POLITECNICO
MILANO 1863

Bitcoin as a Digital Asset:

Correlation and Optimal Portfolio Allocation

Author:
Samuele Vianello

Supervisors:
Daniele Marazzina
Ferdinando M. Ametrano

School of Industrial and Information Engineering
Master of Science in Mathematical Engineering

16th April 2019

Objectives

The main goals of this work are the following:

1. Study the **correlation** of Bitcoin's returns with those of other standard assets by computing the *sample (Pearson's) correlation*, performing tests on its *significance* and computing the *rolling correlations*.
2. **Calibrate** three different continuous-time **models for the multivariate dynamics of the asset prices** (Merton, Heston and Bates) in order to obtain the levels of the correlation in a more sophisticated frameworks that includes *jumps* and *stochastic volatility*.
3. Study the **optimal allocation** for a portfolio that contains Bitcoin and the **diversification properties** of the inclusion of the digital asset.

Introduction

Bitcoin was introduced in (Nakamoto, 2009) as a *peer-to-peer* electronic cash system. It started off as a system that was only used by a niche of people in online cryptography forums but year by year it quickly gained notoriety.

Through its distributed blockchain, Bitcoin is able to solve the problem of double spending without the need for a trusted third party. Transaction are quickly validated by the network, making the transfer of wealth as easy as online data sharing.

Bitcoin's monetary policy based on deterministic supply achieves *scarcity* in the digital realm and mimics the progressive scarcity of gold.

Outline

Dataset I

The dataset is composed of 2163 observations of the prices of *16 assets valued daily* (excluding holidays and weekends) from 19/07/2010 till 2/11/2018 (all data provided by Bloomberg).
The assets are grouped into five classes:

1. **Bitcoin** (btc): Value of a single bitcoin, quoted in dollars.
2. Stock indexes:
    - **S&P500** (sp500): American stock market index based on 500 large company with stock listed either on the NYSE or NASDAQ.
    - **EUROSTOXX 50** (eurostoxx): equity index of eurozone stocks, covering 50 stocks from 11 eurozone countries.
    - **MSCI BRIC** (bric): market cap weighted index designed to measure the equity market performance across the emerging country indexes of Brazil, Russia, India and China.
    - **NASDAQ**(nasdaq): market cap weighted index including all NASDAQ tiers: Global Select, Global Market and Capital Market.

Dataset II

3. Bond indexes:
   - **BBG Pan European** (bond_europe): Bloomberg Barclays Pan-European Aggregate Index that tracks fixed-rate, investment-grade securities issued in different European currencies.
   - **BBG Pan US** (bond_us): BBG US Aggregate Bond Index, a benchmark that measures investment grade, US dollar-denominated, fixed-rate taxable bond market.
   - **BBG Pan EurAgg** (bond_eur): similar to the Pan European but it only considers securities issued in Euros.
4. Currencies:
   - **EUR/USD** (eur): spot value of one Euro in US dollars.
   - **GBP/USD** (gbp): spot value of one British Pound in US dollars.
   - **CHF/USD** (chf): spot value of one Swiss Franc in US dollars.
   - **JPY/USD** (jpy): spot value of one Japanese Yen in US dollars.

Dataset III

5. Commodities:
   - **Gold** (gold): price of gold measured in USD/Oz.
   - **WTI** (wti): price of crude oil used as benchmark in oil pricing and as the underlying commodity in the NYMEX oil future contracts.
   - **Grain** (grain): S&P GPSCI index that measures the performance of the grain commodity market.
   - **Metals** (metal): S&P GSCI Industrial Metals index that measures the movements of industrial metal prices including aluminium, copper, zinc, nickel and lead.

