Replication of Crypto Factor Analysis

Daniel Liang
Master of Science in Finance
University of Illinois
`yuanhui5@illinois.edu`

Zixuan Feng
Master of Science in Finance
University of Illinois
`zf20@illinois.edu`

Xiyu Cai
Master of Science in Finance
University of Illinois
`xiyucai2@illinois.edu`

Weihua Zhang
Master of Science in Finance
University of Illinois
`weihuaz2@illinois.edu`

November 3, 2024

Abstract

1 Introduction

Cryptocurrencies such as Bitcoin are virtual assets that are usually not cleared on centralized exchanges, ensuring their transparency and preventing them from being easily stolen and attacked. In the past few years, these assets have caught the increasing attention of investors because of their extraordinary returns, volatility, and leverage. Two of the most widely discussed topics for them are pricing models and investing preferences over other asset classes.

Because of their vague intrinsic values and virtual designs, some fundamental analyses that apply to common equities will not work for them. Wang and Chong (2021) extended the traditional Fama–French three-factor model to replace the value-to-growth factor with the coin-to-token factor as a proxy and include some volatility, liquidity and macro factors. Poongodi et al. (2020) tried to use market data and some machine learning models, such as linear regression and support vector machine, to make some price predictions of Ethereum. Although these two pieces of research investigate the fair price of cryptocurrencies from two distinct angles, fundamental factors and market technical analysis, there is still a huge room for rational investors to dig into the fair value discovery of these virtual assets.

Besides, many investors are trying to figure out whether cryptocurrency will be a good choice for mid to long-term investments over other investment options, such as bonds and stocks. Because of the non-normal distribution of returns, Han et al. (2021) adopted one non-parametric method, almost stochastic dominance, to compare the performances of factor portfolios of cryptocurrencies to those of stock indexes and bonds.

This report replicates the findings of Han et al. (2021) on cryptocurrency factor portfolios, using data from Wharton Research Data Services (WRDS) and Binance. Following the original methodology, we analyze market data for 82 cryptocurrencies (2014-2024) and trading data for 10 cryptocurrencies (2017-2024). The study employs R to validate the original conclusions on portfolio performance and dominance over traditional financial benchmarks, examining the roles of risk premiums and mispricing.

This paper is organized as follows. **Section 2** summarises the paper intended to replicate, *Cryptocurrency Factor Portfolios: Performance, Decomposition and Pricing Models* by Han et al. (2021), and discusses some of their interesting findings. **Section 3** reviews relevant literature and research on our topics, such as factor models and almost stochastic dominance. **Section 4** describes the data and summary statistics. **Section 5** presents the results generated by similar approaches adopted by Han et al. (2021) with the data described in **Section 4**. **Section 6** discusses some limitations and potential downsides of this project and provides some guidelines for future improvements. **Section 7** concludes this project and compares our findings to those of Han et al. (2021).

2 Paper Summary

Han et al. (2021) argued that cryptocurrency factor portfolios outperform traditional assets, through a combination of almost stochastic dominance (ASD) and factor modelling. The ASD method includes almost first-degree stochastic dominance (AFSD) and almost second-degree stochastic dominance (ASSD), addressing the limitations of traditional metrics, such as mean-variance analysis, to non-normal distributions of returns, which was proven by the paper. AFSD will examine whether a portfolio is preferred without knowledge of investors' risk tolerances, while ASSD will assume the investors are risk-averse to analyse the dominance of a portfolio over another. To identify dominance, the paper set critical thresholds at 5.9% for AFSD and 3.2% for ASSD.

The study constructed 27 factors, divided into 4 main categories - size, momentum, volume and volatility - shown in **Table 1** in **Section 5**. They divided each factor into 5 quantiles respectively, examined the returns of each quantile, and constructed a long-short portfolio by longing the quantile and shorting quantile 1 with weights based on market capitalization. Then, they utilized ASD to compare these factor portfolios' performances to those of S&P 500, of 10-year U.S, of 1-month T-bill and of the cryptocurrency index. Treasury bonds, with five different holding periods, 4-week, 13-week, 26-week, 52-week and 78-week. For the S&P 500, no dominance was found for the holding periods of 4-week and 13-week. Ten factors (LPRC, MAXPRC, MOM1, MOM2, MOM3, ROMO1, RMOM2, RMOM3, RMOM4 and STDPRCVOL) showed AFSD and five of them (LPRC, MAXPRC, MOM1, MOM3 and RMOM2) showed ASSD for longer holding horizon. Extending the investment horizon, an increasing number of factors showed dominance, as high returns would compensate for the high volatility. For 52-week, besides of above factors, seven new factors (MARCAP, VOLPRC, VOLSCALE, RETVOL, RETSKEW, MEANABS and IDIOVOL) showed AFSD and eight new factors (MARCAP, MOM2, RMOM-1,3,4 VOLPRC, STDPRCVO, MEANABS and IDIOVOL) showed ASSD. For 78-week, VOL became an AFSD factor and two more factors (RETVOL and RETSKEW) showed ASSD, with a total of 18 and 16 factors respectively. For ASD against T-bond, T-bill and cryptocurrency index, similar trends and factors were shown. After all, nine dominance factors (MOM1, MOM2, MOM3, RMOM1, RMOM2, RMOM3, VOLPRC, STDPRCVOL and MEANABS) were found, dominating all four benchmarks for both AFSD and ASSD.

After obtaining nice dominance factors, they analysed the contributions of long-leg and short-leg to the dominating performances of long-short portfolios. Long legs of all dominance factors with some of the non-dominance factors dominated all four benchmarks for AFSD and ASSD with a holding period of 78 weeks, and all of the long legs of dominance factors showed AFSD and most of them showed ASSD for a 52-week holding period. For a shorter holding period, almost no AFSD and ASSD were found, except against the cryptocurrency index for a 26-week holding period. However, for short legs, none of the factors showed dominance at any holding horizons, indicating the dominance of long-short portfolios was mostly contributed by the long leg of the portfolios.

Lastly, the paper tried to examine the source outperformance. The paper applied a three-factor model for the cryptocurrency market (including market, size, and momentum factors) to explain portfolio returns. To more accurately capture the sources of returns, the model was extended to include additional mispricing factors and fundamental cryptocurrency factors (such as electricity cost and computing power), aiming to further explore whether the abnormal returns were generated by risk premiums or market mispricing. Four dominant factor portfolios had statistically significant positive alphas and low R-squared values when using a three-factor model, suggesting that their outperformance could not be fully explained by the model and was likely generated by mispricing.

In summary, the paper found nine dominant factors that dominated four selected benchmarks over longer holding horizons. The long leg contributed most to the dominance of long-short portfolios constructed by these factors. Mispricing rather than risk premia was the main reason for the outperformance of the cryptocurrency portfolios.

3 Hypothesis Overview

As discussed in **Section 2**, three objectives and hypotheses were discussed and tested by Han et al. (2021).

Hypothesis 1: Cryptocurrency factor portfolios perform better than stock the S&P, US T-bonds, US T-bills and cryptocurrency index at 4-week, 13-week, 26-week, 52-week and 78-week investment horizons.

To test this hypothesis, Han et al. (2021) established long-short value-weighted portfolios for 27 factors respectively and compared their performances by ASD against four benchmarks. The factors included the various factors used to construct the portfolios, such as size, momentum, volume, and volatility. Using a new dataset, this project will adopt a similar approach and factors to construct portfolios for the investment horizons mentioned above and the 1-week investment horizon and compare the performances against the S&P 500.

Hypothesis 2: The dominance of cryptocurrency factor long-short portfolios mainly comes from the long legs rather than the short legs.

Han et al. (2021) utilized ASD to compare long-only factor portfolios and short-only factor portfolios against four benchmarks for the same investment horizons above. This project will also adopt the same techniques to compare long-only factor portfolios and short-only factor portfolios against S&P 500, long-short portfolios and each other to examine the dominance of one leg.

