Research Article
Special Issue on Econometrics and Business Analytics

Siyun He and Rustam Ibragimov*

Predictability of cryptocurrency returns: evidence from robust tests

https://doi.org/10.1515/demo-2022-0111
received September 26, 2021; accepted March 31, 2022

Abstract: The paper provides a comparative empirical study of predictability of cryptocurrency returns and prices using econometrically justified robust inference methods. We present robust econometric analysis of predictive regressions incorporating factors, which were suggested by Liu, Y., & Tsyvinski, A. (2018). Risks and returns of cryptocurrency. NBER working paper no. 24877; Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency. The Review of Financial Studies, 34(6), 2689–2727, as useful predictors for cryptocurrency returns, including cryptocurrency momentum, stock market factors, acceptance of Bitcoin, and Google trends measure of investors’ attention. Due to inherent heterogeneity and dependence properties of returns and other time series in financial and crypto markets, we provide the analysis of the predictive regressions using both heteroskedasticity and autocorrelation consistent (HAC) standard-errors and also the recently developed t-statistic robust inference approaches, Ibragimov, R., & Müller, U. K. (2010). t-statistic based correlation and heterogeneity robust inference. Journal of Business and Economic Statistics, 28, 453–468; Ibragimov, R., & Müller, U. K. (2016). Inference with few heterogeneous clusters. Review of Economics and Statistics, 98, 83–96. We provide comparisons of robust predictive regression estimates between different cryptocurrencies and their corresponding risk and factor exposures. In general, the number of significant factors decreases as we use more robust t-tests, and the t-statistic robust inference approaches appear to perform better than the t-tests based on HAC standard errors in terms of pointing out interpretable economic conclusions. The results in this paper emphasize the importance of the use of robust inference approaches in the analysis of economic and financial data affected by the problems of heterogeneity and dependence.

Keywords: Bitcoin, cryptocurrencies, predictive regressions, robust inference, HAC, t-statistic inference

MSC 2020: 62P20, 91B84

1 Introduction

This paper focuses on how different factors predict the returns of cryptocurrencies. We provide a comparative analysis of predictability of cryptocurrency returns using econometrically justified robust inference methods that account for inherent heterogeneity and dependence in the time series dealt with. We focus on five predictive models in the study by Liu and Tsyvinski [29,30] who provide the first detailed empirical analysis and estimates of predictive regressions for cryptocurrency returns, with the assessment of significance of their coefficients based on i.i.d. standard errors in the former working paper and the use of...
heteroskedasticity and autocorrelation consistent (HAC) Newey-West standard errors in some of the predictive models in its latter recently published version.

We provide a study of the robustness of the empirical analysis of the predictive models by using both HAC inference based on Newey-West standard errors [35] (see also [3] for HAC inference approaches) and the newly developed $t$-statistic robust inference approaches proposed in the study by Ibragimov and Müller [22]. Following Liu and Tsyvinski [29,30], some of the predictive factors used in the analysis are unique for cryptocurrency and relate to crypto risk, and others are traditional factors such as those related to stock market risk. Three different major cryptocurrencies are analyzed in this paper: Bitcoin, Ripple, and Ethereum.

As is well known, many important economic and financial variables, including returns’ time series, are characterized by inherent heterogeneity, autocorrelation and nonlinear dependence, volatility clustering, excess kurtosis, heavy tailedness, and outliers, as in the case of GARCH or AR-GARCH-type models for financial and cryptocurrency returns (see, among others, the review and discussion in [4,8,9,13–15,20,26,33,42]). Naturally, as compared to developed financial and economic markets dealt with in the most of the literature, the properties of heterogeneity, volatility clustering, nonlinear dependence, and heavy tailedness are typically even more pronounced in emerging and cryptocurrency markets that are subject to more frequent and more pronounced external and internal shocks.

It is important to use inference methods that account for autocorrelation and heterogeneity in their analysis (see, among others, [20], and references therein). In general, the i.i.d. standard errors are not valid under heterogeneity, autocorrelation or dependence, and hence so are the results of $t$-tests since standard errors may be underestimated or overestimated.

The most widely used approach in econometrics, economics, and finance to conduct asymptotically valid inference under heterogeneous and dependent data is that based on consistent variance estimators. In the context of time series, asymptotically valid inference is traditionally based on the popular HAC standard errors with the adjustments for serial correlation and dependence derived by [3,35].

This paper focuses on econometrically justified inference in predictive regressions for cryptocurrency returns and prices using robust methods. Due to inherent heterogeneity and dependence properties of returns and other time series in financial and crypto markets, we provide the analysis of the predictive regressions using (HAC) Newey-West standard errors [35].

While widely used and quite general, HAC inference procedures often have poor finite sample properties, especially when observations exhibit pervasive and pronounced correlations (see, among others, the discussion in [22,23], Section 3.3 in [20], and references therein). Motivated by this, we further present the analysis of the predictive regressions using recently developed $t$-statistic robust inference approaches [22,23]. Robust tests in the approaches are based on $t$-statistics in group estimates of model parameters and do not require consistent estimation of their limiting variances, as shown in HAC methods. In the context of predictive regressions, robust large sample inference is conducted as follows: the time series is partitioned into $q \geq 2$ groups of consecutive observations, the predictive regression is estimated for each group, and then a standard $t$-test of level 5% with the resulting $q$ estimates of a regression parameter of interest is conducted (see [22], Section 3.3 in [20] for details). The $t$-statistic robust inference approaches have been shown to have appealing finite sample performance for time series, panel, clustered and spatially correlated data exhibiting heterogeneity, autocorrelation, and dependence of largely unknown form that are typical for real-world financial and economic markets (see, in addition to the aforementioned references, the recent study by [16]). In addition, according to the numerical results in the aforementioned papers and several other works, finite sample properties of the $t$-statistic approaches to robust inference

---

1 On consistent standard errors, see also [51] for an OLS regression with independent but not necessarily identically distributed errors; [43] for clustered data that includes panel data as a special case; and [12] for spatially dependent data ([46] provided a detailed discussion and empirical examples on consistent standard errors in a wide range of important models and heterogeneity and dependence settings in econometrics). In the context of inference in long-horizon regressions, [49] provided the adjustment for serial correlation in the form of the re-scaled $t$-statistic $t / \sqrt{T}$ that is different from that in the Newey-West $t$-statistics used in this paper.
compare favorably with the widely used robust inference methods based on consistent variance estimation and alternative procedures, including those based on HAC and clustered standard errors. The finite sample performance of the $t$-statistic approaches is often better than those of the latter methods, especially in the case of pervasive and pronounced heterogeneity, autocorrelation, dependence, and heavy tailedness in the data (see, for instance, simulation results in Table 1 in [22] for comparisons of finite-sample performance of the $t$-statistic robust approaches with the traditional HAC methods in inference on time-series regressions).

