1. Utilize the identification model (VAR) to determine whether there is a significant correlation between covid cases and CSAD values.
2. Estimate a model that takes the relationship of the previous model into account when predicting new values.

I only consider significant lags in the table below for the sake of conciseness:

Lag/Variable	Coefficient	Std. Error	p-value
	CSAD		
L1.	.2036495	.0072643	0.000
L2.	.0668588	.0074125	0.000
L3.	.076404	.0074273	0.000
L4.	.0649849	.0074476	0.000
L5.	.0422524	.0074626	0.000
L6.	.0309559	.0074688	0.000
L7.	.0381514	.0074722	0.000
L8.	.0282066	.0074763	0.000
L9.	.0320845	.007478	0.000
L11.	.02961	.0074813	0.000
L13.	.0443491	.0074849	0.000
L15.	.015835	.0074924	0.035
L16.	.0275653	.0074929	0.000
L19.	-.0146841	.0074935	0.050
L20.	.0196034	.0074942	0.009
L24.	.0346924	.0074948	0.000
L27.	-.0157807	.0074959	0.035
L28.	.0154383	.0074963	0.039

L35.	.0255338	.0074978	0.001
L38.	.025208	.0074972	0.001
L39.	.0196346	.0074993	0.009
L43.	.018336	.0074972	0.014
L44.	-.0205288	.0074983	0.006
L55.	-.0230755	.0074957	0.002
L56.	.0220524	.0074972	0.003
L59.	.017865	.0074941	0.017
L64.	.0190622	.007494	0.011
L72.	.0171385	.0074775	0.022
L73.	-.0175363	.0074758	0.019
L78.	.0169193	.0074267	0.023
L79.	-.0169236	.0074118	0.022
L80.	.0168049	.0072632	0.021
	Exogenous Variables		
Non-pandemic Related Events	.0297212	.0158099	0.060
Covid Cases	.0593143	.0269475	0.028

Figure 22. Table containing the results of the VAR model (1/80 Lags) for the USA market.

Lag/Variable	Coefficient	Std. Error	p-value
	CSAD		
L1.	.2180741	.0072637	0.000
L2.	.0709916	.0074344	0.000
L3.	.0669327	.0074493	0.000
L4.	.0476973	.0074645	0.000
L5.	.0321768	.0074726	0.000
L6.	.0408218	.0074761	0.000
L7.	.0187217	.0074818	0.012
L8.	.0468854	.0074829	0.000
L9.	.0264636	.0074897	0.000
L10.	.0210379	.0074922	0.005
L11.	.0181728	.0074937	0.015
L13.	.0638715	.0074937	0.000
L16.	.0321405	.0075077	0.000
L21.	.0178193	.0075067	0.018
L24.	.022477	.0075075	0.003
L25.	-.0151478	.0075092	0.044
L34.	.0166813	.0075071	0.026
L39.	.0350074	.0075073	0.000
L53.	-.0157813	.007507	0.036
L54.	.0165485	.0075079	0.028
L55.	-.0174775	.0075088	0.020
L60.	-.0179209	.0075066	0.017
L61.	.0181906	.0075075	0.015
L62.	-.0148795	.0075083	0.048
L63.	.0170053	.0075091	0.024

L64.	.0205531	.0075095	0.006
L68.	-.0147306	.0074932	0.049
L72.	.0166429	.0074893	0.026
L77.	.0156317	.0074638	0.036
L78.	.0287179	.0074488	0.000
	Exogenous Variables		
Non-pandemic Related Events	0.0347937	0.016330	0.033
Covid Cases	-0.0072682	0.006716	0.279

Figure 23. Results for the VAR model (1/80 Lags) for the Europe market

There are approximately 32-33 significant lags in each market's model. To briefly explain, a lag is a number that indicates how far in the past a value exists in a dataset. For example, if we are interested in looking for the second lag, that would be one unit of observation before the current data point. As a result, if a lag is significant, it means that the past value (the lagged value) has a significant impact on the current value. From these significant lags, I find there is a significant proportion of previous values that influence the current value of Y_t.

