MASTERARBEIT / MASTER'S THESIS

Master Thesis Title

verfasst von / submitted by

Jorge Eduardo Frías Navarrete

Submitted in total fulfillment of the requirement of the degree of:

Master of Science (WU)

Matrikelnummer / student ID number:	012329686
Studium / degree programme:	Quantitative Finance
University:	Vienna University of Economics and Business
Betreut von / Supervisor:	Univ.Prof. David Preinerstorfer, Ph.D.

Wien, August 2025 / Vienna, August 2025

Table of contents

List of Figures

List of Tables

Preface

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

```
1 + 1
```

```
[1] 2
```

1 Introduction

This is a book created from markdown and executable code.

See (knuth84?) for additional discussion of literate programming.

```
1 + 1
```

```
[1] 2
```

2 Summary

In summary, this book has no content whatsoever.

```r
1 + 1
```

```
[1] 2
```

3 Data

3.1 Data extraction and sample construction

I collect daily cryptocurrency data on open, high, close, and low (OHCL) prices, 24-hour volume, and market capitalization (calculated as the cryptocurrency's USD price multiplied by its circulating supply) from CoinCodex, a website-data provider that gathers and aggregates data from more than 400 exchanges. I extract the data using the CoinCodex API as follows:

1. I access the list of all available cryptocurrencies and extract each cryptocurrency short-name, also called "slug". At the time of writing this, there are 14,775 cryptocurrencies listed, excluding stablecoins.

2. I construct an URL for each cryptocurrency and use it to retrieve the metadata from the API. I parsed the JSON API response into a dataframe and extract the OHCL prices, volume, and market capitalization daily data. I exclude those observations with non-zero or missing values in any of these fields.

Out of the 14,775 cryptocurrencies listed, only XXX (COMPLETE ++++++++++++++) contained available data. Following the methodology of Bianchi & Babiak (2021) and Mercik et al. (2025), I apply a series of filters to clean the dataset and remove possible innacuracies:

1. Non-zero and missing values. As mentioned earlier, I exclude observations with non-zero or missing prices, volume, or market capitalization.
2. Small market-capitalization. Similar to Liu et al. (2022), I consider only cryptocurrencies with a market capitalization greater than one million USD. Therefore, I remove observations for coins whose market capitalization falls below this minimum threshold, which allows for the possibility that a coin may become "small" after a certain period or event.
3.

The sample period ranges from January 1st, 2014, to May 31st, 2025.

From Mercik (2025) Exclude: - Stable coins such as USDT, USDC, DAI, etc. regardless of whether they are centrally managed or algorithmically stabilized. - Exclude coins pegged to or reflecting the value of precious metals - Coins used as guarantees for derivatives platforms.

Organize week in the following way: the first seven days of the year forms the first week, and the first 51 weeks of the year consists of 7 days each. The 52th week of the year consists of the last eight days and, in case of a leap year (as 2016, 2020, and 2024), of nine days.

Similar to Liu et al, I construct a daily cryptocurrency market return as the value-weighted average return of all the cryptocurrencies in the sample. For cryptocurrencies $i = 1, ..., N$, the daily market return at time t is computed as:

$$r_t^M = \frac{\sum_{i=1}^{N} r_{it} \cdot marketcap_{it}}{\sum_{i=1}^{N} marketcap_{it}}$$

The cryptocurrency market excess return (`CMKT`) is constructed as the difference between the cryptocurrency market return and the risk-free rate. To proxy the risk-free rate, I used the (daily) 1-month Treasury bill rate from the FRED.

3.2 Variable description

3.2.1 Volume shock

Following Bianchi et al. (2022), the volume shock is defined as the log-deviation of trading volume from its rolling average (over 30 or 60 days) for cryptocurrency i at time t. For $m \in \{30, 60\}$ periods, the volume shock is estimated as:

$$v_{i,t} = \log(\text{Volume}_{i,t}) - \log\left(\frac{1}{m}\sum_{s=1}^{m} \text{Volume}_{i,t-s}\right)$$

3.3 Risk

3.3.1 Realized volatility (rvol)

Using the volatility estimator of Yang and Zhang (2000), I compute the daily realized volatility based on OHCL prices over a rolling 30-day window. For $n > 1$ number of periods, the volatility estimate at time t is:

$$\sigma_t = \sqrt{\sigma_O^2 + k\sigma_C^2 + (1-k)\sigma_{RS}^2}$$

where σ_{RS}^2 is the variance estimator of Rogers et al. (1994), and σ_O^2, σ_C^2, k are defined as follows:

$$\sigma_O^2 = \frac{1}{n-1}\sum_{i=1}^{n}(o_i - \bar{o})^2,$$

$$\sigma_C^2 = \frac{1}{n-1}\sum_{i=1}^{n}(c_i - \bar{c})^2,$$

$$k = \frac{\alpha - 1}{\alpha + \frac{n+1}{n-1}}$$

with $o = ln\, O_t - ln\, C_{t-1}$, and $c = ln\, C_t - ln\, O_t$. Here, C_{t-1} denotes the last days' closing price and O_t the current day's opening price. I set the constant $\alpha = 1.34$ as suggested by Yang and Zhang (2000) to be the best value in practice.

Moskowitz et al. (2012)

References

Bianchi, D., & Babiak, M. (2021). Mispricing and Risk Compensation in Cryptocurrency Returns (SSRN Scholarly Paper 3935934). Social Science Research Network. https://doi.org/10.2139/ssrn.3935934

Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common Risk Factors in Cryptocurrency. The Journal of Finance, 77(2), 1133–1177. https://doi.org/10.1111/jofi.13119

Mercik, A., Będowska-Sójka, B., Karim, S., & Zaremba, A. (2025). Cross-sectional interactions in cryptocurrency returns. International Review of Financial Analysis, 97, 103809. https://doi.org/10.1016/j.irfa.2024.103809

Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. Journal of Financial Economics, 104(2), 228–250. https://doi.org/10.1016/j.jfineco.2011.11.003