Outline

Empirical Correlation

The empirical correlation is computed using Pearson's sample correlation formula on the daily log-returns obtained from the price dataset.

	btc	bric	sp500	eurostoxx	nasdaq	bond_europe	bond_us	bond_eur	eur	gbp	chf	jpy	gold	wti	grain	metal
btc	100.0	1.4	4.4	4.1	3.6	1.4	-1.8	1.9	2.3	0.7	2.5	-1.1	-0.2	0.8	3.5	2.7
bric	1.4	100.0	48.4	57.1	47.4	19.6	-15.1	19.8	20.1	24.2	8.1	-16.2	12.9	30.3	15.2	43.2
sp500	4.4	48.4	100.0	62.0	94.9	13.3	-34.1	15.0	18.4	21.1	0.1	-22.2	-0.6	35.1	15.2	34.9
eurostoxx	4.1	57.1	62.0	100.0	56.2	42.0	-27.8	44.7	48.6	41.9	21.2	-16.5	9.3	32.7	15.2	46.5
nasdaq	3.6	47.4	94.9	56.2	100.0	10.9	-31.4	12.3	15.1	18.6	-1.8	-21.4	-1.0	29.3	14.1	32.4
bond_europe	1.4	19.6	13.3	42.0	10.9	100.0	19.4	98.4	91.9	61.8	60.1	39.9	42.8	14.5	11.6	26.8
bond_us	-1.8	-15.1	-34.1	-27.8	-31.4	19.4	100.0	14.5	-0.4	-5.0	14.0	38.6	21.3	-21.1	-6.8	-17.1
bond_eur	1.9	19.8	15.0	44.7	12.3	98.4	14.5	100.0	94.5	53.6	58.4	37.0	40.7	14.3	11.3	28.0
eur	2.3	20.1	18.4	48.6	15.1	91.9	-0.4	94.5	100.0	57.1	59.4	31.2	36.7	18.2	13.5	31.0
gbp	0.7	24.2	21.1	41.9	18.6	61.8	-5.0	53.6	57.1	100.0	35.5	14.2	24.7	21.9	11.9	26.1
chf	2.5	8.1	0.1	21.2	-1.8	60.1	14.0	58.4	59.4	35.5	100.0	36.7	37.1	6.7	7.5	20.8
jpy	-1.1	-16.2	-22.2	-16.5	-21.4	39.9	38.6	37.0	31.2	14.2	36.7	100.0	39.5	-6.5	2.1	-3.1
gold	-0.2	12.9	-0.6	9.3	-1.0	42.8	21.3	40.7	36.7	24.7	37.1	39.5	100.0	14.7	13.7	32.0
wti	0.8	30.3	35.1	32.7	29.3	14.5	-21.1	14.3	18.2	21.9	6.7	-6.5	14.7	100.0	17.8	36.0
grain	3.5	15.2	15.2	15.2	14.1	11.6	-6.8	11.3	13.5	11.9	7.5	2.1	13.7	17.8	100.0	20.7
metal	2.7	43.2	34.9	46.5	32.4	26.8	-17.1	28.0	31.0	26.1	20.8	-3.1	32.0	36.0	20.7	100.0

The results clearly show that:

- Bitcoin has **low correlation** with every asset.
- Assets in the same class usually have a **high correlation** among them.

Correlation Significance

We perform two tests on the significance of the correlation between Bitcoin and each of the other assets, both of which investigate the following hypotheses:

$$\mathbf{H_0}: \quad \rho = 0 \quad \textit{vs.} \quad \mathbf{H_1}: \quad \rho \neq 0.$$

Pearson's Test: by computing *Pearson's t-statistic*, which under the null hypothesis is distributed as a Student's t, we can obtain the *p-value* of the test and compare it to the confidence level of 95%.

Permutation test: by permuting the sample of a pair of asset and computing Pearson's correlation on the permuted data a large number of times, we can reconstruct an empirical distribution for the possible correlations. Once we obtained this distribution, we can compute the *p-value* of the test and compare it to the chosen confidence level.

Correlation Significance

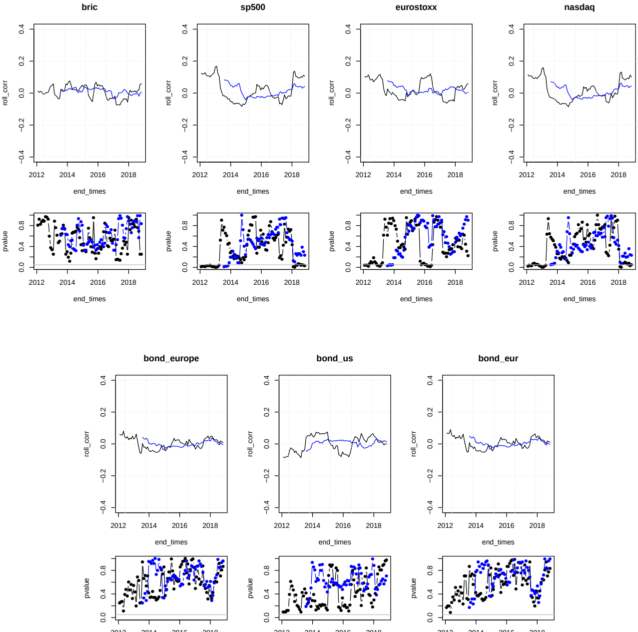
The tables report values of the correlation of Bitcoin with the other assets and the *p-values* of the two tests. All p-values are higher then the threshold of 5% (100% minus the confidence level) except for S&P500.