Hypothesis 3: The outperformance of cryptocurrency factor portfolios comes from mispricing rather than the risk premia.

The paper by Han et al. (2021) constructed two regression models with and without four mispricing factors and two cryptocurrency fundamental factors (electricity and computing power) and examined the alphas and r-squares of each model. This project will not test this hypothesis due to the lack of data for electricity and computing power and the relatively small dataset.

4   Literature Review

A large body of literature has explored factors that generate excess returns. Fama and French (1993) introduced the three-factor model, which expanded on the Capital Asset Pricing Model (CAPM) by adding two additional factors: size (SMB, small minus big) and value (HML, high minus low), demonstrating that small-cap stocks and high book-to-market ratio stocks tend to outperform large-cap stocks and low book-to-market ratio stocks. Building on this, Carhart (1997) added a momentum factor (MOM) to explain the continuation of short-term stock returns. With the development of the cryptocurrency market, Liu et al. (2019) proposed a three-factor model tailored to cryptocurrencies, identifying market size and momentum as key factors that significantly affect cryptocurrency returns. Other studies, such as Liu & Tsyvinski (2020), also indicated that cryptocurrency returns are primarily driven by network factors and investor attention. A recent systematic literature review by Peng et al. (2023) further identified key factors influencing cryptocurrency pricing, including market supply and demand, technological advancements, and social media.

When the data exhibits high skewness, stochastic dominance better reflects the performance of assets compared to mean-variance-based regression. Bali, Brown, and Demirtas (2013) explored the performance of hedge funds compared to stocks and bonds, utilizing both classical and ASD rules to assess dominance. The study found that the return distributions of hedge fund portfolios, as well as equity and bond returns, deviate significantly from normality, making classical selection rules inadequate for explaining investor preferences. In contrast, ASD provides a more suitable framework since it does not rely on parametric specifications of preferences or assumptions about asset returns. Overall, the findings suggest that ASD is more effective than classical methods in evaluating hedge fund performance under non-normal return distributions. Post (2003) demonstrated that the Fama and French market portfolio is inefficient when compared to benchmark portfolios based on market capitalization and book-to-market equity ratios, with Second-Order Stochastic Dominance (SSD) providing a better evaluation framework. In this paper, the authors also used First-Order Stochastic Dominance (FSD) and Second-Order Stochastic Dominance (SSD) to measure the performance of each factor.

Long-short zero-investment strategies were widely adopted and discussed in factor portfolios. Israel and Moskowitz (2013) examined the role of shorting, firm size, and time in explaining market anomalies. They found that long positions account for the majority of size, value, and momentum profits, while shorting plays a lesser role, especially for momentum. Additionally, the value premium decreases with firm size and is weakest among the largest stocks, whereas momentum profits show no consistent relationship with firm size. Overall, the premium for momentum strategies tends to be higher than that for value, particularly in large-cap stocks. Similarly, Blitz et al. (2019) found that factor premiums are stronger on the long side. Frazzini and Pedersen (2014) showed that high-beta portfolios have lower risk-adjusted returns than low-beta ones. Barroso and Santa-Clara (2015) highlighted that momentum strategies have high returns but also

significant risks, which can be managed through volatility scaling. Daniel and Moskowitz (2016) discussed momentum crashes, noting that momentum profits primarily come from long positions.

5 Replication

This project intends to replicate the methodology and factors of Han et al. (2021) by using relatively recent data to examine the consistency of their findings for out-of-sample data. This section will mainly discuss the data used by this project, codes adopted for the replication and results obtained.

5.1 Data

Han et al. (2021) collected data for the top 400 cryptocurrencies based on their market capitalizations in USD over the period 1st January 2014 to 31st December 2019 on opening price, high price, low price, close price, volume and market capitalization from Coinmarketcap.com, S&P 500 and 10-year T-bond from CRSP and risk-free rate from Kenneth French's website. This project will use more recent but smaller data, including daily market capitalization data of 82 cryptocurrencies from 1st January 2014 to 30th September 2024 extracted from Binance, daily trading data of 10 cryptocurrencies (ada, avax, bnb, btc, doge, eth, sol, ton, trx and xrp) corresponding to up to 4 stablecoins (busd, usdc, usdt and usdp) from 17th August 2017 to 1st October 2024 from Binance, daily risk-free rate from 14th October 2013 to 30th August 2024 from WRDS, daily S&P 500 data from 10th January 2017 to 8th September 2023 from WRDS, and weekly S&P 500 data from 3rd January 2016 to 13th October 2024 from WRDS. After obtaining raw data, these data will be merged according to dates.

Han et al. (2021) established 27 factors for cryptocurrencies, shown in **Table 1**. This project replicates most of these factors and builds zero-investment long-short portfolios based on them. Compared to the replicated paper, the age factor was removed due to lack of data. The factors BETA, BETA^2, IDIOVOL, and DELAY were also excluded because the date coverage in this project is short, leading to potential data loss and reduced reliability. Then, this project uses these factors to examine ASD against .

5.2 Analysis of Dominance Factors

5.2.1 Data Preparation

Figure 2 shows the data preparation. Before the calculation of the return of value-weighted cryptocurrency index, data1, which contains market capitalization data for top cryptocurrencies, and data2, which contains trading data for cryptocurrencies are cleaned to have the same asset names and one column of market capitalization data. Two datasets then will be merged based on asset names and dates.

```
#1 Data Cleaning
data1 <- as.data.frame(readRDS("MktCapTopCoins.rds"))
head(data1)

data2 <- as.data.frame(readRDS("Top10_cryptoStablePairs.rds"))
head(data2)

distinct(data2,market)

str(data1)
str(data2)
sum(is.na(data1["CapMrktEstUSD"]))
sum(is.na(data2))

#new col: merge the right mktcap
data1$CapMkt <- ifelse(!is.na(data1$CapMrktCurUSD), data1$CapMrktCurUSD, data1$CapMrktEstUSD)
#head(data1)
data1 <- data1[, c("asset", "time", "CapMkt")]

# before merge, get "asset"
data2$asset <- sapply(strsplit(as.character(data2$market), "-"), function(x) x[2])

merged_data <- merge(data1, data2, by = c("asset","time"))
sum(is.na(merged_data))
```

Figure 1: Codes for Data Preparation - Cryptocurrencies Data

To calculate the return of the value-weighted (market-capitalization-weighted) cryptocurrency index, daily log returns will be calculated as **Equation 1** for each cryptocurrency based on its closing price and daily market-capitalization-weighted index return will be computed as **Equation 2** as the weighted sum of the log

Table 1: The constructions of different factors. (Han et al., 2021)

