Department of Political Economy

King's College London

Dissertation proposal cover sheet

Programme: History and Political Economy BSc (Hons)

Module code: 6SSPP352

Candidate number: AE13330

Title of your proposed dissertation: Measuring the Effects of Herd Behavior on Cryptocurrency markets

Supervisor name: Dr. Teresa Esteban-Casanelles

Word count*: 1575

* Please see p. 5 of the Module Handbook for details of what is and is not included in the word count for undergraduate dissertations in the Department of Political Economy.

Title

Herding in Cryptocurrency Markets: Identification and Forecasting under COVID-19

Aims of the Research

The Efficient Market Hypothesis (EMH) posits that asset prices fully reflect all available information[1]. However, real-world complexities in investor behavior often lead to violations of this hypothesis. The COVID-19 pandemic amplified behavioral economic phenomena from 2019 to 2022, drawing increased attention to their impact on markets. Understanding such behavior is valuable for individuals, firms, and macroeconomic policymakers. Cryptocurrency markets, known for their susceptibility to these behaviors, experienced heightened activity during the pandemic-induced recession, characterized by less-experienced investors influenced by social media. This is known as herding, where investors follow outside advice rather than their own. This paper aims to analyze the influence of herding on future growth in global cryptocurrency markets, considering the pandemic period (PP) and an out-of-sample post-pandemic period (PPP). The study frames its objectives through the following hypotheses:

Herding Identification

H_1: Herd behavior is not significant in any of the markets examined during PP.

H_2: The frequency and magnitude of herd behavior does not significantly differ between PP and PPP.

Growth Forecasting

H_3: Herd behavior during PP does not significantly influence the future growth prospects of the markets chosen.

[1] Fama, Eugene F., ‘Efficient Capital Markets: A Review of Theory and Empirical Work’, *Journal of Finance*, 25, (New York 1970), 414

H_4: The future growth prospects of markets during PP and PPP do not diverge significantly.

The analysis will utilize time-series modeling to identify patterns of herding in hourly price data of the ten foremost currencies in the four largest markets: the United States, Europe, India, and South Korea. The results of this analysis will subsequently inform the construction of a second time-series model that will forecast future growth potential.

Theoretical Framework

The concept of herding draws from various economic, psychological, and financial frameworks. Keynesian economics, with its notion of 'animal spirits,' was among the first to recognize the influence of psychological factors like loss and risk aversion[2]. Irving Janis expanded on this, highlighting individuals' preference for conformity in economic decisions over critical thinking[3]. The evolution of this idea into rational decision-making versus conformity is evident in Kahnemann and Tversky's prospect theory and Timur Kuran's study on information cascades[4]. Prospect theory operationalizes behavior patterns using a value function of expected utility. Richard Thaler's work in behavioral finance further contributes by building upon Kahnemann and Tversky's ideas through mental accounting and the endowment effect[5]. In terms of econometric models, cross-sectional absolute dispersion (CSAD) and cross-sectional standard deviation (CSSD) play a crucial role not only in economics and finance but also in broader statistical contexts such as political science. Christie and Huang (2000) and Chang et al.'s focus on herding in equity markets heavily influence the application of CSAD and CSSD in this study. These metrics are key

[2] Keynes, John M., *The General Theory of Employment, Interest, and Money*, (Cambridge 1936), 161
[3] I. Janis, 'Groupthink: Psychological Studies of Policy Decisions and Fiascoes', (1982)
[4] D. Kahnemann, and A. Tversky, 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica*, 47, (1979),
[5] R. Thaler, *Misbehaving: The Making of Behavioral Economics,* (2015)

components of the analysis used to verify H_1 and H_2. H_3 and H_4 will be verified using the time-series model autoregressive integrated moving average (ARIMA), introduced by Box and Jenkins. The results from CSAD and CSSD models will inform the forecasting analysis of future market growth through a modified ARIMA model known as SARIMAX.