One should also note that the moment assumptions for the validity of HAC inference methods, including the approaches based on Newey-West standard errors, such as the often imposed condition that the regressors and the regression errors have finite eighth moments (see Section 16.4 in [46]), are typically too restrictive for predictive regressions for financial returns. For example, according to many studies in the literature, the distributions of financial returns are heavy tailed with tail indices smaller than four, thus implying infinite fourth moments and, in the case of emerging markets, even tail indices smaller than 2 are not uncommon (see the review in Section 1.2 in [20] and the references therein). On the other hand, $t$-statistic robust inference approaches can be used under pronounced heavy tailedness in the data, even including the case of infinite variances (see the discussion in Section 3.3 in [20]).

These conclusions suggest that the $t$-statistic robust inference approaches are useful complements to the widely used procedures based on consistent (e.g., HAC, including Newey-West) standard errors and motivate the use of both of them in robust inference.²

We provide comparisons of robust predictive regression estimates between different cryptocurrencies and their corresponding risk and factor exposures. We further contrast the results from the robust predictive regression analysis with the conclusions implied by the i.i.d. standard errors. In general, the number of significant factors disappears or decreases as we use more robust $t$-tests, and the $t$-statistic robust inference approaches appear to perform better than the $t$-tests based on HAC standard errors in terms of pointing out interpretable economic conclusions. In particular, this is observed for predictive regressions incorporating cryptocurrency momentum (see Section 3.1) that indicates that cryptocurrency returns exhibit the stylized fact of the absence of linear autocorrelations and linear dependence similar to financial returns (see [13] for a review and discussion of this and other key stylized facts for financial markets). In addition, the aforementioned property is confirmed for predictive regressions for Bitcoin returns incorporating price-to-acceptance ratio (an analog of price-to-dividend ratio for stocks, see Section 3.3) that points to apparent/potential absence of intrinsic/fundamental value of Bitcoin, in contrast to financial assets with pricing models incorporating dividends. According to the results in the paper, the aforementioned disappearance of significance under more robust $t$-statistic approaches also holds for predictive regressions for cryptocurrency returns incorporating Google trends (Section 3.4), indicating that cryptocurrency prices are affected by investors’ sentiment/attention to a lesser extent as compared to financial asset prices.

To our knowledge, this paper is the first to apply several econometrically justified robust inference approaches in examining the predictability of cryptocurrency returns, including $t$-statistic robust inference methods. Its results, in particular, illustrate the importance of using HAC and other robust inference methods in the analysis of the statistical significance of different factors in predictive models for financial and crypto returns and other important economic and financial variables exhibiting heterogeneity and dependence over time.

The results in this paper point to the advantages of complementing empirical analyses of dependent and heterogeneous economic and financial data using HAC and other (e.g., clustered) consistent standard errors by the simple-to-implement $t$-statistic robust inference approaches that do not require consistent estimation of limiting variances of estimators dealt with. The analysis in the paper further emphasizes that statistical conclusions from econometrically justified robust tests are different from those implied by i.i.d. standard errors.

² The working paper version [21] of [22] provided, in Section 3.4, a simple to implement test that can be used to determine which robust inference procedure – that based on the $t$-statistic approach or consistent (e.g., HAC) standard errors – is more appropriate to use in the analysis of dependent and heterogeneous data at hand.
The main goal of the paper is to emphasize the importance of the use of robust inference approaches in the analysis of economic and financial data affected by the problems of heterogeneity and dependence, and predictive regressions for cryptocurrency returns are mainly used for illustration and comparisons of conclusions implied by different inference methods.

At the same time, the results and robust inference methods used in this paper may be helpful in further analysis of predictive models for cryptocurrency markets and the development of economic models for them. The results in this paper further point out to several similarities between cryptocurrency and financial markets. The empirical applications of robust inference methods in the paper may inspire future researches on robust tests of crypto markets’ efficiency or the robust analysis of stylized facts of cryptocurrency returns.

Many papers in the literature have focused on the analysis of empirical properties of cryptocurrency markets. [1] emphasizes the importance of using correct data sources in empirical analyses of crypto markets, especially in the case of time series of returns and prices of crypto market indices. It is widely agreed that cryptocurrencies suffer from much higher volatility compared to the traditional currencies (see, among others, the review and discussion in Ch. 8 in [2,4,5,8,14,52]). The Bitcoin market efficiency is discussed by [34,48,50]. [11] presents an econometric analysis of bubbles in the Bitcoin market using the methodology developed in [39,40]. According to the analysis in [6], cryptocurrencies are highly exposed to tail risk within crypto markets but not in other global markets. Nguyen et al. [36] show that the right tail risk among cryptocurrencies is more pronounced as compared to the left one. Trimborn and Härdle [47] proposed the construction of an index that quickly reacts to cryptocurrency market changes.

Several authors have also focused on the analysis of predictability of cryptocurrency prices and returns. Poyser [41] summarized factors that drive the price of cryptocurrency into two categories: internal factors and external factors. Sovbetov [45] used cointegration models and tests to identify factors that influence the prices of five major cryptocurrencies. Lintilhac and Tourin [28] applied cointegration models in the analysis of pairs trading strategies in the bitcoin markets.