Nonetheless, what I am more interested in looking at is the significance of the variable for covid cases. I calculate a p-value of 0.028 for the cases variable in the USA market and a p-value of 0.279 in the Europe market. As a result, the USA cases variable is significant at the 5% level while the Europe cases variable is not. The coefficients of the variable in the USA and European markets are 0.0593 and -0.0073, respectively. In the USA market, the coefficient implies that for every unit increase in the cases variable, the endogenous variable, which is CSAD, increases by 0.0593. This implies that herding decreases for every unit increase in cases. In contrast, the Europe market shows an increase in herding because CSAD decreases by -0.0073. However, since only the USA cases variable is significant at the 5% level, I only have sufficient evidence to claim that herding decreases in the USA market, but not enough to claim that it increases in the Europe market. Additionally, the event dummy variable is significant only in the Europe market and causes CSAD values to increase by approximately 0.03. This means that exogenous events unrelated to the pandemic may also reduce herding, which is an important observation for reasons that will be explained shortly. If we recall our first hypothesis, it states:

H_1: Herding behavior increases significantly in response to the uncertainty created by changes in daily COVID-19 cases.

If we consider the places where the cases variable was significant, i.e., the USA market, the dispersion increases by 0.0593, which indicates that herding decreases significantly in response to an increase in covid cases, which leaves H_1 unsatisfied.

I conclude that herding does not increase in response to an increase in covid cases, and in fact, shows no significant change aside from a small decrease in the USA market, but why is this the case? The absence of an increase in herding can be specifically attributed to Richard Thaler's idea of a nudge coupled with a new concept called psychological reinforcement. A nudge is a man-made environmental signpost that is designed to influence a person's decisions in a predictable manner without their knowledge. This can come in the form of many things we see every day, such as cashback rewards or an opt-out instead of opt-in healthcare system[31]. In the case of cryptocurrency markets, nudges exist to prevent an

[31] Thaler, Richard, and Sunstein, Cass R., *Nudge: Improving Decisions about Health, Wealth, and Happiness*, (New Haven 2008)

investor from making unwise investment decisions and to inform them of the dangers of heightened market volatility. Exchanges such as Bitstamp and Bitthumb are always trying to accomplish this[32]. This may come in the form of fine print, advertisements, as well as warnings that come up during transactions. This has an effect of preventing an investor from succumbing to biases such as herding. The other component, called psychological reinforcement, has the same effect but differs in its origins. This is because it is a natural component of the environment that influences a person's experiences or preferences. This may come in the form of market volatility in close prices or experiences of financial losses after certain investment decisions. What this and nudging does is change the investors' reference point. In prospect theory, a reference point is determined by a person's experiences, biases, and preferences, and they evaluate decisions and outcomes from this reference point[33]. As a result, I posit that the reason cryptocurrency markets are so resilient to exogenous events like the pandemic is due to the changing of the reference point. Cryptocurrency markets are more volatile and fluctuate more radically than other investment environments. As a result, over time, investors working in these markets naturally become desensitized to losses and periods of heightened volatility simply because it is such a common occurrence in cryptocurrency markets. As a result, they learn to anticipate exogenous shocks, such that when a black swan event like the pandemic occurs the impact is only in the short-run and investors are more likely to make independent, rational decisions because their experiences have raised their reference point of risk. This is corroborated by how only short-run volatility

[32] Understanding Crypto Investments: A Comprehensive Guide', *Bitstamp Learn,* 2023
[33] Kahnemann and Tversky, 'Decision Under Risk', 275

spikes were observed with a similar exogenous event, the Russia-Ukraine War. This change in an investor's reference point, caused by nudging and psychological reinforcement, may be the reason that no significant change or decrease in herding is observed.