Category	Factor Used	Definition
Size	MARCAP	Log last day market capitalization in the portfolio formation week
Size	LPRC	Log last day price in the portfolio formation week
Size	MAXPRC	The maximum price of portfolio formation week
Size	Age	The number of existent weeks that listed on Coinmarketcap.com
Momentum	MOM1	One-week momentum
Momentum	MOM2	Two-week momentum
Momentum	MOM3	Three-week momentum
Momentum	MOM4	Four-week momentum
Momentum	MOM8	Eight-week momentum
Momentum	RMOM1	One-week risk-adjusted momentum based on Sharpe ratio
Momentum	RMOM2	Two-week risk-adjusted momentum based on Sharpe ratio
Momentum	RMOM3	Three-week risk-adjusted momentum based on Sharpe ratio
Momentum	RMOM4	Four-week risk-adjusted momentum based on Sharpe ratio
Momentum	RMOM8	Eight-week risk-adjusted momentum based on Sharpe ratio
Volume	VOL	Log average daily volume in the portfolio formation week
Volume	VOLPRC	Log average daily volume times price in the portfolio formation week
Volume	VOLSCALE	Log average daily volume times price then divided by market capitalization in the portfolio formation week
Volatility	RETVOL	The standard deviation of daily returns in the portfolio formation week
Volatility	RETSKEW	The skewness of daily returns in the portfolio formation week
Volatility	RETKURT	The kurtosis of daily returns in the portfolio formation week
Volatility	MAXRET	The maximum daily return of the portfolio formation week
Volatility	STDPRCVOL	Log standard deviation of dollar volume in the portfolio formation week
Volatility	MEANABS	The mean absolute daily return divided by dollar volume in the portfolio formation week
Volatility	BETA	The regression coefficient of β_{MKTI} in $R_i - R_f = \alpha_i + \beta_{MKTI}MKT + \epsilon_i$. The model is estimated by using the daily return of previous 365 days before formation week.
Volatility	BETA^2	Beta squared
Volatility	IDIOVOL	The idiosyncratic volatility is measured as the standard deviation of the residual after estimating $R_i - R_f = \alpha_i + \beta_{MKTI}MKT + \epsilon_i$. The model is estimated by using the daily return of previous 365 days before formation week.
Volatility	DELAY	The improvement of R^2 in $R_i - R_f = \alpha_i + \beta_{MKTI}MKT + \beta_{MKTI-1}MKT_{-1} + \beta_{MKTI-2}MKT_{-2} + \epsilon_i$ compared to regression that only uses MKT, where MKT_{-1} and MKT_{-2} are lagged one- and two-day market index return. The model is estimated by using the daily return of previous 365 days before formation week.

daily returns for each cryptocurrency, weighted by shares of their market capitalizations on the date. **Figure 3** shows the code in R for this procedure.

$$\text{log\_daily\_return} = \log\left(\frac{\text{price\_close}}{\text{previous day price\_close}}\right) \qquad \text{(Equation 1)}$$

$$\text{index\_return} = \sum(\text{log\_daily\_return} \times \text{weight}) \qquad \text{(Equation 2)}$$

Figure 3 and **Figure 4** shows the Market-Capitalization-Weighted Index Return and the Market-Capitalization-Weighted Index Value (with a base of 100 on 17th August 2017), respectively. These plots together provide insights into the daily fluctuations and longer-term trends in the top cryptocurrency market. The index return plot shows the day-to-day variability, while the index value plot demonstrates how the overall market capitalization changes cumulatively over time.

```
#2 crypto log return
unique(merged_data$asset)    #top 10 crypto

merged_data <- merged_data %>%
  group_by(asset) %>%
  arrange(time) %>%
  mutate(log_daily_return = log(price_close / lag(price_close)))

##Crypto index return
index_data <- merged_data %>%
  group_by(time) %>%
  mutate(total_cap = sum(CapMkt, na.rm = TRUE)) %>%   #total market cap
  ungroup()

index_data <- index_data %>%
  mutate(weight = CapMkt / total_cap)         #weight calculation

index_data <- index_data %>%
  group_by(time) %>%
  summarise(index_return = sum(log_daily_return * weight, na.rm = TRUE))
```

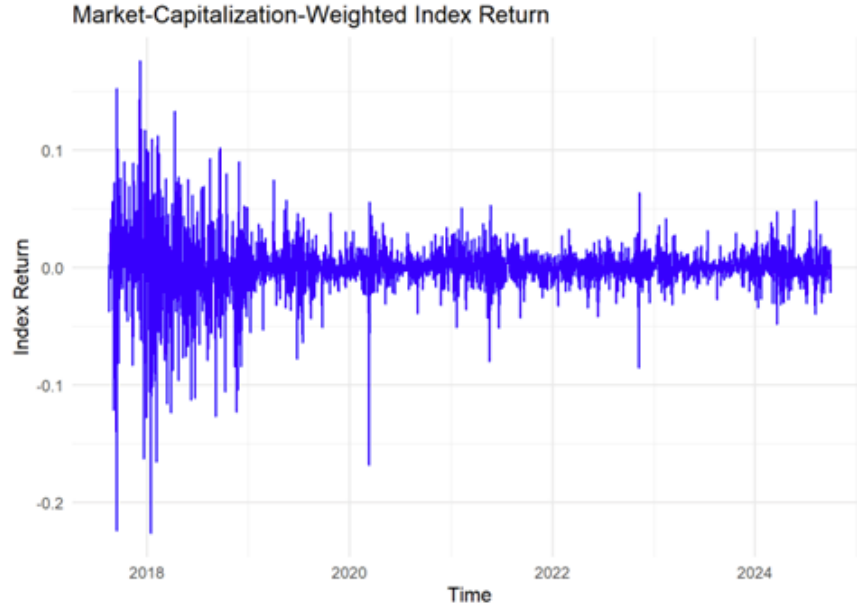Figure 2: Codes for Calculation of the Value-Weighted Cryptocurrency Index Return

Figure 3: Daily Returns of a Hypothetical Index based on the Market Capitalization

5.2.2 Benchmark: Risk-free and S&P500

Han et al. (2021) used S&P500, risk-free rates and value-weighted cryptocurrency index as benchmarks against cryptocurrency factor portfolios. Thus, this project combines the other two benchmarks' data with the value-weighted cryptocurrency index return constructed as **Section 5.2.1** to compare the performance of cryptocurrencies with traditional financial benchmarks of risk-free rates and S&P500.

As shown in **Figure 5**, the S&P 500, the value-weighted cryptocurrency index and risk-free rates show no dominance over each other. Besides, the results by Han et al. (2021) indicated these three benchmarks yielded similar outcomes. To simplify the process, this project sets the S&P 500 as the only benchmark against cryptocurrency factor portfolios, assuming that dominance over the S&P 500 will indicate dominance over others.

Figure 4: The Index Value of the Market-Capitalization-Weighted Index

```
##### 2.Benchmark: Riskfree and S&P500

rf <- read.csv("Rf.CSV")
sp500 <- read.csv("SP500.csv")


rf$time <- ymd(rf$time)
sp500$time <- ymd(sp500$time)

benchmark <- merge(rf, sp500, by = "time") %>%
  select(time,RF,sprtrn)


index_data$time <- ymd(index_data$time)
crypo_index <- merge(index_data,benchmark,by = "time") %>%
  select(time,index_return,RF,sprtrn)

head(crypo_index)
```

5.2.3   ECDFs Calculating

To calculate the Empirical Cumulative Distribution Functions (ECDFs) for the S&P 500 returns (sp_cdf), risk-free rate (rf_cdf), and the crypto index returns (crypto_cdf), an ECDF function is applied to return the cumulative probability up to a given value in the input data. For example, for a value x, the ECDF function returns the proportion of data points that are less than or equal to x. The plot compares the distributions of returns for the crypto index, S&P 500, and the risk-free rate. By observing the shape of the CDFs, we can identify which asset is riskier (steeper increases indicate more concentrated values), which has more variability, and how they perform relative to each other.