Research Problem and Literature Review

The studies by Chang et al. were the first to explore herding in markets, using CSAD and CSSD models to detect herding in American, South Korean, and Japanese equity markets. Their analysis revealed herding in developing markets, specifically the South Korean market[6]. This research forms the basis for subsequent studies on cryptocurrency, such as Larisa Yarovaya et al.'s work, which identified unconditional herding in the USD, JPY, and EUR markets but not in the KRW market, with herding contingent on market conditions[7]. Studies by Ki Hong Choi et al. and Kumar et al. present varied results, focusing on herding identification. Phillipas et al. emphasize sentiment scores, incorporating Twitter mentions and the CBOE volatility index[8]. However, something that is common to all these studies is a lack of foresight. This paper addresses these gaps through PP and PPP, enabling a comparative narrative and informing a new approach—forecasting—raising questions about herding's influence on future market performance.

[6] C. L. Chang, M. McAleer, and Y. Wang, 'Herding Behaviour in Energy Stock Markets During the Global Financial Crisis, SARS, and ongoing COVID-19', *Renewable and Sustainable Energy Reviews,* 134, (2020)
[7] L. Yarovaya, R. Matkovskyy, and A. Jalan, 'The Effects of a "Black Swan" Event on Herding Behavior in Cryptocurrency Markets', *Journal of International Financial Markets, Institutions, and Money,* (2021)
[8] D. Philippas, H. Rjiba, K. Guesmi, and S. Goutte, 'Media Attention and Bitcoin Prices', *Finance Research Letters,* 30 (2019)

Research Methodology

Before any significant analysis can begin, the data must be cleaned, and variables must be added to the dataset (STATA). This involves removing outliers, standardizing data formats, removing duplicate entries, and converting data types. At present, the dataset containing hourly price data only contains variables such as trade volume, date, open and close price, and high and low price. As a result, volatility (VIX), financial ratios (e.g., P/E, dividends), CSAD, CSSD, and market capitalization must be added or calculated. The paper will conduct the study using an analytical framework designed for the purpose of this dissertation, called CONEX, which is defined below:

1. The comparative overlay for nexus examination (CONEX) is used when the examination of a single period can lead to inaccuracies in forecasting. The data from PP potentially includes events unique to that period that may cause issues with forecasting accuracy. The nexus in this case is between the pandemic and post-pandemic periods. Including an out-of-sample period can provide comparisons between the results of the two periods. The comparative overlay represents comparing the elements within each period's analysis. Here, the overlay represents comparing the forecasting results of markets with and without significant herding.

The CSAD and CSSD formulas are below, whose values will be calculated using STATA:

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

Where N is the number of assets, R_i is the return of the i-th investment and $\bar{R}$ is the mean return across all assets.

$$CSSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(R_i - R)^2}$$

CSAD measures the absolute difference between individual asset returns. A low CSAD indicates a clustering of returns, which can be an indicator of herding. Likewise, the CSSD measures the standard deviation of individual asset returns around the mean return. The CSAD and CSSD metrics are the key tools used to verify H_1 in PP only, but the CSAD and CSSD metrics from both periods must be compared to verify H_2.

After concluding the identification analysis in both time periods, a multivariate approach will be used to forecast the dependent variables. Time series models are different from regressions or other models in that they lack a distinct independent variable. The SARIMAX model utilizes the future variable as the dependent variable, while the independent variable is the past or current value of the same variable. Setting up the SARIMAX model requires careful consideration of seasonality and exogenous variables to ensure the validity of the results. Cryptocurrency markets are known to have strong seasonal trends, such as the year-end rally, the January effect, and crypto events or conferences. Additionally, the analysis requires the inclusion of herding and pandemic variables to effectively incorporate their effects on future growth. The herding variables are the hourly CSSD and hourly trade volume. The pandemic variables are daily COVID cases and the daily unemployment rate. The four dependent variables that will be analyzed are market capitalization, asset price, price-to-earnings ratio (P/E), and return-on-investment ratio (ROI). The general equations for each SARIMAX model are below:

$$\text{Cap}_t = \beta_0 + \phi_1\text{Cap}_{t-1} + \theta_1\varepsilon_{t-1} + \beta_2\text{Cases}_t + \beta_3\text{Rate}_t + \beta_4\text{Volume}_t + \beta_5\text{CSSD}_t + \varepsilon_t$$