As the models show, nudging and psychological reinforcement may prevent herding from increasing because investors are being conditioned to expect exogenous events. However, many studies, like the ones shown in the literature review, reveal that herding does increase during periods of heightened volatility, such as during the Asian Financial Crisis in 1997[34]. In contrast, some regions have instead displayed resilience to the pandemic, such the Korean market in Yarovaya et al.'s work. Nonetheless, even though there are contradictions in the behavior of herding during exogenous events, the data highlights that herding is relatively unaffected by the pandemic. This indicates resilience even though the reason for this is not entirely clear.

4.2 Forecasting Results

Now that the theoretical underpinnings of why herding has not increased during the PP have been discussed, we can proceed to the forecasting results. I utilize the same VAR model used previously to forecast the values. I forecast until March 23rd, 2025. The values are shown on a time-series line plot below:

[34] Jirasakuldech, B., 'Empirical Analysis', 1

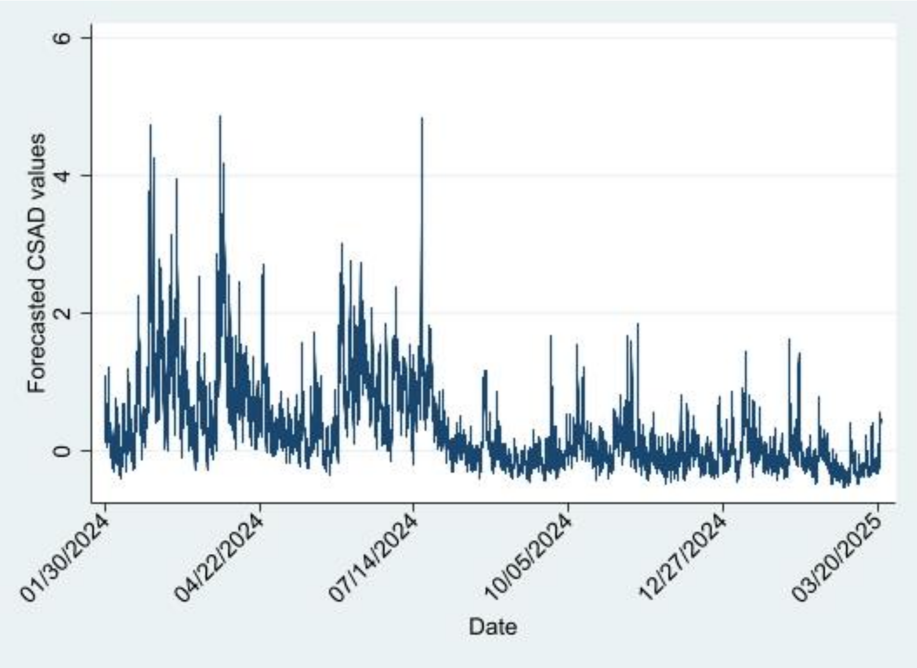

Figure 24. Forecasted USA CSAD values from January 30th, 2024, to March 23rd, 2025.