Liu and Tsyvinski [29,30] provided the first detailed empirical analysis and estimates of predictive regressions for cryptocurrency returns that incorporate several traditional and crypto risk factors, including foreign exchange rates, prices of metal commodities, momentum, and measures of investors’ attention. Among other results, the authors propose, for the first time in the literature, an analog of the price to dividend ratio factor for cryptocurrencies, with, e.g., the intrinsic/fundamental value of Bitcoin proxied by the number of Bitcoin wallet users. The analysis in working paper Liu and Tsyvinski [29] is based on i.i.d. and (i.i.d.) bootstrapped standard errors, while the assessment of the significance of the coefficients in some of the predictive regressions in its recently published version [30] is based on HAC Newey-West standard errors. Based on the empirical analysis, [29,30] reach the conclusion that cryptocurrency returns appear to be less affected by factors related to traditional asset classes, but appear to be predictable by momentum and investors’ attention.³

Liu et al. [31] provided estimates of factor models for the cross-section of returns for a large set of cryptocurrencies and concluded that the returns are explained by three cryptospecific factors, namely, market, size, and momentum.⁴ [7,27,32] documented large differences in bitcoin prices across exchanges located in different countries, and for different fiat currency pairs. [32] argued that these bitcoin discounts are explained by capital market segmentation, capital controls, and weak financial institutions. [27] pointed to the importance of market inefficiencies. Focusing on the most reputable exchanges, [7] showed that the large bitcoin price differences in them are due to costly arbitrage and idiosyncratic risks. Among other results, [7] provided estimates, with Newey-West HAC standard errors, for the three-factor model for cryptocurrency returns in [31] incorporating a crypto market, size, and momentum factors (see Table 4 in [7]). The authors further relate idiosyncratic risk measured as residuals from the estimated three-factor

³ Predictive regressions for cryptocurrency returns and prices may be particularly useful for determining and testing factors important in valuation models for cryptocurrencies (see [5,24,25,37,38] for pricing and equilibrium models for bitcoin and other cryptocurrency markets).

⁴ To our knowledge, the analysis in [31] is based on i.i.d. and i.i.d. bootstrapped standard errors, similar to [29].
regressions to a set of portfolio characteristics and find that it is higher for portfolios containing exchange-
currency pairs with higher liquidity and execution risks.

This paper is organized as follows. The data used in the analysis are described in Section 2. Section 3
presents the results of the empirical analysis of predictive regressions for cryptocurrency returns and
factors considered. Section 4 makes some concluding remarks. Appendix provides several tables on the
empirical analysis referred to in the paper.

2 Data

The data frequency is specified throughout the paper. Sections 3.1 and 3.2 deal with predictive regressions
for daily cryptocurrency returns using daily data on the returns and stock market factors. Section 3.3
provides estimation results in predictive regressions for weekly Bitcoin returns on the proxy for acceptance
of Bitcoin given by the ratio of the spot price of Bitcoin to the number of wallet users (the data on the latter
number is updated weekly). Section 3.4 provides the estimates for predictive regressions of weekly crypto-
currency returns on the proxies for weekly investors’ attention measured using the data on Google trends.
Weekly cryptocurrency returns are calculated using the data on daily cryptocurrency prices and returns.

We use Yahoo Finance as the main data source for cryptocurrency prices.⁵ The time interval for Bitcoin
is from 01/01/2011 to 31/05/2018, for Ripple is from 04/08/2013 to 31/05/2018, and for Ethereum is from 07/
08/2015 to 31/05/2018.⁶

As in [29,30], the CAPM [44], Fama-French three-factor [17], Carhart four-factor [10], Fama-French five-
factor [18], and Fama-French six-factor [19] data are obtained from Kenneth French’s website.⁷ The data on
the number of Bitcoin wallet users are from blockchain.info.⁸ Further, as in [29,30], the frequencies of
words “bitcoin,” “ripple,” and “ethereum” searched on Google are measured by Google Trends.⁹

3 Empirical analysis

3.1 Cryptocurrency momentum

The analysis of cryptocurrency momentum is based on the daily returns of Bitcoin, Ripple, and Ethereum.¹⁰
The tables on the empirical results in this and subsequent sections provide the estimates of the coefficients
of predictive regressions considered (e.g., the regression of daily cryptocurrency returns on their lags in the
momentum analysis) and the HAC t-statistics for the coefficients based on Newey-West standard errors.¹¹
For the t-statistic robust inference approach in [22], the tables also provide the t-statistics in group esti-
mates of the predictive regression coefficients calculated using q = 4 and q = 8 groups of consecutive time

---

5 Cryptocurrency prices are retrieved from: https://finance.yahoo.com/. According to Yahoo Finance, prices are provided by CoinMarketCap. The market opens at 12:00 AM and closes at 11:59 PM UTC time. The dataset for cryptocurrency prices is thus different from that in [29,30]. As discussed in Section 1, the estimates of predictive regressions for cryptocurrency returns for the dataset are mainly used for illustration and comparison of conclusions implied by different inference methods, including econometrically justified, e.g., HAC and t-statistic, robust inference approaches. The conclusions in the paper should not be interpreted as evidence for or against a particular potentially predictive factor for crypto returns.

6 The frequency of the data used is indicated in the description of the results in Section 3.

7 Stock factors are retrieved from: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

8 Bitcoin wallet users data are retrieved from: https://www.blockchain.com/.

9 “bitcoin,” “ripple,” and “ethereum” frequency data are retrieved from: https://trends.google.com/trends/.

10 Usually, one expects to observe weaker predictability as the time horizon goes longer.

11 The automatic bandwidth selection described in [35] is used throughout.
In the $t$-statistic robust inference approaches, it is concluded that the predictive regression coefficient is significant at the 5% level if the absolute value of the $t$-statistic calculated using its $q$ group estimates is greater than the 0.975-quantile of a Student-$t$ distribution with $q - 1$ degrees of freedom (so that, for instance, the quantiles of a Student-$t$ distribution with 3 degrees of freedom are used in the case of $q = 4$ groups and with 7 degrees of freedom in the case of $q = 8$ groups reported in the tables).