7

```
> head(crypo_index)
        time index_return   RF     sprtrn
1 2017-08-17  0.000000000 0.004 -0.015437
2 2017-08-18 -0.037746151 0.004 -0.001835
3 2017-08-21  0.012494450 0.004  0.001163
4 2017-08-22 -0.008892551 0.004  0.009941
5 2017-08-23  0.018408599 0.004 -0.003454
6 2017-08-24  0.040773924 0.004 -0.002074
```

```
###3 CDF
sp_cdf <- ecdf(crypo_index$sprtrn)
rf_cdf <- ecdf(crypo_index$RF)
crypto_cdf <- ecdf(crypo_index$index_return)

cdf_data <- data.frame(
  value = c(crypo_index$sprtrn, crypo_index$RF, crypo_index$index_return),
  cdf = c(sp_cdf(crypo_index$sprtrn), rf_cdf(crypo_index$RF), crypto_cdf(crypo_index$index_return)),
  variable = rep(c("S&P500", "RF", "Crypto_index"), each = length(crypo_index$sprtrn))
  )

# plot
ggplot(cdf_data, aes(x = value, y = cdf, color = variable)) +
  geom_step() +
  labs(title = "Cumulative Distribution Functions",
       x = "Return Value",
       y = "Cumulative Probability") +
  scale_color_manual(values = c("blue", "darkgreen", "orange")) +
  theme_minimal() +
  theme(legend.title = element_blank())
```
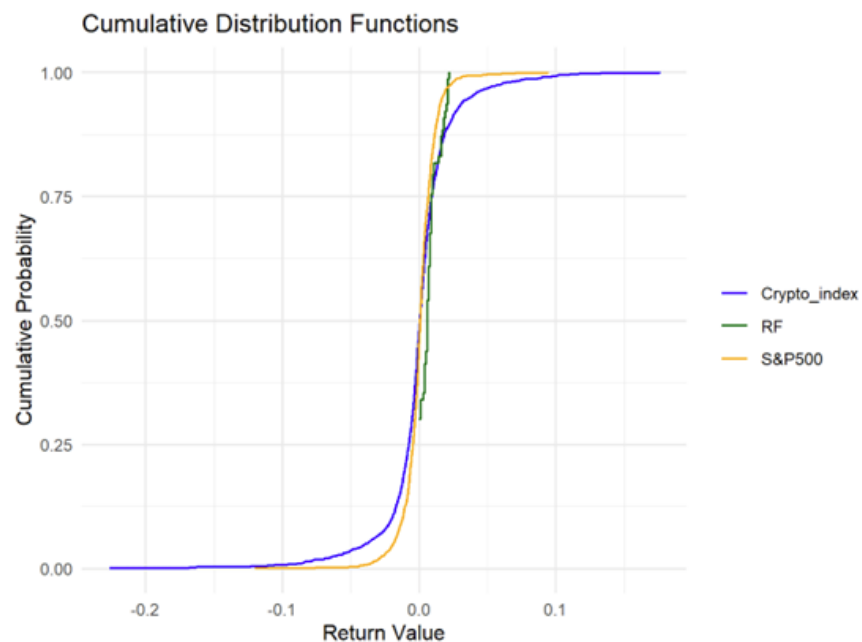

Figure 5: Cumulative Distribution Function for Risk free, SP500 and crypto index

8

```r
#####Into function for further use
cdf_plot <- function(data1, data2, labels) {
  # Combine the two dataframes
  combined_data <- data.frame(
    value = c(data1, data2),
    variable = rep(labels, each = length(data1))
  )
  # Generate ECDFs for each variable
  ecdf1 <- ecdf(data1)
  ecdf2 <- ecdf(data2)
  # Create a data frame with the CDF values
  cdf_data <- data.frame(
    value = c(data1, data2),
    cdf = c(ecdf1(data1), ecdf2(data2)),
    variable = rep(labels, each = length(data1))
  )
  # Plot
  print(ggplot(cdf_data, aes(x = value, y = cdf, color = variable)) +
    geom_step() +
    labs(title = "Cumulative Distribution Functions",
         x = "Return Value",
         y = "Cumulative Probability") +
    scale_color_manual(values = c("blue", "darkgreen")) +
    theme_minimal() +
    theme(legend.title = element_blank()))
}
#cdf_plot(crypo_index$sprtrn, crypo_index$RF, c("S&P500", "RF"))
```

Then the cdf_plot function is introduced as a reusable tool for comparing the distribution of two different sets of data. This helps the later modeling of AFSD and ASSD.

5.2.4   AFSD & ASSD

Almost Stochastic Dominance is a concept used in decision theory and finance to compare distributions of returns or outcomes, allowing for some tolerance of violations in the preference ordering. To test whether AFSD and ASSD exist, we calculate the ratio of violation area over the total enclosed area for each portfolio and compare these ratios (ϵ_1 for AFSD and ϵ_2 for ASSD) to determine if they are lower than AFSD (5.9

For AFSD, the FSD violation range is given by:

$$R_1(F_H, F_L) = \{s \in (r_1, r_2) : F_L(s) < F_H(s)\}$$

The empirical violation area is defined as:

$$\epsilon_1 = \frac{\int_{R_1} [F_H(s) - F_L(s)]\, ds}{\int_{\min}^{\max} [F_H(s) - F_L(s)]\, ds}$$

If $\epsilon_1 < 5.9\%$, AFSD exists.

9

```
#4.1 AFSD
AFSD <- function(X, Y, interval = 0.001) {

  F_H <- ecdf(X)
  F_L <- ecdf(Y)
  s_range <- seq(min(c(X, Y)), max(c(X, Y)), by = interval)

  numerator <- 0   # Numerator: interval where the integral does not satisfy the condition
  denominator <- 0  # Denominator: integral absolute difference

  for (i in 2:length(s_range)) {
    s <- s_range[i]
    s_prev <- s_range[i - 1]

    diff_HL <- F_H(s) - F_L(s)
    diff_HL_prev <- F_H(s_prev) - F_L(s_prev)
    delta_s <- s - s_prev

    abs_diff <- (abs(diff_HL) + abs(diff_HL_prev)) / 2 * delta_s
    denominator <- denominator + abs_diff

    # Numerator: The integral is computed only if the condition is not satisfied
    #(F_H(s) > F_L(s))
    if (diff_HL > 0) {
      num_diff <- (diff_HL + diff_HL_prev) / 2 * delta_s
      numerator <- numerator + num_diff
    }
  }

  ratio <- numerator / denominator

  if(ratio>0.5){
    ratio <- 1-ratio
  }

  return(list(
    numerator = numerator,
    denominator = denominator,
    ratio = ratio
  ))
}
```

For ASSD, the SSD violation range is given by:

$$R_2(F_H, F_L) = \left\{ s \in (F_H, F_L) : \int_0^r [F_L(s) - F_H(s)]\, ds < 0 \right\}$$

The empirical violation area is defined as:

$$\epsilon_2 = \frac{\int_{R_2} [F_H(s) - F_L(s)]\, ds}{\int_{\min}^{\max} [F_H(s) - F_L(s)]\, ds}$$

If $\epsilon_2 < 3.2\%$, ASSD exists.

10

```
#4.2 ASSD

ASSD <- function(X, Y, interval = 0.001) {

  F_H <- ecdf(X)
  F_L <- ecdf(Y)

  s_range <- seq(min(c(X, Y)), max(c(X, Y)), by = interval)

  numerator <- 0
  denominator <- 0

  integral_FX <- 0
  integral_FY <- 0

  for (i in 2:length(s_range)) {
    s <- s_range[i]
    s_prev <- s_range[i - 1]

    diff_HL <- F_H(s) - F_L(s)
    diff_HL_prev <- F_H(s_prev) - F_L(s_prev)

    delta_s <- s - s_prev
    integral_FX <- integral_FX + (F_H(s) + F_H(s_prev)) / 2 * delta_s
    integral_FY <- integral_FY + (F_L(s) + F_L(s_prev)) / 2 * delta_s

    # Denominator: Calculates the integral of the absolute difference
    abs_diff <- (abs(integral_FX - integral_FY) + abs(diff_HL_prev)) / 2 * delta_s
    denominator <- denominator + abs_diff

    # Denominator: Calculates the integral of the absolute difference
    abs_diff <- (abs(integral_FX - integral_FY) + abs(diff_HL_prev)) / 2 * delta_s
    denominator <- denominator + abs_diff
    # Numerator: The integral is computed only when second-order stochastic
    #dominance is not satisfied (integral_FX > integral_FY)
    #Molecules: Calculate integrals only when second-order random dominance is
    #not satisfied (integrat_FX>integrat_FY)
    if (integral_FX > integral_FY) {
      num_diff <- (integral_FX - integral_FY + diff_HL_prev) / 2 * delta_s
      numerator <- numerator + num_diff
    }
  }

  ratio <- numerator / denominator

  if(ratio>0.5){
    ratio <- 1-ratio
  }

  return(list(
    numerator = numerator,
    denominator = denominator,
    ratio = ratio
  ))
}
```

5.2.5 Factor Calculation

Asset characteristics are calculated based on 22 factors related to key aspects of size, momentum, volume, and volatility.