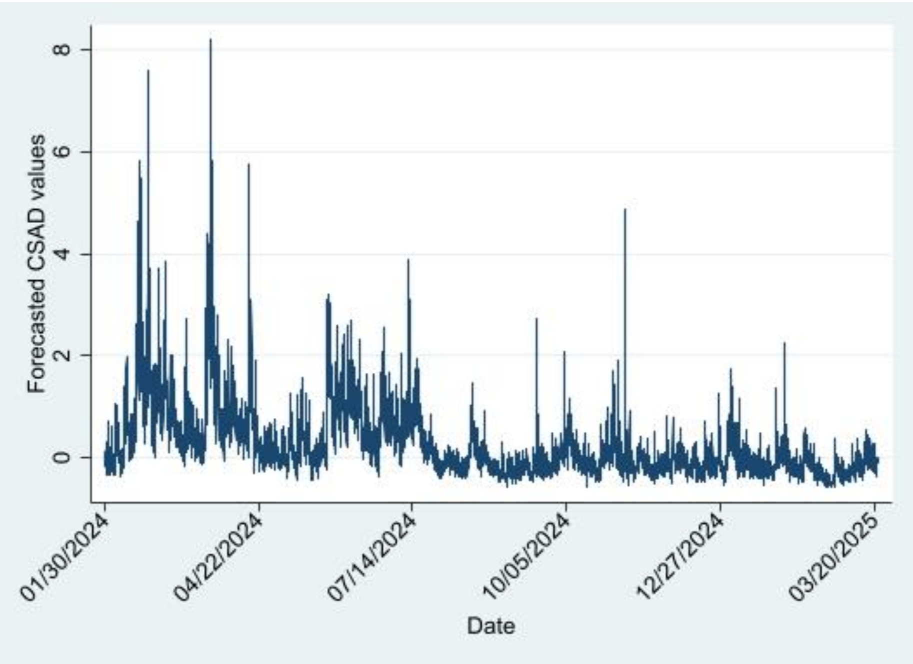

Figure 25. Forecasted Europe CSAD values from January 30th, 2024, to March 23rd, 2025.

As shown above, there are periods of volatility in the plot, and we can also compare the mean CSAD values in the PP and the forecasted period in the following table:

Region	CSAD (PP)	CSAD (Forecasted)	Pre-Pandemic Period
USA	0.0032	0.0033	0.0028
Europe	0.0031	0.0024	0.0030

Figure 26. Table showing mean CSAD values of the PP versus the forecasted period.

If we recall H_2, it states:

> H_2: The forecast period is expected to see a decrease in herding compared to the PP because of a lower level of uncertainty created by the absence of daily COVID-19 cases.

As shown by the data, herding is predicted to decrease slightly during the forecasted period for the USA market and increase for the Europe market. However, I find that if we analyze the period of 6 months prior to the pandemic's emergence, I get a mean CSAD value of 0.0028 and 0.0030 for each region. This implies that higher levels of herding were observed prior to the pandemic, which lines up with my claims about psychological reinforcement causing a reduction in herding. Despite this, a decrease is indeed observed in the USA market during the forecast period in comparison to PP, which fulfills H_2. This is not the case for Europe, where herding increases, leaving H_2 unsatisfied. Nonetheless, the results of comparing three different periods show that market resilience is a more plausible conclusion for herding' evolution than what H_2 predicts.

Conclusion and Limitations:

I conclude that cryptocurrency markets in the United States and Europe during the PP were relatively resilient to the pandemic, showing no signs of major deviations in herding patterns save from a small reduction in herding in the USA market. Although the reasons for this are difficult to discern, I consider an approach that builds upon existing literature to introduce a new concept of psychological reinforcement, which is a long-term, experiential component of the decision-making process that conditions an individual's standard or reference point of loss and risk aversion. The cryptocurrency market serves as a prime example of where such a bias may emerge, as it involves highly volatile and time-sensitive trends. Given my use of the existing theories on the topic, I can safely hypothesize that given what was elucidated from the models, a conditioning bias such as psychological reinforcement could exist and play a significant role in decision-making under risk. The resilience of the American and European cryptocurrencies markets nonetheless goes against what various studies like Kumar et al. and Yarovaya et al found, where an increase in herding was observed. However, our conclusions are supported by Bouri et al. and Corbet et al., who show how markets may be resilient and that herding may not fluctuate significantly in response to shocks but may increase in intensity. The discrepancy between the results of each study is not just because of a difference in motive and focus, but most likely because a concrete mathematical solution to a behavioral problem is difficult to ascertain. Behavioral biases generally do not follow the laws of mathematics, and thus there will always be uncertainties and contradictions. Consequently, it is more a matter of determining the most appropriate model, even if it is not perfect. That is what this study aims to contribute to.