One would expect the significance of a predictive regression coefficient to decrease when using more robust tests (e.g., HAC $t$-tests as compared to those based on i.i.d. standard errors, and similar for $t$-statistic robust inference approach). As illustrated by the results in this and subsequent sections, the exact conclusions of the results depend on the structure of autocorrelation and dependence in the time series considered. For instance, as is well known, for positively autocorrelated time series, HAC $t$-statistics are expected to be smaller than their i.i.d. counterparts, and thus, a coefficient found to be significant using an i.i.d. $t$-test may not appear to be such when a HAC $t$-test is used; the comparisons and conclusions on significance are preserved in the case of negative autocorrelation. If the significance of a factor, such as cryptocurrency momentum and other factors considered in the paper, disappears with more robust tests, this indicates that the corresponding cryptocurrency apparently does not have strong exposure to this factor.

Tables 1–3 present the estimation results for regressions of future daily cryptocurrency returns ($R_{t+h}$, $h = 1, \ldots, 7$, i.e., up to 7-day ahead returns) on the current daily cryptocurrency return ($R_t$).

In line with the above, overall, HAC $t$-statistics in Tables 1–3 appear to be considerably smaller than the $t$-statistics based on i.i.d. standard errors. In particular, the (nonvalid) i.i.d. $t$-tests indicate significant momentum effects for Ripple over the whole time horizon. In contrast, according to HAC $t$-tests, the current Ripple return statistically significantly (and positively) predicts only the first 2-day ahead returns on the cryptocurrency. Similarly, according to HAC $t$-tests, the current Bitcoin return positively and statistically significantly predicts the 2-day, 3-day, and 5-day ahead returns.

In contrast, interestingly, HAC $t$-tests indicate statistically significant predictability of the 5-day ahead Ethereum return, while no momentum effects for the cryptocurrency are significant according to (nonvalid) i.i.d. standard errors.

According to the most of the results for the $t$-statistics robust inference approaches in Tables 1–3, no momentum effects appear to be significant for the cryptocurrencies considered, with the only exception of the 1-day ahead return on Ripple, where both HAC $t$-test and the $t$-statistic approach with $q = 8$ groups indicate significance. This conclusion is important as it indicates that, similar to financial returns (see Ch. 2 in [9], and the review in [13]), cryptocurrency returns appear to exhibit the stylized facts of the absence of linear autocorrelation. Thus, importantly, it appears that the current cryptocurrency returns cannot be used to predict future returns using linear predictive regressions.

### 3.2 Stock market factors

Similar to [29,30], in analogy to the analysis of predictability of daily stock excess returns, we consider CAPM, Fama-French three-factor, Carhart four-factor, Fama-French five-factor, and Fama-French six-factor models. All data used are daily. We regress daily excess cryptocurrency return on the stock market factors...
The estimates and significance analysis for the corresponding predictive regressions for Bitcoin returns are summarized in Table 4. The estimates and significance analysis for predictive regressions for Ripple and Ethereum are provided in Tables A1 and A2 in Appendix. In Table 4 and Tables A1 and A2, the mentioned in these five models. We subtract the risk-free returns (1-month US Treasury bill rate) from Bitcoin, Ethereum, and Ripple returns in order to obtain the excess returns. In total, there are six stock market factors considered: excess return on the stock market (Mkt.RF), market capitalization (SMB), value premium (HML), momentum (MOM), profitability (RMW), and investment (CMA). Detailed constructions of these factors could be found in [10,17–19,44].
Table 4: Bitcoin and stock factors

|       | CAPM       | Excess Bitcoin return |
|-------|------------|------------------------|
|       |            | 3-F        | 4-F        | 5-F        | 6-F        |
| Mkt.RF| −0.015     | −0.137     | −0.133     | −0.025     | −0.019     |
|       | [0.932]    | [0.270]    | [0.271]    | [0.049]    | [0.052]    |
|       | (−0.088)   | (−0.633)   | (−0.641)   | (−0.111)   | (−0.087)   |
|       | (−0.032)   | [−0.595]   | [−0.602]   | [−0.413]   | (−0.553)   |
| SMB   | 0.586      | 0.545      | 0.597      | 0.546      |
|       | [1.488]    | [1.454]    | [1.214]    | [1.174]    |
|       | (0.947)    | (0.860)    | (0.914)    | (0.820)    |
|       | [1.099]    | [0.969]    | [0.994]    | [0.906]    |
| HML   | 0.167      | 0.023      | −0.282     | −0.471     |
|       | [1.132]    | [0.971]    | [0.646]    | [0.523]    |
|       | (0.446)    | (0.063)    | (−0.666)   | (−1.162)   |
|       | [0.808]    | [0.488]    | [−0.406]   | (−0.674)   |
| MOM   | −0.002     | [−0.164]   | [−0.208]   | (−1.295)   |
|       | (−1.103)   | [−4.944]   | (−0.319)   |
| RMW   | 0.184      | 0.162      | −0.694     | −0.702     |
|       | [0.325]    | [0.294]    | (0.231)    | (0.364)    |
| CMA   | 1.260      | 1.322      | [0.247]    | [0.262]    |
|       | (1.902)    | (2.102⁰)   | [1.175]    | [1.284]    |
|       | [1.284]    |           |           |           |
| Alpha | 0.009      | 0.009      | 0.009      | 0.009      |
|       | [3.521⁰]   | [3.569⁰]   | [3.571⁰]   | [3.591⁰]   |
|       | (3.200⁰)   | (3.173⁰)   | (3.163⁰)   | (3.183⁰)   |
|       | [2.393]    | [2.383]    | [2.387]    | [2.450]    |

The excess return is given by the difference between the Bitcoin return and the risk-free return. The values in the first row of each stock factor are coefficients of each stock factor. The t-statistics in square brackets are based on i.i.d. standard errors, and those in round brackets are based on HAC (Newey-West) standard errors. The values in curly brackets are the values of the t-statistic in q = 4 group estimates of the corresponding predictive regressions for t-statistic robust inference approach. t-statistics with an asterisk are significant under the 5% significance level.