Momentum Factors: Momentum measures the rate of price change over time and is calculated using lagged values of the asset prices.

- **Lagged Values**: Lag1, Lag2, Lag3, Lag4, Lag8 represent the asset's lagged prices over different time periods (7, 14, 21, 28, and 56 days).
- **Momentum (mom)**: These factors are calculated as the difference between the current price and the lagged prices over different time horizons (mom1, mom2, mom3, mom4, mom8). This shows how much the asset's price has changed over those time intervals.

Volatility Factors: Volatility factors measure the risk and dispersion of returns over various time periods.

- **Standard Deviation (std)**: Standard Deviation are calculated over rolling windows of 7, 14, 21, 28, and 56 days (std1, std2, std3, std4, std8). This provides an estimate of how much the price fluctuates within the specified period.
- **Risk-Adjusted Momentum (rmom)**: Momentum is adjusted by dividing by the corresponding standard deviation to account for risk. These are also calculated over different time windows (rmom1, rmom2, rmom3, rmom4, rmom8)

Size Factors: The size of an asset is generally measured by its market capitalization (*CapMkt*), which is the product of price and the number of shares outstanding.

- **Log-Price (lprc)**: The natural logarithm of the price to stabilize the variance.

- **Max Price (maxprc)**: The maximum price observed over a rolling 7-day window.

Volume Factors: Volume factors reflect the trading activity and liquidity of the asset.

- **Log-Volume (vol)**: The natural logarithm of the average volume over a rolling 7-day window.

- **Volume-Price Product (volprc)**: The product of volume and price, representing the total traded value.

- **Volume-Scaled Price (volscale)**: Volume-price product divided by market capitalization to assess the liquidity relative to the asset size.

The process will build portfolios described in **Setion 5.2.6** and compare the portfolios based on weekly return calculations to evaluate if one strategy or factor dominates another.

Quantile: Quantile formation involves sorting assets based on factor values and dividing them into equalsized groups. This method is used to classify assets according to their factor characteristics. For each factor, the assets are divided into five quantiles, and this categorization helps analyze the performances of portfolios constructed from each quantile.

By categorizing the assets into quantiles for each of these factors, the code enables analysis of portfolio returns, risks, and other metrics. Quantile-based sorting allows for examining if certain factor characteristics (e.g., high momentum or low volatility) result in superior performance.

5.2.6 Portfolio Building

This project then analyzes financial portfolios, focusing on constructing portfolios based on different factors such as size, momentum, volatility, and volume. Portfolio returns are calculated using the value-weighted methods as Han et al. (2021), with the ultimate goal of assessing the performance of different quantile-based portfolios.

Step 1: Portfolio Building
Portfolios are built based on different factors, such as size, momentum, volatility, and volume, combined with returns of the following week.

Step 2: Value-Weighted Portfolio Calculation
This project calculates returns for a value-weighted portfolio as the replicated paper did. The weights of each asset (weights_port) are calculated by dividing the market value of each asset by the total market value for the period. This normalization ensures that the sum of the weights equals 1. If holding_period is 1, it uses the default returns (e.g., 1-week weekly returns). For longer holding periods, it uses cumulative returns over that period. This approach ensures that larger assets (by market value) have a more significant impact on the portfolio's returns, following a value-weighted methodology.

Step 3: Summary Calculations for Different Factors
The average returns are calculated for different quantiles (1 to 5) across various factors.

The Quantile-Based Long/Short Strategy:

- The project divides each factos into five quantiles and the **Table 2** shows the returns for each quantile from 1 to 5.

- **Long Positions**: Assets in the top quantile (5).

- **Short Positions**: Assets in the bottom quantile (1).

- The column (5-1) computes portfolio returns as the difference between the returns from long and short positions.

Table 2: Quantiles for the factors

	1	2	3	4	5	5-1
return_maxret	-0.023867625	-5.051694e-03	0.0042286835	0.019828017	0.08439965	0.108267277
meanabs	0.028590067	1.122526e-02	0.0203354148	0.024668671	0.01333350	-0.015256565
return_retkurt	0.020529818	1.038705e-02	0.0074547599	0.003765309	0.02915543	0.008625609
return_retskew	0.006646157	-1.495630e-04	0.0027632592	0.027389337	0.04086133	0.034215172
return_retvol	-0.003479828	1.181149e-05	0.0002414904	0.009836626	0.07250270	0.075982533
stdprcvol	0.011253833	1.725408e-02	0.0091490198	0.033103162	0.01358263	0.002328798

	1	2	3	4	5	5-1
mom1	-0.003327996	0.0038378029	0.007257764	0.04650253	0.04150269	0.04483069
mom2	-0.004454876	0.0035915450	0.024453987	0.03132944	0.03114343	0.03559831
mom3	-0.001193049	0.0088756042	0.027616219	0.02554294	0.03256921	0.03376226
mom4	0.007066979	0.0098438222	0.019729658	0.02935637	0.02838154	0.02131456
mom8	0.006727528	0.0134684199	0.025734144	0.02959754	0.02147988	0.01475235
rmom1	-0.045146292	-0.0087186656	0.023455778	0.03104889	0.09689537	0.14204166
rmom2	-0.027465950	-0.0085896236	0.006204910	0.02267248	0.09184784	0.11931379
rmom3	-0.024381104	-0.0118076128	0.009543391	0.01936385	0.08775118	0.11213229
rmom4	-0.020748661	-0.0003192778	0.005171271	0.02146868	0.07746622	0.09821488
rmom8	-0.017141282	-0.0008870510	0.008680772	0.01481967	0.09421955	0.11136083

	1	2	3	4	5	5-1
size	0.01168492	0.01767259	0.02948269	0.008758709	0.01392489	0.002239973
lprc	0.03143679	0.01124451	0.01640280	0.020063821	0.01369361	-0.017743185
maxprc	0.03143461	0.01136381	0.01845733	0.018946010	0.01372685	-0.017707765

	1	2	3	4	5	5-1
vol	0.01288615	0.023274081	0.02152722	0.009517958	0.03305016	0.020164009
volprc	0.03134705	0.011586255	0.01696765	0.020307837	0.01372685	-0.017620202
volscale	0.02009964	0.009158605	0.02127446	0.009630901	0.01320121	-0.006898431

5.2.7 AFSD and ASSD testing and plotting

This analysis involves testing AFSD and ASSD across different factors and holding periods, using log returns for the S&P 500 index as a benchmark.

Step 1: Set the Benchmarks of Risk-Free and S&P 500
Firstly, the following code shows ways to read and prepare weekly log return data for the S&P 500. This step converts the Date column to a Date format, computes the weekly log returns, and adjusts the dates.

```
##Benchmark weekly
s_p500_weekly <- read.csv("S_P_500_Weekly.csv")
s_p500_weekly$Date <- as.Date(s_p500_weekly$Date, format = "%m/%d/%Y")
s_p500_weekly$Price <- as.numeric(gsub(",", "", s_p500_weekly$Price))
s_p500_weekly <- s_p500_weekly %>%
  rename(week = Date)
s_p500_weekly <- s_p500_weekly %>%
  mutate(log_return = log(Price / lag(Price)))
s_p500_weekly <- s_p500_weekly %>%
  select(week, log_return) %>%
  filter(!is.na(log_return))
# debug, correct the dates between data
s_p500_weekly <- s_p500_weekly %>%
  mutate(week = week + 1)
```

Step 2: AFSD and ASSD Testing

Different factor-based long-short portfolio returns will be compared with the S&P 500 returns. AFSD and ASSD are used to measure the dominance relationships between the long-short portfolio performance and the benchmark.