Nonetheless, our model does not account for various exogenous variables that could shed light on the subject, namely covid deaths, which could bring some enlightenment to the discussion despite the complexities it would introduce into the models. Inflation factors, VIX, P/E, as well as market sentiment would be important factors to incorporate as well. The models themselves also had high AIC and BIC values, and perhaps a more complex M-switch model with an effective countermeasure for autocorrelation could provide more accurate results. VAR is a very flexible model, but it does assume linearity as well as homoskedasticity and stationarity. The lack of data for the Asian market also forces me to leave out a considerably large cryptocurrency market, which prevents any analysis of potential relationships that would overturn the conclusions made in this paper. The complexity of the models also had to be reduced due to data constraints as well as hardware capabilities. A significantly complex model, such as a Markov-switching model, would take a long time to get results from, and it would also overcomplicate the model beyond our analytical capabilities. Nonetheless, this paper accomplishes quite a great deal despite the constraints and provides a comprehensive analysis of herding during the pandemic.

Bibliography

Abdulnasser, Hatemi-J, ‘Multivariate Tests for Autocorrelation in the Stable and Unstable VAR Models’, *Economic Modelling* 21 2004

Adrian, Tobias, Iyer, Tara, and Qureshi, Mahvash S., ‘Crypto Prices Move More in Sync With Stocks, Posing New Risks’ *The International Monetary Fund* 2022, Available at: https://shorturl.at/abiyY, accessed 21 March 2024

Bianchi, Francesco, ‘Methods for Measuring Expectations and Uncertainty in Markov-Switching Models’, *Journal of Econometrics* 190 2016

Chang, C.-L., McAleer, M., Wang, Y.-A. ‘Herding Behaviour in Energy Stock Markets during the Global Financial Crisis, SARS, and Ongoing COVID-19.’ *Renewable and Sustainable Energy Reviews,* 134 (2000)

Marco Cipriani and Antonio Guarino, ‘Introduction’, in *Herd Behavior in Financial Markets: An Experiment with Financial Market Professionals,* (2008), 4

Clatworthy, Laura, ‘What is Happening with Crypto Regulation.’ *Rosenblatt,* 14 November 2023. Available at: https://shorturl.at/dkntZ 30 March 2024

Conlon, Thomas, and McGee, Richard, ‘Safe Haven or Risky Hazard? Bitcoin during the Covid-19 Bear Market,’ *Finance Research Letters* 35 (2020)

Corbet, Shaun, Hou, Greg, Hu, Yang, Larkin, Charles, and Oxley, Les, ‘Any Port in the Storm: Cryptocurrency Safe-havens during the COVID-19 Pandemic’ *Economics Letters,* 194 2020

Davidson, Helen, ‘First Covid-19 case happened in November, China government records show - report,’ *The Guardian,* 2020, Available at: https://shorturl.at/dkntZ Accessed 30 March 2024

Espinosa-Méndez, Christian, & Arias, Javier. 'COVID-19 effect on herding behaviour in European capital markets.' *Equilibrium Quarterly Journal of Economics and Economic Policy*, 15 (2020)

Fama, Eugene F., 'Efficient Capital Markets: A Review of Theory and Empirical Work.' *Journal of Finance,* 25 (1970)

Gärling, T., Kirchler, E., Lewis, A., & van Raaij, F. (2009). 'Psychology, Financial Decision Making, and Financial Crises.' *Psychological Science in the Public Interest,* 10

Ge Huang, Vicky, 'BlackRock Launches First Tokenized Fund on Ethereum Blockchain' *The Wall Street Journal* 2024, Available at: https://shorturl.at/bZ235, accessed 19 March 2024

Godoy, Jody, 'US SEC Seeks $2 Billion from Ripple Labs, Chief Legal Officer Says' *Reuters* 2024, Available at: https://shorturl.at/hpuV3, accessed 29 February 2024