The t-statistics in square brackets are based on i.i.d. standard errors, and those in round brackets are based on HAC (Newey-West) standard errors. The values in curly brackets are the values of the t-statistic in q = 4 group estimates of the corresponding predictive regressions for t-statistic robust inference approach.

According to the tables, all alphas appear to be strongly significant for the three cryptocurrencies, and nearly none of the factors are significant, even when i.i.d. standard errors are used.¹³ Although the risk-free market returns are not statistically significant for these cryptocurrencies, we notice that the coefficients on them are negative in all models for the three cryptocurrencies considered. The overall conclusion using HAC inference methods is that Bitcoin, Ethereum, and Ripple prices are indifferent to changes in the stock market.

Remarkably, momentum (MOM) factors in Fama-French four-factor models are strongly significant for Bitcoin and Ethereum returns according to the t-statistic robust inference approaches. Momentum is the

---

¹³ According to HAC t-test, the investment factor (CMA) is significant (at 5% level) for Bitcoin returns in the Fama-French six-factor model and at 10% level in the five-factor model. The value premium (HML) is significant at 10% level for Ripple, also in the Fama-French five-factor and six-factor model. For Ethereum, none of the factors are significant.
factor that determines the rate of acceleration on stock price change, namely, the trend of prices. Usually, investors will increase the asset when its momentum is positive and decrease the asset when its momentum is negative. For Bitcoin, MOM has negative coefficients. Thus, when the momentum of stocks increases, the returns of Bitcoin will decrease correspondingly. Bitcoin is highly risky, and the demand for them is highly elastic. When stock prices show upward potentials, the relative yield of Bitcoin decreases compared to stocks. Therefore, the demand for cryptocurrencies decreases drastically, which causes the prices and returns of Bitcoin to decrease. In addition, Bitcoin can be used to hedge momentum risks in a diversified investment portfolio. However, interestingly, Ethereum has a positive MOM coefficient; this suggests that Ethereum investors have similar positions in the stock market and Ethereum returns comove more with high prior profit firms than low prior profit ones. Unlike Bitcoin, Ethereum does not intend to form a new monetary system but to facilitate contracts and decentralized applications on the Ethereum platform. One may analogize the Ethereum platform as the stock market but using Ether (ETH) instead of fiat money as the median of exchange. Thus, investors using the Ethereum platform tend to have business connections with the stock market, and they can invest more using Ether when the stock market has a upward moving trend.

Importantly, for Ripple, one observes a decrease in the significance of the value (HML) factor in Fama-French five- and six-factor models compared to the three-factor model according to the $t$-statistic robust inference. In contrast, its significance increases in the five- and six-factor models as compared to the three-factor model according to HAC tests (the corresponding values of HAC $t$-statistics increase and the HML factor becomes significant in the five- and six-factor models according to HAC $t$-tests although it was not significant in the three-factor model). According to the conclusions in [18], the HML becomes redundant in the five-factor model for stock returns that includes these factors. Notably, according to [18], it is highly correlated with the CMA with a correlation coefficient of $-0.7$.

The decrease in the significance of HML in five-factor model according to $t$-statistic robust inference approach as compared to HAC $t$-test accords with the conclusion in [18], and illustrates how the conclusions based on $t$-statistic robust inference approaches may be used to complement and improve upon the conclusions from the HAC methods.

### 3.3 Acceptance of Bitcoin

Liu and Tsyvinski [29,30] proxy the intrinsic/fundamental value of Bitcoin by the number of Bitcoin wallet users and use the ratio of the Bitcoin price to the number of wallet users as an analogue of the price-to-dividend ratio in predictive regressions for financial returns. This is motivated by the interpretation of the price-to-dividend ratio as a measure of the gap between an asset’s market and intrinsic/fundamental values (see [29,30] for the discussion). In this section, we present the analysis of robust predictive regressions for Bitcoin returns with the same price-to-dividend proxy used as a regressor. More precisely, the results are provided for robust inference approaches applied to regressions of weekly Bitcoin returns on the lagged ratio of the spot price of Bitcoin to the number of wallet users (the data on the latter number are updated weekly). It is natural to view the number of Bitcoin wallet users as a measure of acceptance of Bitcoin and to refer to the latter ratio as price-to-acceptance ratio, denoted by $(P/A)_t$.

We regress the future weekly Bitcoin returns ($R_{t+h}$, $h = 1,\ldots, 7$, i.e., up to 7-week ahead returns) on the current weekly price-to-acceptance ratio of Bitcoin $(P/A)_t$. The results of the predictive regression analysis are provided in Table 5. The number of users for Ethereum and Ripple is currently unavailable, so the price-to-acceptance ratios of them will not be discussed in this paper. It is interesting to observe that the cryptocurrency price-to-acceptance proxy is insignificant over all time horizons according to $t$-tests with

---

14 The use of the number of wallet users as a potential predictor for cryptocurrency returns and prices may be further motivated by valuation and pricing models for cryptocurrencies that naturally incorporate the number of users, measures of consumer adoption, and related variables, among other factors (see, for instance, [5,37], and references therein)
Table 5: Price-to-acceptance ratio of Bitcoin

|        | $R_{t-1}$ | $R_{t-2}$ | $R_{t-3}$ | $R_{t-4}$ | $R_{t-5}$ | $R_{t-6}$ | $R_{t-7}$ |
|--------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| $(P/A)_t$ | 0.113     | 0.001     | 0.207     | 0.163     | 0.140     | 0.303     | ~0.170    |
| i.i.d.  | 0.506     | 0.006     | 0.927     | 0.729     | 0.626     | 1.357     | ~0.761    |
| HAC     | 7.606*    | 0.078     | 12.882*   | 9.958*    | 8.156*    | 17.862*   | ~10.936*  |
| $t$-Statistic ($q = 4$) | 1.584 | 1.023 | 0.669 | 0.589 | 0.503 | 0.504 | 0.503 |

*The values in the first row are coefficients of price-to-acceptance ratios of Bitcoin. The $t$-statistics in the second row are based on i.i.d. standard errors, and those in the third are based on HAC (Newey-West) standard errors. The values in the fourth row are the values of the $t$-statistic in $q = 4$ group estimates of the corresponding predictive regressions for $t$-statistic robust inference approach. $t$-statistics with an asterisk are significant under the 5% significance level.