The following codes take the size based portfolios as an example to calculate the AFSD and ASSD. Same techniques and methodologies will be applied to momentum, volatility and volume based portfolios.

```
## AFSD ASSD testing
### Size
Size_ASD <- data.frame("AFSD_Ratio"=1:length(row.names(Summary_Size)),"ASSD_Ratio"=1:length(row.names(Summary_Size)),row.names = row.names(Summary_Size))
size_port$week <- as.Date(size_port$week, format = "%Y-%m-%d")
match_dates <- merge(s_p500_weekly, size_port[1], by="week")
for(i in 1:length(row.names(Summary_Size))){
  factors <- row.names(Summary_Size)[i]
  if(factors=="size"){
    factors <- ""
  }
  long_short_performance <- value_weighted_factor_portfolio(size_port,factors,5) - value_weighted_factor_portfolio(size_port,factors,1)

  Size_ASD[i,1] <- AFSD(match_dates$log_return,long_short_performance)[3]
  Size_ASD[i,2] <- ASSD(match_dates$log_return,long_short_performance)[3]
  cdf_plot(long_short_performance, match_dates$log_return, c(factors,"S&P 500"))
}
```

Step 3: Factor Analysis by Holding Periods

After constructing and analysing portfolios with a holding period of 1 week, this project extends the holding period similar to Han et al. (2021) and analyzes portfolios' performances over these holding periods (4 weeks, 13 weeks, 26 weeks, 52 weeks, and 78 weeks). The following codes take the size based portfolios as an example to calculate the portfolios performances for holding periods longer than one week. Same techniques and methodologies will be applied to momentum, volatility and volume based portfolios.

```
### For 4 weeks, 13-week, 26-week, 52-week and 78-week
match_dates <- match_dates %>%
  mutate(log_return_4 = lead(slide_dbl(log_return, sum, .before = 3, .complete = TRUE),3),
         log_return_13 = lead(slide_dbl(log_return, sum, .before = 12, .complete = TRUE),12),
         log_return_26 = lead(slide_dbl(log_return, sum, .before = 25, .complete = TRUE),25),
         log_return_52 = lead(slide_dbl(log_return, sum, .before = 51, .complete = TRUE),51),
         log_return_78 = lead(slide_dbl(log_return, sum, .before = 77, .complete = TRUE),77))

momentum_port <- momentum_port %>%
  mutate(across(names(rtns)[2:11],
                list("4" = ~ lead(slide_dbl((.x+1), prod, .before = 3, .complete = TRUE)-1,3),
                     "13" = ~ lead(slide_dbl((.x+1), prod, .before = 12, .complete = TRUE)-1,12),
                     "26" = ~ lead(slide_dbl((.x+1), prod, .before = 25, .complete = TRUE)-1,25),
                     "52" = ~ lead(slide_dbl((.x+1), prod, .before = 51, .complete = TRUE)-1,51),
                     "78" = ~ lead(slide_dbl((.x+1), prod, .before = 77, .complete = TRUE)-1,77)),
                .names = "{.col}_{.fn}"))
```
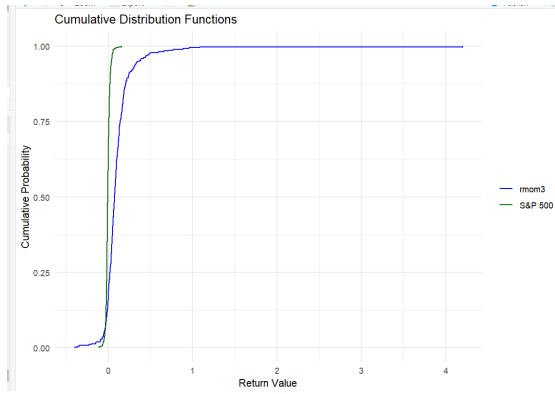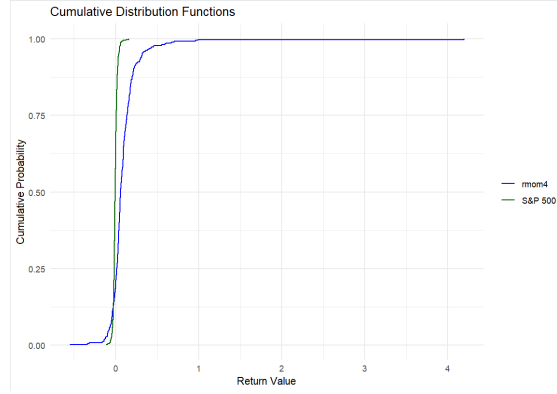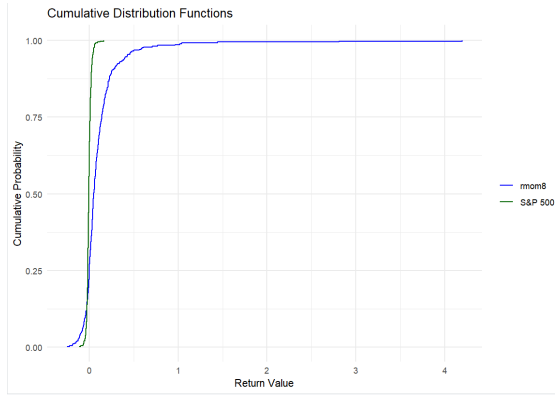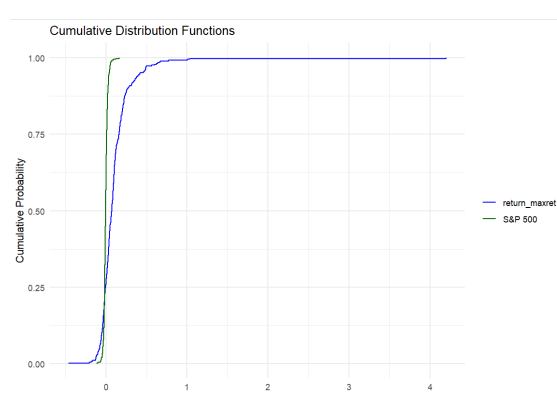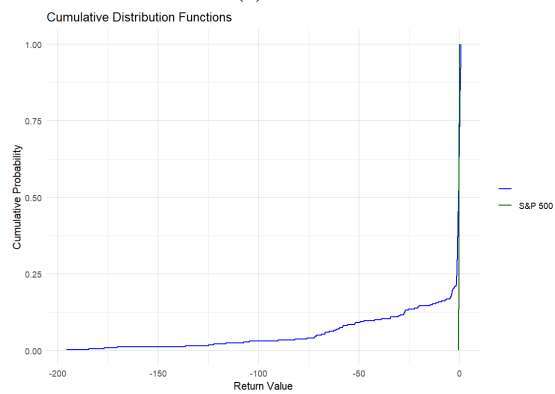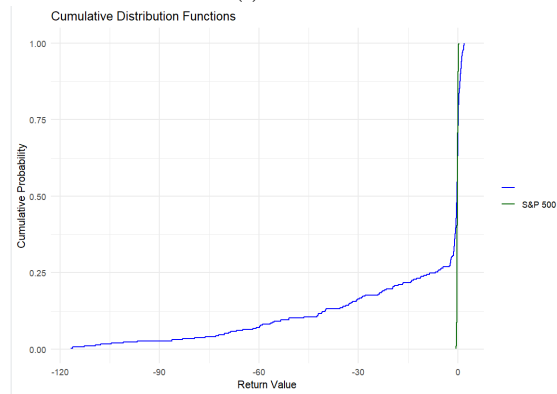
14
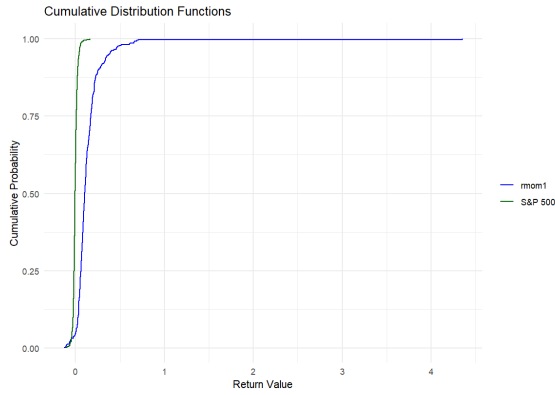
(a) mom1
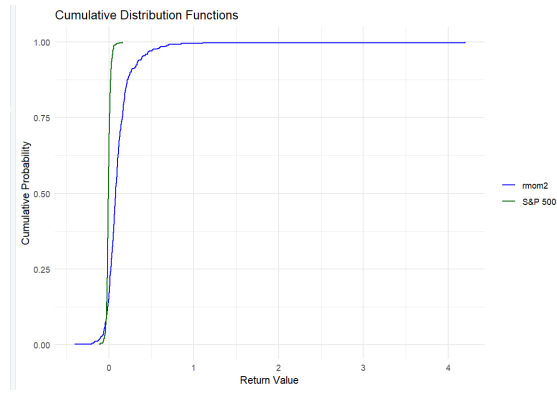
(b) mom2