Goodell, John W., and Goutte, Stephane, 'Co-movement of COVID-19 and Bitcoin: Evidence from Wavelet Coherence Analysis,' *Finance Research Letters* 38, 2021

https://www.imf.org/en/Blogs/Articles/2020/04/14/blog-weo-the-great-lockdown-worst-economic-downturn-since-the-great-depression, accessed 22 April 2024

https://www.mordorintelligence.com/industry-reports/cryptocurrency-market, accessed 14 April 2024

Jirasakuldech, B., and Emekter, Riza, 'Empirical Analysis of Investors' Herding Behaviours during the Market Structural Changes and Crisis Events: Evidence from Thailand,' *Global Economic Review*, 50 (2021)

Kahneman, Daniel, & Tversky, Amos. 'Prospect Theory: An Analysis of Decision under Risk.' *Econometrica*, 47 (1979)

Lux, Thomas, ‘Herd Behavior, Bubbles, and Crashes’, *The Economic Journal*, 105 (1995)

Munawar, Hafiz Suliman, Khan, Sara Imran, Ullah, Fahim, Kouzani, A., & Mahmud, M. P. ‘Effects of COVID-19 on the Australian Economy: Insights into the Mobility and Unemployment Rates in Education and Tourism Sectors.’ *Sustainability* (2021)

Naipal, Karan, ‘Swift vs Ripple: 6 key areas of comparison for international business payments’ *Medium* 2023, Available at: https://shorturl.at/oABE5, accessed 24 March 2024

Napoletano, E., ‘What is Litecoin? How Does it Work?’, *Forbes* 2022, Available at: https://shorturl.at/gDMN9, accessed 30 March 2024.

John von Neumann and Oskar Morgenstern, in *Theory of Games and Economic Behavior*, (Princeton 1944)

Nietzsche, Friedrich, ‘The Way of the Creating One’, *Thus Spoke Zarathustra: A Book for All and None*, (London, 1961)

Quarterly market share of selected cryptocurrencies, based on market cap 2013-2023 (https://www.statista.com/statistics/730782/cryptocurrencies-market-capitalization, accessed 30 March 2024)

Thaler, Richard, and Sunstein, Cass R., *Nudge: Improving Decisions about Health, Wealth, and Happiness*, (New Haven 2008)

Thaler, Richard, in *Misbehaving: The Making of Behavioral Economics*, (New York 2015)

Trotter, Wilfred, *Instincts of the Herd in Peace and War* (London, 1916)

‘Understanding Crypto Investments: A Comprehensive Guide’, *Bitstamp Learn*, 2023

WHO COVID-19 Dashboard (https://data.who.int/dashboards/covid19/cases?n=c, accessed 23 March 2024)

Yarovaya, Larisa, Matkovskyy, Roman, and Jalan, Alaksha, ‘The Effects of a "Black Swan" Event (COVID-19) on Herding Behaviour in Cryptocurrency Markets,’ *Journal of International Financial Markets, Institutions & Money*, 75 (2021)

Appendix

Stata Do-file:

//We first start by calculating a unique observation id. This is done so that Stata recognizes the dates in the dataset. We create a unique id because the command tsset, which is used to set the time variable, does not work for repeated values.

egen obs_id = seq(), from(1)

tsset obs_id

//here we start the process of calculating the CSAD values, which begins with calculating the hourly returns of each cryptocurrency

gen btc_return = (btcusdclose / l1.btcusdclose) - 1

gen eth_return = (ethusdclose / l1.ethusdclose) - 1

gen ltc_return = (ltcusdclose / l1.ltcusdclose) - 1

gen xrp_return = (xrpusdclose / l1.xrpusdclose) - 1

gen btc_return = (btceurclose / l1.btceurclose) - 1

gen eth_return = (etheurclose / l1.etheurclose) - 1

gen ltc_return = (ltceurclose / l1.ltceurclose) - 1

gen xrp_return = (xrpeurclose / l1.xrpeurclose) - 1

// calculate the average return of the equally weighted portfolio at each time point (repeat the process for the European values)