3.4 Google trends

Similar to [29,30], to measure how investors’ attention affects the returns of Bitcoin, we use the Google Trends of words “bitcoin,” “ripple,” and “ethereum” as proxies. In this section, we consider the weekly returns of the three cryptocurrencies. In order to maintain the consistency of Google Trends, we use the last day of the previous 180-day interval as the first day of the new 180-day interval, calculate the ratio of that overlapped day’s Google Trend, and use the ratio to normalize the new interval. Then, we calculate the average of every seven-day Google Trends and use these averages as proxies for weekly investors’ attention.¹⁶⁻¹⁷ We regress the future weekly cryptocurrency return ($R_{t+h}$, $i = 1, \ldots, 7$, i.e., up to 7-week ahead returns) on the current weekly Google Trend. The results of the predictive regression analysis are provided in Tables 6–8.

The $t$-tests based on (nonvalid) i.i.d. standard errors indicate the statistical significance of Google Trends of “bitcoin” and “ripple” over time horizon between 1 week and 7 weeks and the statistical significance of Google Trends of “ethereum” over time horizon between 1 week and 2 weeks. The conclusions based on i.i.d. standard errors suggest that the current Google Trends can predict crypto returns over at least future half-month period, and thus, the crypto price is affected by investors’ attention that may accord with a priori expectation.

However, the statistical significance of Google Trends disappears when HAC $t$-test or $t$-statistic robust inference approaches are used. Similar to the case of price-to-acceptance ratio analysis, the use of $t$-statistic

---

¹⁵ As discussed in Section 3.1, the fact that HAC $t$-statistics are larger than their i.i.d. counterparts points out negative autocorrelation in the data.

¹⁶ In their analysis, [29,30] construct the deviation of Google searches for the words “Bitcoin” in a given week compared to the average of those in the preceding four weeks. (Google search data for “Bitcoin” minus the average of previous four weeks, normalized.)

¹⁷ The conclusions on the significance of the coefficients in the considered predictive regressions using daily data are generally the same as those using weekly data reported in this section.
robust inference methods may be preferred in the settings due to relatively small samples of weekly observations used in the analysis. We cannot conclude that when investors’ attention is proxied by the Google Trends, Bitcoin, Ripple, and Ethereum returns are affected by such a factor.

4 Conclusion

In order to examine the predictability in cryptocurrency markets, this paper analyses four different models on cryptocurrency returns and three different t-tests. The regular i.i.d. t-test reported for illustrative
purposes is not valid due to dependence and heterogeneity in the time series of cryptocurrency returns and their predictive factors. This motivates the use of HAC t-tests that account for heteroskedasticity and autocorrelation in the data. Further, motivated by poor statistical properties of HAC t-tests in finite samples reported in previous studies, we also provide the analysis of predictive regressions for cryptocurrency returns using the t-statistic robust inference approaches developed in [22].

According to HAC t-tests, cryptocurrency returns have exposures to some of the factors in all of the models considered. For Bitcoin, momentum and price-to-acceptance ratio have a relatively strong influence and Ripple is relatively strongly influenced by its momentum. Ethereum appears to be not affected by the momentum and stock factors.

When t-statistic robust inference approaches are used, the number of significant exposures drastically decreases. None of the cryptocurrencies appear to be exposed to their momentum. In the stock factor models, one observes differences in factors that appear to be significant according to HAC t-tests and t-statistics robust inference approaches. In particular, the MOM is not significant for Bitcoin and Ethereum returns according to HAC t-test, but is it significant according to t-statistic robust inference. The conclusions on significance point out to the interaction between the cryptocurrency market and the stock market. We also observe the changes in the significance of the HML in different Fama-French models that appear to be in accordance with conclusion in [18] and that may indicate further advantages of complementing the results of HAC t-test with those based on t-statistic robust inference approaches.

For the two Bitcoin-specific factors, price-to-acceptance ratio and the Google trend, we do not find any significance at all for weekly return prediction using the t-statistic robust inference approaches. These approaches are preferred to HAC methods in the case of relatively small samples of weekly observations considered due to poor performance of the latter in finite samples.

Throughout the paper, we discuss the advantages and limitations of different approaches to robust inference under heterogeneity and dependence in data. Overall, the t-statistic robust inference approaches appear to perform better than the HAC t-tests in terms of pointing out interpretable economic conclusions.

The results and conclusions in the paper emphasize the necessity in the use of econometrically justified inference methods that account for autocorrelation and heterogeneity in observations. They further emphasize the usefulness of t-statistic robust inference approaches as complements to widely used HAC methods.

Further research on the topic may focus on the development of pricing models for cryptocurrencies incorporating the factors that appear to have predictive power for crypto prices and returns. It would also be of interest to provide econometrically justified and robust analysis of predictive regressions incorporating further factors, such as, importantly, the measures of liquidity, network security, and nonfundamental uncertainty that enter valuation models for cryptocurrencies (see [37,38]). It would also be of interest to consider applications of two-sample t-statistic robust inference approaches (see [23]) in robust tests of structural breaks in the coefficients of predictive regressions in cryptocurrency markets and their comparisons across different cryptocurrencies. The research in these directions is currently underway by the authors and co-authors.

Acknowledgments: We thank two anonymous referees, Andrea Buraschi, Emiliano Pagnotta, Artem Prokhorov and Johan Walden and the participants at the 1st Inaugural International Conference on Econometrics and Business Analytics (iCебA) and the seminars at CЕBA, St. Petersburg State University, and the joint meeting of the 2nd Workshop in Applied Econometrics and the VII International Conference on Modern Econometric Tools and Applications (Higher School of Economics, Moscow and Nizhny Novgorod, Russia) for helpful comments. Rustam Ibragimov gratefully acknowledges support provided by the Russian Foundation for Basic Research, Project No. 20-010-00960.