$$\text{Price}_t = \gamma_0 + \phi_2\text{Price}_{t-1} + \theta_2\varepsilon_{t-1} + \gamma_2\text{Cases}_t + \gamma_3\text{Rate}_t + \gamma_4\text{Volume}_t + \gamma_5\text{CSSD}_t + \varepsilon_t$$

$$\text{P/E}_t = \delta_0 + \phi_3\text{P/E}_{t-1} + \theta_3\varepsilon_{t-1} + \delta_2\text{Cases}_t + \delta_3\text{Rate}_t + \delta_4\text{Volume}_t + \delta_5\text{CSSD}_t + \varepsilon_t$$

$$\text{ROI}_t = \lambda_0 + \phi_4\text{ROI}_{t-1} + \theta_4\varepsilon_{t-1} + \lambda_2\text{Cases}_t + \lambda_3\text{Rate}_t + \lambda_4\text{Volume}_t + \lambda_5\text{CSSD}_t + \varepsilon_t$$

Where β_0, γ_0, λ_0, and δ_0 represent constants, and ϕ_1 to ϕ_4 represent the coefficients of the autoregressive order, and θ_1 to θ_4 represent the coefficients of the moving average order. The coefficients β_1, γ_1, λ_1, and δ_1 to β_4, γ_4, λ_4, and δ_4 are the coefficients for each of the exogenous variables. These are general equations for forecasting, as the order of differencing (*d*), and seasonal order (P, D, Q) must be identified before the specific equations can be created. This is done to address autocorrelation, stationarity, and temporal dependencies. Failure to address these issues can lead to biases, unreliable forecasts, and incorrect statistical inferences.

Unlike H_1 and H_2, H_3 and H_4 both rely on comparing the forecasting results of PP and PPP. PPP's forecasting results represent 'normal' trends and must be compared with the results of PP to identify abnormalities. The results can be interpreted through the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and p-values. These metrics are used to assess the model's fit and the significance of each variable. Diagnostic and forecasting plots can be used to plot future data against past and current data to visualize trends. Identifying limitations in the data and models, such as changes caused by sudden shocks, will be possible as well.

Bibliography

Chang, Chia Lin, McAleer, Michael, and Wang, Yu-Ann, ‘Herding Behaviour in Energy Stock Markets During the Global Financial Crisis, SARS, and ongoing COVID-19’, *Renewable and Sustainable Energy Reviews*, 134, (Taichung 2020)

Choi, Ki-Hong, Kang, Song-Hoon, and Yoon, Seong-Min, ‘Herding Behavior in Korea’s Cryptocurrency Markets’, 54, (London 2022)

Christie, William G., and Huang, Roger D., ‘Following the Pied Piper: Do Individual Returns Herd around the Market?’, *Financial Analysts Journal*, 51, (Nashville 1999)

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

Janis, Irving, ‘Groupthink: Psychological Studies of Policy Decisions and Fiascoes’, (Boston 1982)

Kahnemann, Daniel, and Tversky, Amos, ‘Prospect Theory: An Analysis of Decision Under Risk’, *Econometrica*, 47, (Palo Alto 1979)

Keynes, John M., *The General Theory of Employment, Interest, and Money*, (Cambridge 1936).

Kumar, Ashish, ‘Empirical Investigation of Herding in Cryptocurrency Markets under Different Market Regimes’ *Behavioral Finance Review*, 13, (London 2021)

Philippas, Dionisis, Rjiba, Hatem, Guesmi, Khaled, and Goutte, Stephane, ‘Media Attention and Bitcoin Prices’, *Finance Research Letters*, 30 (Paris 2019)

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

Yarovaya, Larisa, Matkovskyy, Roman, and Jalan, Akanksha, 'The Effects of a "Black Swan" Event on Herding Behavior in Cryptocurrency Markets', *Journal of International Financial Markets, Institutions, and Money*, (Amsterdam 2021)