Common dynamic factors for cryptocurrencies and multiple pair-trading statistical arbitrages

Marco Patacca

from joint work with
Gianna Figà-Talamanca and Sergio Focardi

Department of Economics,
University of Verona,
marco.patacca@univr.it

3^{rd} Crypto Asset Lab Conference
Milano, November 4, 2021

Outline of the talk

- Introduction and motivation
- Preliminary analysis
- Dynamic factor model
- Market neutral strategy
- Concluding remarks

Introduction

- BitCoin is the first decentralised digital currency, which provides a solution to the problem of trust in a currency system.
- Opposite to traditional transactions, which are based on trust in financial intermediaries, this system relies on network, on fixed rules and on cryptography.
- The entire system is based on an open source software created in 2009 by a computer scientist known under the pseudonym Satoshi Nakamoto, and it generates a peer to peer network which includes a high number of computers connected to each other through the Internet.

Introduction

In the years following the birth of Bitcoin many other cryptocurrencies have been created, generally called Altcoins.

Nowadays there are more than 3000 Altcoins in the cryptocurrencies world but only a few of them have a relevant market capitalization.

In our paper we focus on four cryptocurrencies: Bitcoin, Ethereum, Litecoin, Monero.

We made this choice to combine a double need:

- currencies with significant market cap;
- availability of at least three years time series data.

Goals

As several cryptocurrencies are now available and can be traded on exchanges, we want to explore the possibility of forming market neutral strategies, where gains and losses depend only on the relative behavior of assets.

Preliminary analysis

In this paper we consider the daily prices, observed from the website `https://coinmarketcap.com/`, of Bitcoin, Ethereum, Litecoin and Monero, in the period from January 1st, 2016 to November 30, 2019.

Table: Summary Statistics of Cryptocurrencies prices

	BTC	ETH	LTC	XMR
Min.	364.33	0.94	3.00	0.43
Q_1	818.41	12.79	4.05	11.27
Median	4035.57	175.05	48.82	54.56
Mean	4830.50	226.47	57.59	79.10
Q_3	7833.04	298.33	78.77	105.11
Max.	19497.40	1396.42	358.34	469.20
St. Dev.	3974.16	248.35	60.12	87.70

Preliminary analysis

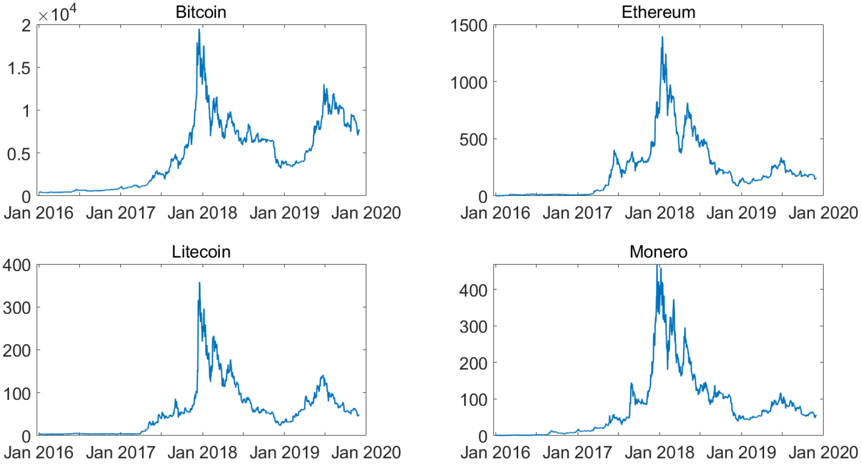


Figure: Price behaviour for Bitcoin (top-left), Ethereum (top-right), Litecoin (bottom-left), Monero (bottom-right) from January 1st, 2016 to November 30, 2019.