Conflict of interest: The authors state no conflict of interest.
Appendix

Tables A1 and A2

Table A1: Ripple and stock factors

|        | Excess Ripple return |
|--------|-----------------------|
|        | 3-F       | 4-F       | 5-F       | 6-F       |
| CAPM   |           |           |           |           |
| Mkt.RF | -0.142    | -0.138    | -0.147    | -0.137    |
|        | [1.213]   | [1.104]   | [0.927]   | [0.974]   |
|        | (-0.267)  | (-0.247)  | (-0.236)  | (-0.224)  |
|        | (0.154)   | (-0.146)  | (0.111)   | (-0.442)  |
| SMB    | 0.191     | 0.210     | -0.090    | -0.057    |
|        | [0.822]   | [0.912]   | [0.318]   | [0.434]   |
|        | (0.221)   | (0.234)   | (-0.099)  | (-0.061)  |
|        | (-0.009)  | (-0.109)  | (-0.371)  | (-0.461)  |
| HML    | -1.076    | -1.040    | -1.728    | -1.672    |
|        | [0.068]   | [0.254]   | [-0.733]  | [-0.544]  |
|        | (-1.406)  | (-1.295)  | (-1.945)  | (-1.849)  |
|        | (-2.812)  | (-2.027)  | [-2.209]  | [-1.965]  |
| MOM    | 0.001     | 0.001     |           |           |
|        | [0.509]   | [0.609]   |           |           |
|        | (0.141)   | (0.255)   |           |           |
|        | (-0.587)  | (0.158)   |           |           |
| RMW    |           |           | -1.406    | -1.410    |
|        |           |           | [-1.640]  | [-1.649]  |
|        |           |           | (-1.137)  | (-1.135)  |
|        |           |           | [-1.246]  | [-1.232]  |
| CMA    | 1.562     | 1.589     |           |           |
|        | [1.180]   | [1.217]   |           |           |
|        | (1.212)   | (1.261)   |           |           |
|        | [1.610]   | [1.701]   |           |           |
| Alpha  | 0.015     | 0.015     | 0.015     | 0.015     |
|        | [2.926*]  | [2.920*]  | [3.012*]  | [2.991*]  |
|        | (3.175*)  | (3.769*)  | (3.525*)  | (3.306*)  |
|        | (4.303*)  | (4.508*)  | (4.426*)  | (4.366*)  |

The excess Ripple returns are defined as the difference between Ripple returns and the risk-free rate. The values in the first row of each stock factor are coefficients of each stock factor. The t-statistics in square brackets are based on i.i.d. standard errors, and those in round brackets are based on HAC (Newey-West) standard errors. The values in curly brackets are the values of the t-statistic in q = 4 group estimates of the corresponding predictive regressions for t-statistic robust inference approach. t-statistics with an asterisk are significant under the 5% significance level.
Table A2: Ethereum and stock factors

|          | Excess Ethereum return |
|----------|-------------------------|
|          | 3-F | 4-F | 5-F | 6-F |
| Mkt.RF   | −0.542 | −0.527 | −0.506 | −0.529 | −0.495 |
|          | [0.029] | [0.288] | [0.516] | [0.298] | [0.590] |
| SMB      | (−0.780) | (−0.800) | (−0.782) | (−0.776) | (−0.738) |
|          | (0.005) | (−0.444) | (−0.484) | (−0.670) | (−0.733) |
| HML      | −0.460 | 0.093 | −0.257 | −0.128 |
|          | [−1.644] | [−1.005] | [−1.896] | [−1.267] |
|          | (−0.029) | (0.120) | (−0.324) | (−0.160) |
|          | (0.163) | (−0.075) | (0.116) | (−0.064) |
| MOM      | 0.003 | 0.004 | 0.003 |
|          | [2.518] | [2.640] |
|          | (0.634) | (0.720) |
|          | (−3.534) | (−1.037) |
| RMW      | −1.227 | 0.093 |
|          | [−0.927] | [−0.932] |
|          | (−1.151) | (−0.145) |
|          | (−0.202) | (−0.171) |
| CMA      | 1.567 | 1.677 |
|          | [1.029] | [1.301] |
|          | (1.142) | (1.181) |
|          | (−0.083) | (−0.061) |
| Alpha    | 0.013 | 0.013 |
|          | [3.393] | [3.418] |
|          | (3.426) | (3.731) |
|          | (2.478) | (2.364) |

The excess Ethereum returns are defined as the difference between Ethereum returns and the risk-free rate. The values in the first row of each stock factor are coefficients of each stock factor. The t-statistics in square brackets are based on i.i.d. standard errors, and those in round brackets are based on HAC (Newey-West) standard errors. The values in curly brackets are the values of the t-statistic in q = 4 group estimates of the corresponding predictive regressions for t-statistic robust inference approach. t-statistics with an asterisk are significant under the 5% significance level.

References

[1] Alexander, C., & Dakos, M. (2020). A critical investigation of cryptocurrency data and analysis. *Quantitative Finance*, 20, 173–188.
[2] Ammous, S. (2018). *The Bitcoin standard: The decentralized alternative to Central Banking*. Hoboken, New Jersey: John Wiley & Sons.
[3] Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariant matrix estimation. *Econometrica*, 59, 817–858.
[4] Baur, D. G., Dimpfl, T., & Kuck, K. (2018). Bitcoin, gold and the US dollar-A replication and extension. *Finance Research Letters*, 25, 103–110.
[5] Bolt, W., & van Oordt, M. R. C. (2019). On the value of virtual currencies. *Journal of Money, Credit and Banking*, 52, 835–862.
[6] Borri, N. (2019). Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance*, 50, 1–19.
[7] Borri, N., & Shakhnov, K. (2020). The cross-section of cryptocurrency returns. Working paper, LUISS University and the University of Surrey.
[8] Bouoiyour, J., & Selmi, R. (2016). Bitcoin: A beginning of a new phase? *Economics Bulletin*, 36, 1430–40.
[9] Campbell, J. Y., Lo, A. W., & MacKinlay, A. (1997). The econometrics of financial markets. Princeton: Princeton University Press.