egen avg_return = rowmean(btc_return eth_return ltc_return xrp_return)

```
//calculate the CSAD for each cryptocurrency (repeat for european values)

gen abs_dev_btc = abs(btc_return - avg_return)

gen abs_dev_eth = abs(eth_return - avg_return)

gen abs_dev_ltc = abs(ltc_return - avg_return)

gen abs_dev_xrp = abs(xrp_return - avg_return)

//calculate the mean CSAD (repeat the process for the european CSAD value)

egen CSAD_usa = rowmean(abs_dev_btc abs_dev_eth abs_dev_ltc abs_dev_xrp)

// after loading the case data, we manually remove all the data that is not useful to us, and then we aggregate the case data into a single variable for daily cases we do this using a loop. We only need to do this for the Europe data because the USA data is already aggregated for us.

egen europe_cases = rowtotal(germany_cases france_cases spain_cases netherlands_cases italy_cases)

//now we expand it to fit the hourly price data

expand 24

// the event variable is created manually without code in an excel file

egen event_variable = .

// normalizing the variables using z-scores so that they are on the same scale

foreach var of varlist csad_usa cases_usa {

    summarize `var'
```

```
    scalar `var'_mean = r(mean)

    scalar `var'_sd = r(sd)

}

// Calculate z-scores for the variables

foreach var of varlist csad_usa cases_usa {

    gen `var'_z = (`var' - `var'_mean) / `var'_sd

}

//plotting the variables

tsline norm_csad_usa norm_cases_usa

//we repeat the same for the european csad and cases variables

// We start with the Markov-Switching Model

mswitch dr norm_csad norm_cases_usa event_dummy, vce(robust)

// calculate the residuals

predict residuals, residuals

// We calculate the AIC

estat ic

// ACF and PACF plots

ac residuals

pac residuals
```

```
// Calculate the Ljung-Box statistic for heteroskedasticity

wntestq residuals

// VAR model where we repeat the same process with the M-switch model, we also do the same with subsequent var models and the ARIMA models

var norm_csad, lags(1/80) exog(norm_cases_usa event_dummy)

predict residuals, residuals

ac residuals

pac residuals

estat ic

wntestq residuals

// ARIMA model

arima norm_csad norm_cases_usa event_dummy, arima(1,0,1) vce(robust)

predict residuals, residuals

ac residuals

pac residuals

estat ic

wntestq residuals

// we now move to forecasting, we restimate the model and then predic the next 29500 values

var norm_csad, lags(1/80) exog(norm_cases_usa event_dummy)
```

```
//forecast the values and place them into a new variable

tempfile forecasts

forvalues i = 1(500)30000 {

    predict forecasted_norm_csad`i' if _n > 22 & _n <= (`i' + 22)

    save `forecasts'_`i', replace

}

//we calculate the MAE, MSE, and RMSE measures. this process, including the forecasting is repeated and each model's measures are compared with one another

summarize actual_variable forecasted_variable, meanonly

di "MAE: " abs(r(mean))

// Compute absolute differences

gen abs_diff = abs(actual_variable - forecasted_variable)

// Compute squared differences

gen squared_diff = abs_diff^2

// Calculate MSE

sum squared_diff, meanonly

di "MSE: " r(mean)

// Calculate RMSE

di "RMSE: " sqrt(`mse')
```

//plot the forecasted variable as well as the original variable

tsline forecasted_norm_csad_usa norm_csad_usa

// we then summarize each mean for the forecasted period, the real period, and the pre-period to compare mean csad values

sum forecasted_norm_csad_usa

sum norm_csad_usa

sum pre_norm_csad_usa

// end do-file

Supplementary files available upon request:

USA Dataset (PP)

USA Dataset (PPP)

Europe Dataset (PP)

Europe Dataset (PPP)

Forecast Data (USA)

Forecast Data (Europe)