







# 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

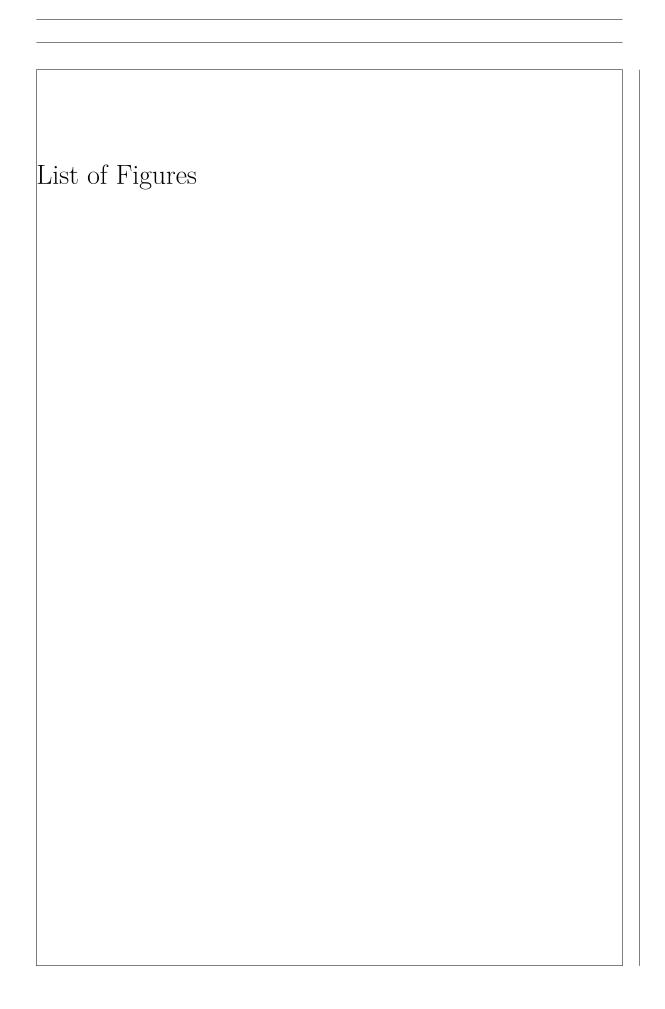
University: Vienna University of Economics and Business

Betreut von / Supervisor: Univ.Prof. David Preinerstorfer, Ph.D.

Wien, August 2025 / Vienna, August 2025

# Table of contents

Pr	reface	5					
1	Introduction	6					
2	2 Literature review						
3	Methodology						
4	Data	9					
	4.1 Data extraction and sample construction	9					
	4.2 Variable description	10					
	4.3 Risk	10					
$\mathrm{R}\epsilon$	eferences	12					



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

Mirar Liu et al (y el otro paper similar) donde habla de las criptomonedas, e inspirarnos de ahi. Quiza la otra tesis en esto tambien.

Empezar con una introduccion de criptomonedas, del mercado, de la gran volatilidad, grandes retornos. Mencionar coin Market cap, la capitalizacion del mercado total de criptomonedas.

Mencionar articulos o reportes donde mencionen la importancia de este mercado, cuantas personas en promedio tienen criptos en su portafolio. Mencionar sucesos recientes importantes, como la introduccion de criptomonedas en algunos exchanges, de futuros en CME, de indices en XXX, del boom en la crisis de COVID (ver Mercik, donde menciona sucesos importantes).

mencionar a Ross, que comprobo la estructura lineal de los factores: – Esto quiza en introduccion

\*\* Ejemplo de frase See (knuth84?) for additional discussion of literate programming.\*\*

1 +	1				
[1]	2				

### 2 Literature review

Hacerla similar a la descripciond e la hipoteca inversa. Mencionar las dos corrientes de litertura: por una parte, los modelos de factores construidos como managed-strategies. Mencionar los modelos de factores mas conocidos, y alguna de la literatura importante en este tema: fama french modelo de tres factores Fama & French (1993), FF modelo de cinco factores Fama & French (2015) Citar la coleccion de Chen and ZImmerman A. Y. Chen & Zimmermann (2021) con la gran cantidad de factores en su dataset.

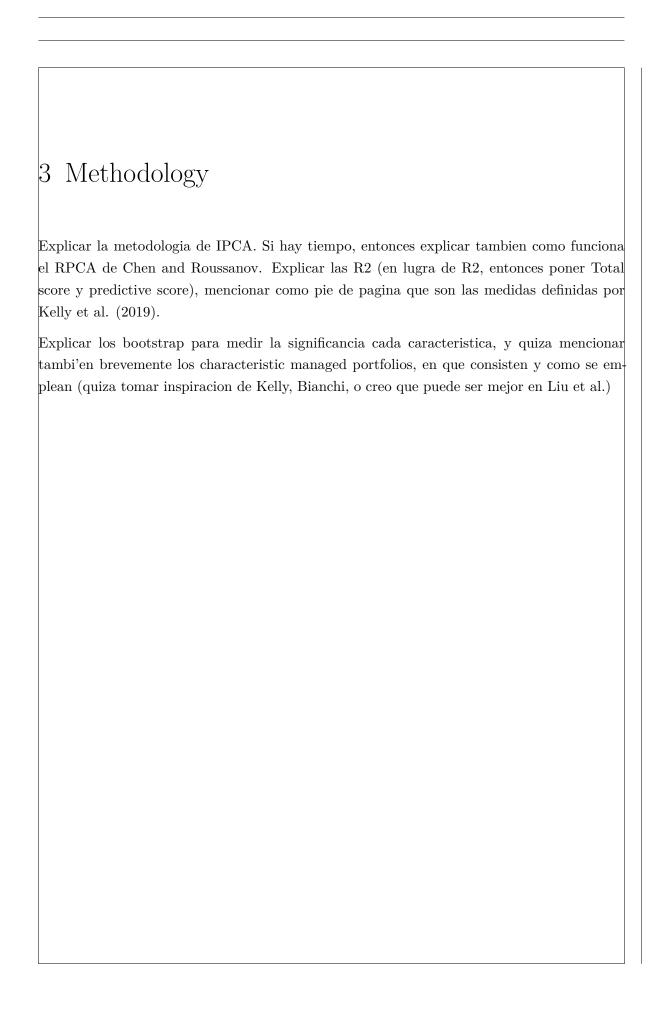
Por otra, una corriente basada en la estadistica que construye los factores mediante modelos puramente estadisticos, y que asume que los factores son latentes. Mencionar los dos autores pioneros de PCA: Chamberlain & Rothschild (1983) and Connor & Korajczyk (1986).