[10] Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of Finance, 52, 57–82.

[11] Cheung, A., Roca, E., & Su, I.-J. (2015). Crypto-currency bubbles: An application of the Phillips-Shi-Yu (2013) methodology on Mt. Gox bitcoin prices. Applied Economics, 47, 2348–58.

[12] Conley, T. G. (1999). GMM estimation with cross sectional dependence. Journal of Econometrics, 92, 1–45.

[13] Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. Quantitative Finance, 1, 223–236.

[14] Dyhrberg, A. (2016). Hedging capabilities of bitcoin. Is it the virtual gold? Finance Research Letters, 16, 139–144.

[15] Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). Modelling extremal events for insurance and finance. New York: Springer.

[16] Esarey, J., & Menger, A. (2019). Practical and effective approaches to dealing with clustered data. Political Science Research and Methods, 7, 541–559.

[17] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33, 3–56.

[18] Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116, 1–22.

[19] Fama, E. F., & French, K. R. (2018). Choosing factors. Journal of Financial Economics, 128, 234–252.

[20] Ibragimov, M., Ibragimov, R., & Walden, J. (2015). Heavy-tailed distributions and robustness in economics and finance. Lecture notes in statistics (Vol. 214). New York: Springer.

[21] Ibragimov, R., & Müller, U. K. (2007). t-Statistic based correlation and heterogeneity robust inference. Harvard Institute of Economic Research Discussion Paper No. 2129.

[22] Ibragimov, R., & Müller, U. K. (2010). t-Statistic based correlation and heterogeneity robust inference. Journal of Business and Economic Statistics, 28, 453–468.

[23] Ibragimov, R., & Müller, U. K. (2016). Inference with few heterogeneous clusters. Review of Economics and Statistics, 98, 83–96.

[24] Ibragimov, R., Parlour, C. A., & Walden, J. (2019). Cryptocurrencies: Intrinsic value and bubbles. Working paper, Imperial College Business School and Haas School of Business, the University of California at Berkeley.

[25] Ilyidogan, E. (2019). Incentive mechanism and economic model of blockchain based cryptocurrencies. Working paper, Imperial College Business School.

[26] Katsiampa, P. (2017). Volatility estimation for bitcoin: A comparison of garch models. Economics Letters, 158, 3–6.

[27] Krückeberg, S., & Scholz, P. (2020). Decentralized efficiency? arbitrage in bitcoin markets. Financial Analysts Journal, 76, 135–152.

[28] Lintilhac, P. S., & Tourin, A. (2017). Model-based pairs trading in the bitcoin markets. Quantitative Finance, 17, 703–716.

[29] Liu, Y., & Tsyvinski, A. (2018). Risks and returns of cryptocurrency. NBER working paper No. 24877.

[30] Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency. The Review of Financial Studies, 34(6), 2689–2727.

[31] Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common risk factors in cryptocurrency. The Journal of Finance, 77(2), 1133–1177.

[32] Makarov, I., & Schoar, A. (2019). Trading and arbitrage in cryptocurrency markets. Journal of Financial Economics, 135, 293–319.

[33] McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative risk management: Concepts, techniques and tools (Revised Edition). Springer.

[34] Nadarajah, S., & Chu, J. (2017). On the inefficiency of Bitcoin. Economics Letters, 150, 6–9.

[35] Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55, 703–708.

[36] Nguyen, L. H., Chevapatrakul, T., & Yao, K. (2020). Investigating tail-risk dependence in the cryptocurrency markets: A LASSO quantile regression approach. Journal of Empirical Finance, 5, 333–355.

[37] Pagnotta, E. (2018). Bitcoin as decentralized money: Prices, mining and network security. Working paper, Imperial College Business School.

[38] Pagnotta, E., & Buraschi, A. (2018). An equilibrium valuation of Bitcoin and decentralized network assets. Working paper, Imperial College Business School.

[39] Phillips, P. C. B., Shi, S., & Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? International Economic Review, 52, 201–226.

[40] Phillips, P. C. B., Shi, S., & Yu, J. (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500. International Economic Review, 56, 1043–1077.

[41] Poyser, O. (2017). Exploring the determinants of Bitcoin’s price: An application of Bayesian structural time series. arXiv preprint arXiv: http://arXiv.org/abs/arXiv:1706.01437.

[42] Rachev, S., Menn, C., & Fabozzi, F. J. (2005). Fat-tailed skewed asset return: Implications for risk management, portfolio selection, and option pricing. Hoboken: John Wiley & Sons.

[43] Rogers, W. H. (1993). Regression standard errors in clustered samples. Stata Technical Bulletin, 13, 19–23.

[44] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. The Journal of Finance, 19, 425–442.
[45] Sovbetov, Y. (2018). Factors influencing cryptocurrency prices: Evidence from Bitcoin, Ethereum, Dash, Litecoin, and Monero. *Journal of Economics and Financial Analysis*, 2, 1–27.

[46] Stock, J. H., & Watson, M. W. (2019). *Introduction to econometrics* (4th ed.). Harlow: Pearson.

[47] Trimborn, S., & Härdle, W. K. (2018). CRIX an Index for cryptocurrencies. *Journal of Empirical Finance*, 49, 107–122.

[48] Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters*, 148, 80–82.

[49] Valkanov, R. (2003). Long-horizon regressions: Theoretical results and applications. *Journal of Financial Economics*, 68, 201–232.

[50] Wei, W. C. (2018). Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, 168, 21–24.

[51] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817–838.

[52] Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. In L. David, & K. Chuen (Eds.), *Handbook of digital currency* (pp. 31–43). Amsterdam: Elsevier.