(c) mom3

(d) mom4

(e) mom8

(f) maxret

(g) marketcap-52weeks

(h) marketcap-78weeks

Figure 6: CDF of factor portfolios against benchmark

Step 4: Identifying Dominant Factors
After calculations of ϵ_1 for AFSD and ϵ_2 for ASSD mentioned in **Section 5.2.4** for each long-short portfolio for each factor, the following codes test whether factors exhibit dominance according to AFSD and ASSD criteria for a specified holding period. This project uses similar thresholds (5.9% for AFSD and 3.2% for ASSD) to Han et al. (2021) as mentioned before in **Section 2** to determine dominance. The **Figure 6** shows AFSD graphs for the dominant factors with corresponding dominant holding periods. Suggested by Han et al. (2021), the direction of dominance of long-short portfolios will not be a matter, because switching the long-short quantiles will solve this issue. Based on the data used by this project, rmom1 - rmom8 and maxret shows dominance against S&P 500 for the holding period of one week, and marketcap shows dominance the holding periods of 52 weeks and 78 weeks.

```
### Names of Dominant Factors
for(j in 0:5){
  if(j==0){holding_period <- 1}else{
  holding_period <- list_period[j]}
  for(i in 1:length(row.names(Summary_Size))){
    factors <- row.names(Summary_Size)[i]
    if(Size_ASD[i,(2*j+1)] <0.059){
      if(Size_ASD[i,(2*j+2)]<0.032){
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD and ASSD"))
      } else{
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD but not ASSD"))
      }
    }
  }
  for(i in 1:length(row.names(Summary_momentum))){
    factors <- row.names(Summary_momentum)[i]
    if(Momentum_ASD[i,(2*j+1)] <0.059){
      if(Momentum_ASD[i,(2*j+2)]<0.032){
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD and ASSD"))
      } else{
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD but not ASSD"))
      }
    }
  }
  for(i in 1:length(row.names(Summary_volatility))){
    factors <- row.names(Summary_volatility)[i]
    if(Volatility_ASD[i,(2*j+1)] <0.059){
      if(Volatility_ASD[i,(2*j+2)]<0.032){
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD and ASSD"))
      } else{
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD but not ASSD"))
      }
    }
  }
  for(i in 1:length(row.names(Summary_volume))){
    factors <- row.names(Summary_volume)[i]
    if(Volume_ASD[i,(2*j+1)] <0.059){
      if(Volume_ASD[i,(2*j+2)]<0.032){
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD and ASSD"))
      } else{
        print(paste0(factors," for a holding period of ", as.character(holding_period)," weeks dominates the AFSD but not ASSD"))
      }
    }
  }
}
```

5.3 Contributions of Dominances

After identifying the dominant factors, this project tries to figure out whether long-legs, short-legs or both contribute to the dominance of these dominant factors.

The following codes take rmom1 as an example, creating factor-based long-only, short-only and long-short portfolios. The codes are used to create a long-only portfolio by investing in assets in the top quantile (5), a short-only portfolio by shorting assets in the bottom quantile (1) and a long-short portfolio by combining two portfolios. Same techniques will be applied to rmom2 - rmom8 for one week investment horizon and CapMkt for 52-week and78-week investment horizons. Then, the AFSD and ASSD are utilized to calculate the Almost Stochastic Dominance statistics to test whether:

- The long-only portfolio dominates the benchmark.

- The short-only portfolio dominates the benchmark.

- The short-only portfolio dominates the long-only portfolio.

- The short-only portfolio dominates the long-short portfolio.

- The long-only portfolio dominates the long-short portfolio.

```
### rmom1 1 weeks
long_only_performance <- value_weighted_factor_portfolio(momentum_port,"rmom1",5)
short_only_performance <- -value_weighted_factor_portfolio(momentum_port,"rmom1",1)
long_short_performance <- long_only_performance+short_only_performance
benchmark <- match_dates$log_return
AFSD_stats <- AFSD(benchmark,long_only_performance)[3]
ASSD_stats <- ASSD(benchmark,long_only_performance)[3]
if(AFSD_stats <0.059){
  if(ASSD_stats <0.032){
    print("rmom1: long_only donminate sp500 AFSD and ASSD.")
  } else{
    print("rmom1: long_only donminate sp500 AFSD but not ASSD.")
  }
} else{
  print("rmom1: long_only not donminate sp500")
}
```

5.3.1 Results

- **Momentum Factors (rmom1, rmom3, rmom4, rmom8):**
    - **Long-Only vs. S&P 500**: For all momentum factors tested, the long-only strategy does not dominate the S&P 500 for both AFSD and ASSD, indicating that a long-only approach may not outperform the benchmark.
    - **Short-Only vs. S&P 500**: Similarly, the short-only strategy fails to dominate the S&P 500, suggesting it does not provide better returns than the benchmark.
    - **Short-Only vs. Long-Only**: In each case, the short-only strategy shows a significant dominant relationship with the long-only strategy for both AFSD and ASSD criteria. After checking the patterns of each portfolio, this project suggests that shorting low-ranked assets yields inferior performance than buying high-ranked assets in the momentum context.
    - **Short-Only vs. Long-Short**: The short-only strategy also shows a significant dominant relationship with the long-short strategy across all momentum factors, indicating that pure shorting is significantly less effective than combining long and short positions.
    - **Long-Short vs. Long-Only**: In some cases (like rmom1), the long-short strategy shows a significant dominant relationship with the long-only strategy, suggesting that including short positions worsens performance over a pure long strategy.

    Generally speaking, the insignificant dominant relationships for long-leg and short-leg with the benchmark, S&P 500, are a result of huge left-tail losses. However, this effect will be mitigated when combining two legs, of which a negative correlation between two legs may be a cause. Besides, long-leg consistently yields better performances, potentially indicating that long-leg contributes more to the long-short strategies.

- **Size Factor for Longer Holding Periods:**
    - As suggested by the patterns of AFSD between CapMkt and the benchmark, shorting the large-size cryptocurrencies (quantile 1) and longing the small-size cryptocurrencies (quantile 5) will yield superior returns than S&P 500. Thus, the long-only strategy here will refer to longing the small-size cryptocurrencies and the short-only strategy will refer to shorting the large-size cryptocurrencies (quantile 1).
    - For both 52-week and 78-week investment horizons, the long-only strategy dominates the S&P 500 in both AFSD and ASSD for a 52-week holding period, while the short-only strategy does not dominate the S&P, indicating that the long-leg contributes more than the short-leg due to huge right-tail positive returns, even though the short leg will further improve the right-tail positive returns by a small extent.