Mencionar modelos recientes con factores dinamicos. Describir brevemente estos modelos de PCA, y despues propuestas de modelacion de factores latentes dinamicos. Entre estos, hablar de Kelly et al. (2019), que propuso el modelo dinamico de factores, y mas recientemente. RPCA Q. Chen et al. (2022) Z. Chen et al. (2024), inspirado en regressiones de fama macbeth, que hace una combinacion de este modelo mas una implementacion de PCA.

Bianchi & Babiak (2021) aplicaron el modelo de Kelly en el mercado de criptos, mencionar otros papers que utilizaron el IPCA (hay otros de Kelly que lo uso para bonos y opciones investigar)

Z. Chen et al. (2024) aplico el RPCA para el cross-section de diferentes asset-classes.

Mencionar literatura extensa tratando de entender los factores de riesgo de las criptomonedas, por ejemplo, Mercik et al. (2025), y otros papers que tengo en mis archivos y en notas.



## 4 Data

### 4.1 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.

- 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.

### 4.2 Variable description

#### 4.2.1 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:

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

#### 4.3 Risk

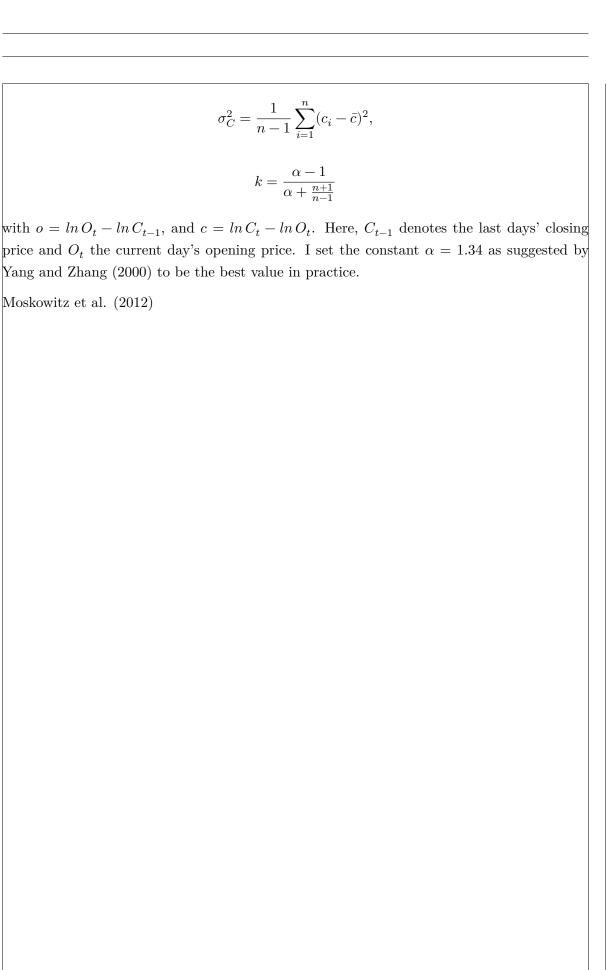
#### 4.3.1 Realized volatility (rvol)

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  $\sigma_{RS}^2$  is the variance estimator of Rogers et al. (1994), and  $\sigma_O^2$ ,  $\sigma_C^2$ , k are defined as follows:

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



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
- Chamberlain, G., & Rothschild, M. (1983). Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. Econometrica, 51(5), 1281–1304. https://doi.org/102307/1912275
- Chen, A. Y., & Zimmermann, T. (2021). Open Source Cross-Sectional Asset Pricing (SSRN Scholarly Paper 3604626). Social Science Research Network. https://doi.org/10.2139/ssrn.3604626
- Chen, Q., Roussanov, N. L., & Wang, X. (2022). Semiparametric Conditional Factor Models in Asset Pricing (SSRN Scholarly Paper 3984633). Social Science Research Network https://doi.org/10.2139/ssrn.3984633
- Chen, Z., Roussanov, N. L., Wang, X., & Zou, D. (2024). Common Risk Factors in the Returns on Stocks, Bonds (and Options), Redux (SSRN Scholarly Paper 4703281). Social Science Research Network. https://doi.org/10.2139/ssrn.4703281
- Connor, G., & Korajczyk, R. A. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. Journal of Financial Economics, 15(3), 373–394 https://doi.org/10.1016/0304-405X(86)90027-9
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010
- Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics, 134(3), 501–524. https://doi.org/10 1016/j.jfineco.2019.05.001
- 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