	bric	sp500	eurostoxx	nasdaq	bond_europe	bond_us	bond_eur
Correlation	1,41%	4,37%	4,12%	3,59%	1,41%	-1,84%	1,92%
Pearson	52,70%	4,10%	5,40%	9,35%	50,35%	39,90%	37,95%
Permutation	51,21%	4,24%	5,55%	9,48%	51,29%	39,28%	37,20%

	eur	gbp	chf	jpy	gold	wti	grain	metal
Correlation	2,29%	0,73%	2,46%	-1,09%	-0,24%	0,77%	3,50%	2,69%
Pearson	29,00%	72,75%	24,75%	61,45%	91,00%	71,55%	10,10%	20,90%
Permutation	28,64%	73,33%	25,34%	61,32%	91,08%	72,00%	10,34%	21,13%

This means that all correlations are **not significantly different from zero**. The only value for which both tests assert that the correlation is *statistically significantly different from zero* is for S&P500.

Rolling Correlation



Outline

Merton Model

In his article (Merton, 1976), Merton presented a **jump diffusion model** with the aim to account for sudden rises and falls of the prices due to the arrival of new information. It is based on the dynamics of a *geometric Brownian motion* (GBM), which are the same dynamics used in Black&Scholes:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

to which Merton added a jump component:

$$\frac{dS_t}{S_t} = (\mu - \lambda\mu_J)dt + \sigma dW_t + Y_t dN.$$

We have that μ is the *drift*, σ is the *volatility*, $N(t)$ is a Poisson process with parameter λ to model the arrival of jumps, Y_t is a stochastic process to model the intensity of the jumps. In the case of Merton, Y_t is distributed as log-normal, namely $y_t = \log(Y_t) \sim \mathcal{N}(\mu_J, \sigma_J^2)$.

Heston Model

Heston model (Heston, 1993) is another improvement from the simple GBM: relaxing the assumption that the volatility is constant, Heston added **a stochastic process to model the variance**. The dynamics are described by the following equations:

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{V_t} dW_t^S,$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_t^V.$$

The variance V_t follows a CIR process (Cox et al., 1985) which is a *mean-reverting* process with mean reversion *level* and *rate* represented respectively by θ and κ. σ_V models the volatility of the variance process. The two Brownian motions W_t^S and W_t^V are correlated through the parameter ρ: $dW_t^S dW_t^V = \rho dt$. The *drift* is modelled by μ.

It can be shown that the variance process is always non negative when the *Feller* condition is satisfied:

$$2\kappa\theta \geq \sigma_V$$

Bates Model

Bates model (Bates, 1996) is the combination of Merton and Heston model. It includes both a **jump component** and **stochastic volatility**:

$$\frac{dS_t}{S_t} = (\mu - \lambda \mu_J)dt + \sqrt{V_t}dW_t^S + Y_t dN,$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma_V \sqrt{V_t}dW_t^V.$$

As before, μ is the *drift*, V_t is the variance following a CIR process with parameters κ, θ and σ_V, ρ is the correlation between the two Brownian motions. The jump component is modelled by Y_t such that $y_t = \log(Y_t) \sim \mathcal{N}(\mu_J, \sigma_J^2)$. and $N(t)$ is the Poisson process with parameter λ.