These results highlight the importance of the contributions of the long legs, even though the short legs will improve the performances by either mitigating huge left-tail losses or improving the huge right-tail returns.

6 Future Work

The original paper is vulnerable to overfitting issues. Firstly, because a large number of factors and several investment horizons were adopted, the dominance factors may be detected by chance. This issue was mitigated by adopting a six-year panel data with 400 cryptocurrencies. However, this project replicates those factors and methodologies with a significantly small dataset, which will be much more vulnerable to overfitting.

Secondly, a six-year observation may not be enough to witness a full picture of the market cycle. The resulting dominance of the original paper may be a result of weak stock market conditions, such as that in 2018, or a strong cryptocurrency market. Given the high volatility of cryptocurrency assets, it is unclear whether a portfolio constructed from historical data will continue to perform well under new future conditions. Furthermore, the original paper constructed portfolios by allowing both buying and short-selling without costs, but it disregards real-world transaction costs and short-selling constraints, which could also impact the portfolio's performance.

Future research could extend the analysis by examining longer observation horizons to include a full market cycle, employing more cryptocurrencies and incorporating different asset benchmarks, such as commodities or real estate, to better understand the performance of cryptocurrency factors over time and across a broader investment landscape. Up-to-date data and a larger number of cryptocurrencies will reduce the possibility of overfitting and provide more reliable information about the current market. Additionally, investigating the impact of various market conditions (e.g., bull vs. bear markets) on factor dominance could provide insights into the conditional performance of cryptocurrency investments. Besides, setting constrains, such as transaction costs and short-selling limitations, will gain more reliable results for real-world strategies.

7 Conclusions

This analysis employs non-parametric almost stochastic dominance (AFSD and ASSD) to assess the performance of various portfolio strategies based on different factors, including momentum, size, volume and volatility. It aims to determine if these strategies can outperform the S&P 500 benchmark over different time horizons, ranging from 1 to 78 weeks.

Our findings indicate that risk-adjusted momentum-based strategies (rmom1, rmom3, rmom4, and rmom8) exhibit dominant factors in the short term (1 week). Furthermore, long-short strategies generally perform better than long-only strategies, particularly for short-term momentum factors. In the longer term (52 and 78 weeks), the size factor (CapMkt) shows significant dominance over the benchmark. Although a few dominant factors were found, the number was significantly less than that found by the original paper, which found nine dominant factors across 26-week to 78-week holding horizons. This may indicate a changing market environment, an overfitting issue of the original paper, or a small dataset of this project.

After obtaining the dominant factors with corresponding dominant investment horizons, this project evaluates the relative contributions of long legs and short legs to long-short portfolios constructed by these dominant factors. This project gains similar results as the original paper, suggesting the long legs contribute more than the short legs. This result does not undervalue the importance of short legs, which improve the performances of long legs by either mitigating huge left-tail losses or improving the huge right-tail returns. Especially, for short-term momentum strategies, long-only approaches do not dominate the benchmark, highlighting the importance of short legs in mitigating huge losses and the limitations of long-only strategy.

Further research should explore a large dataset with longer up-to-date observation horizons and more cryptocurrencies to mitigate overfitting and obtain more instructive information. Moreover, different benchmarks should be examined to check the dominances cryptocurrency factor strategies across larger investment landscape. Thirdly, incorporating some real-world constraints into the investigation will be helpful to gain more practical guidances.

References

Bali, Turan G., Stephen J. Brown, and K. Ozgur Demirtas (Aug. 2013). "Do Hedge Funds Outperform Stocks and Bonds?" en. In: *Management Science* 59.8, pp. 1887–1903. ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.1120.1689. URL: https://pubsonline.informs.org/doi/10.1287/mnsc.1120.1689 (visited on 10/29/2024).

Barroso, Pedro and Pedro Santa-Clara (Apr. 2015). "Momentum has its moments". en. In: *Journal of Financial Economics* 116.1, pp. 111–120. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2014.11.010. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304405X14002566 (visited on 10/29/2024).

Blitz, David, Guido Baltussen, and Pim Van Vliet (Oct. 2020). "When Equity Factors Drop Their Shorts". en. In: *Financial Analysts Journal* 76.4, pp. 73–99. ISSN: 0015-198X, 1938-3312. DOI: 10.1080/0015198X.2020.1779560. URL: https://www.tandfonline.com/doi/full/10.1080/0015198X.2020.1779560 (visited on 10/29/2024).

Carhart, Mark M. (Mar. 1997). "On Persistence in Mutual Fund Performance". en. In: *The Journal of Finance* 52.1, pp. 57–82. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/j.1540-6261.1997.tb03808.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1997.tb03808.x (visited on 10/29/2024).

Daniel, Kent and Tobias J. Moskowitz (Nov. 2016). "Momentum crashes". en. In: *Journal of Financial Economics* 122.2, pp. 221–247. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2015.12.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304405X16301490 (visited on 10/29/2024).

Fama, Eugene F. and Kenneth R. French (Feb. 1993). "Common risk factors in the returns on stocks and bonds". en. In: *Journal of Financial Economics* 33.1, pp. 3–56. ISSN: 0304405X. DOI: 10.1016/0304-405X(93)90023-5. URL: https://linkinghub.elsevier.com/retrieve/pii/0304405X93900235 (visited on 10/29/2024).

Frazzini, Andrea and Lasse Heje Pedersen (Jan. 2014). "Betting against beta". en. In: *Journal of Financial Economics* 111.1, pp. 1–25. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2013.10.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304405X13002675 (visited on 10/29/2024).

Han, Weihao et al. (2021). "Cryptocurrency Factor Portfolios: Performance, Decomposition and Pricing Models". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3857315. URL: https://www.ssrn.com/abstract=3857315 (visited on 10/31/2024).

Israel, Ronen and Tobias J. Moskowitz (May 2013). "The role of shorting, firm size, and time on market anomalies". en. In: *Journal of Financial Economics* 108.2, pp. 275–301. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2012.11.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304405X12002401 (visited on 10/29/2024).

Li, Jingming et al. (Feb. 2019). "Energy consumption of cryptocurrency mining: A study of electricity consumption in mining cryptocurrencies". en. In: *Energy* 168, pp. 160–168. ISSN: 03605442. DOI: 10.1016/j.energy.2018.11.046. URL: https://linkinghub.elsevier.com/retrieve/pii/S0360544218322503 (visited on 10/29/2024).

Liu, Yukun and Aleh Tsyvinski (May 2021). "Risks and Returns of Cryptocurrency". en. In: *The Review of Financial Studies* 34.6. Ed. by Itay Goldstein, pp. 2689–2727. ISSN: 0893-9454, 1465-7368. DOI: 10.1093/rfs/hhaa113. URL: https://academic.oup.com/rfs/article/34/6/2689/5912024 (visited on 10/29/2024).

M., Poongodi et al. (Jan. 2020). "Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system". In: *Computers & Electrical Engineering* 81, p. 106527. ISSN: 0045-7906. DOI: 10.1016/j.compeleceng.2019.106527. URL: https://www.sciencedirect.com/science/article/pii/S0045790618331343 (visited on 10/31/2024).

Post, Thierry (Oct. 2003). "Empirical Tests for Stochastic Dominance Efficiency". en. In: *The Journal of Finance* 58.5, pp. 1905–1931. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/1540-6261.00592. URL: https://onlinelibrary.wiley.com/doi/10.1111/1540-6261.00592 (visited on 10/29/2024).

Wang, Qiyu and Terence Tai-Leung Chong (July 2021). "Factor pricing of cryptocurrencies". In: *The North American Journal of Economics and Finance* 57, p. 101348. ISSN: 1062-9408. DOI: 10.1016/j.najef.2020.101348. URL: https://www.sciencedirect.com/science/article/pii/S1062940820302308 (visited on 10/31/2024).