The variance process is always non negative when the *Feller* condition is satisfied:

$$2\kappa\theta \geq \sigma_V$$

Single Asset Calibration

In order to calibrate each set of parameter ψ, we adopt the **maximum likelihood estimation** (MLE) procedure. MLE consists of maximizing the likelihood function $\mathcal{L}(\psi|X)$ computed for the sample data X over the domain Ψ of all admitted values of the parameters:

$$\widehat{\psi} = \underset{\psi \in \Psi}{\arg\max}\ \mathcal{L}(\psi|X).$$

In our case, X is composed of all observed **daily log-returns** $x_i = \log(S_i/S_{i-1})$ and they are all **i.i.d.** which allows us to compute the likelihood function as:

$$\mathcal{L}(\psi|X) = \prod_{i=1}^{N} f_{\mathcal{D}}(x_i;\psi),$$

where $f_{\mathcal{D}}(x;\psi)$ is the density function (pdf) of the log-returns. In practice, it is a common approach to consider the *log*-likelihood function:

$$\ell(\psi|X) = \log \mathcal{L}(\psi|X) = \sum_{i=1}^{N} \log f_{\mathcal{D}}(x_i;\psi).$$

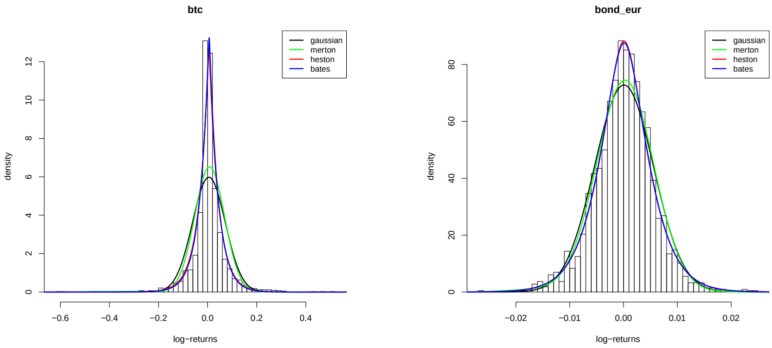
Issues with Calibrating

When performing the calibration, we have to face the following issues:

- **choosing the optimizer**: the trade-off is between fast computation and robust solution. Taking into account the complexity of the shape of the log-likelihood, we implemented the combination of a **global** (*Differential evolution*) and **local** (*quasi-Newton method*) **optimizers**
- **unavailability of the pdf**: in the cases of Heston and Bates models, there are no explicit formulas for the pdf. We have to obtain it from the **characteristic function** via *Fourier inversion*. Luckily, this can be numerically implemented using the **FFT** algorithm.

Graphical Results

The two examples below show that in our implementation of the calibration **Heston** and **Bates** are able to **better fit** the empirical distribution of the log-returns, especially its *leptokurtosis*.



Multi Asset Framework

In order to obtain the values for the correlation matrix for the different models, we followed the *parsimonious approach* presented in (Dimitroff et al., 2011) that only directly models the *asset-asset* correlations and not the *asset-variance* or *variance-variance*.
There are three steps to obtaining the model matrix correlation:

1. **Calibrate** each **single asset model**
2. **Compute** the model **correlation** for **every asset pair**. This is done by simulating the two paths N_{sim} times with ρ as the *model correlation* and computing the *simulated correlation* ρ_{sim} in each scenario. The aim is to find a ρ such that $\mathbb{E}[\rho_{sim}(\rho)] = \rho_{obs}$, where ρ_{obs} is the *sample correlation* observed from our dataset.
3. **Perform Jäckel regularization** to the matrix obtained in step 2 to get a valid correlation matrix as final result.

Graphical Comparison of Correlation Matrices

Outline

Optimal Allocation Framework

The general framework introduced by Markowitz in (Markowitz, 1952) consists of *optimizing* the allocation of wealth based on the **assets returns** and the overall **portfolio risk** that must also take into account the correlation in the asset prices.

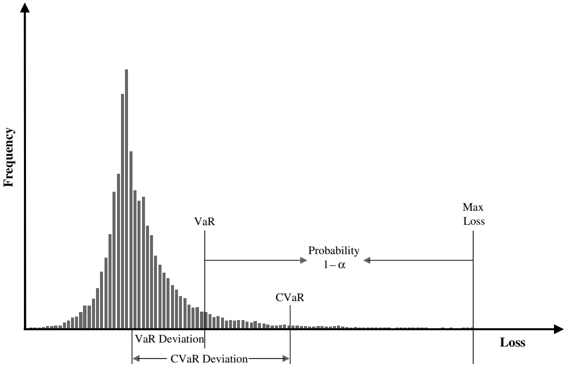
$$
\begin{aligned}
\min_{\mathbf{w} \in \mathbb{R}^N} \quad & R_{ptf}(\mathbf{w}) \\
\text{subject to} \quad & \mathbf{e}^T \mathbf{w} = 1, \\
& \mathbf{r}^T \mathbf{w} = r_{target}, \\
& w_i \geq 0, \text{for } i = 1 \ldots N,
\end{aligned}
$$

Where $\mathbf{w}$ is the vector of weights, $\mathbf{e} = (1, ..., 1)^T$ is a vectot of ones, $\mathbf{r}$ is the vector of expected return of the assets, $R_{ptf}(\mathbf{w})$ is the portfolio risk measure.

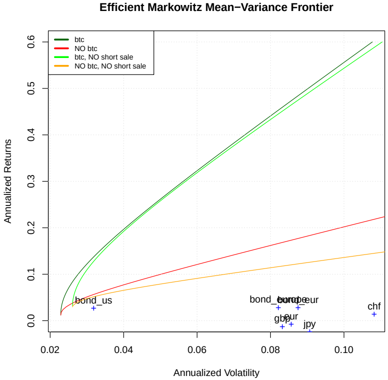
Portfolio Risk Measures

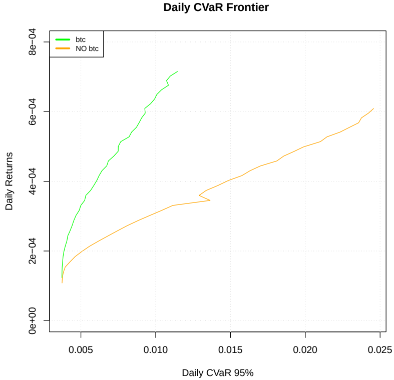
We will consider two different risk measures for the portfolio risk:

1. Portfolio Variance $\sigma^2_{ptf}(\mathbf{w}) = \mathbf{w}^T \Sigma \mathbf{w}$. Σ is the (sample) covariance matrix of the assets.
2. Daily Portfolio $CVaR_{95\%}(\mathbf{w})$: we can obtain the distribution of the returns of the portfolio by multiplying the matrix of daily observation of our dataset to the portfolio weights. Then we compute the $CVaR$ for the portfolio loss distribution.

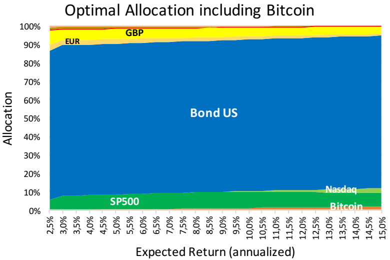


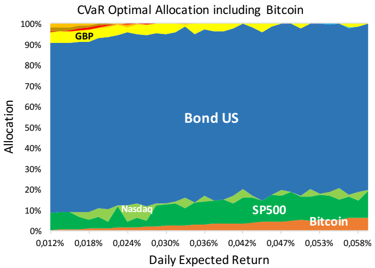
Efficient Frontiers

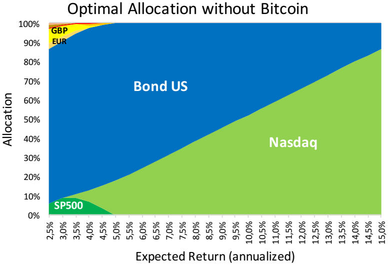


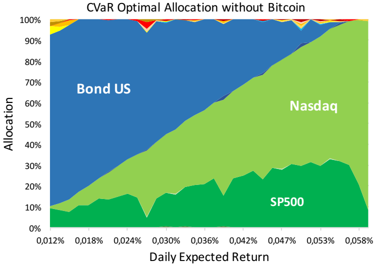


Optimal Allocation









Outline

Conclusions

References I

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 9(1):69–107.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.

Dimitroff, G., Lorenz, S., and Szimayer, A. (2011). A parsimonious multi-asset heston model: Calibration and derivative pricing. *International Journal of Theoretical and Applied Finance (IJTAF)*, 14(08):1299–1333.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–43.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Merton, R. (1976). Option prices when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

Nakamoto, S. (2009). Bitcoin: A peer-to-peer electronic cash system.