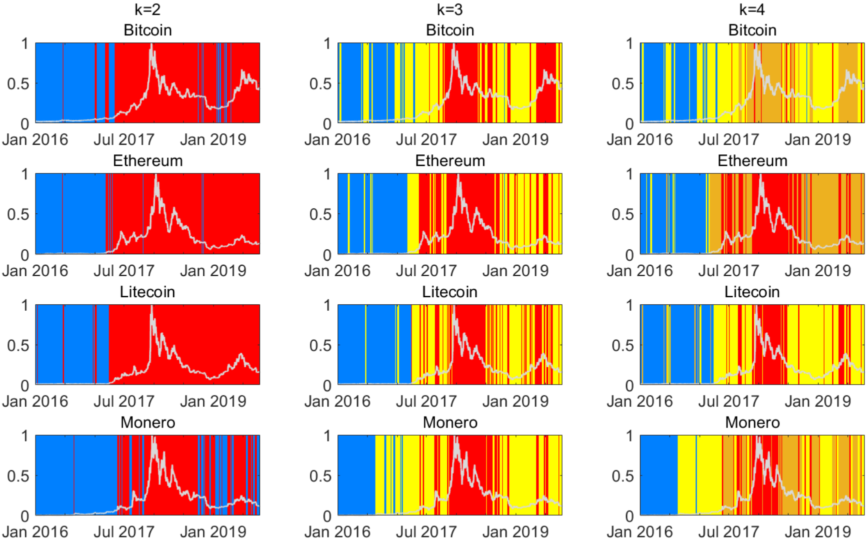
Preliminary analysis

In a companion paper (Figá-Talamanca, Focardi, and Patacca, 2021) we analyze the co-movement among several cryptocurrencies.

Specifically, we perform the analysis using regime switching markov models and the results show

- the existence of common regimes in the analyzed cryptocurrencies (BitCoin, Ethereum, Litecoin, Monero);
- that at least two regimes are statistically significant for all model specifications:
- that the best performance are obtained with three independent regimes or two common regimes.

Preliminary analysis



Preliminary analysis

Table: ADF and KPSS tests of cryptocurrencies daily prices (Panel a) and price differences (Panel b) from January 1, 2016 to November 30, 2019.

	BTC	ETH	LTC	XMR
Panel a: prices				
ADF-Test *p*-value	0.4312	0.2319	0.1546	0.1382
KPSS-Test *p*-value	$<$0.010	$<$0.010	$<$0.010	$<$0.010
Panel b: price differences				
ADF-Test *p*-value	$<$0.001	$<$0.001	$<$0.001	$<$0.001
KPSS-Test *p*-value	$>$0.100	$>$0.100	$>$0.100	$>$0.100

0.001 and 0.100 are respectively the minimum and the maximum *p*-values provided by Matlab® *adftest.m* and *kpsstest.m* functions.

Preliminary analysis

Table: Johansen cointegration test between Bitcoin, Ethereum, Litecoin and Monero prices: data are from January 1, 2016 to November 30, 2019.

r	statValue	cValue	pValue
0	105.0086	40.1751	0.0010
1	57.8639	24.2747	0.0010
2	15.7722	12.3206	0.0130
3	0.5850	4.1302	0.6356

The four analyzed cryptocurrencies are cointegrated with three cointegrating vectors.

Dynamic factor model

Dynamic factor models are models that allow to specify a dynamics for factors and for the processes themselves.

The general specification of a dynamic factor model is

$$\begin{aligned} p_t &= \beta f_t + \epsilon_t \\ \Phi(L) f_t &= \Theta(L) \eta_t \\ \Phi(L) &= 1 - \Phi_1 L - \cdots - \Phi_p L^p \\ \Theta(L) &= 1 - \Theta_1 L - \cdots - \Theta_q L^q \end{aligned}$$

where f_t represent the factors, L is the lag operator, ϵ_t and η_t are the errors processes.

Dynamic factor model

Assume we are given with I different cryptocurrencies and that we model their mutual relationship through a dynamic factor model.
For $i = 1, 2, \dots, I$ and $t = 1, 2, \dots, T$:

$$\begin{aligned} p_{i,t} &= \alpha_i + \sum_{k=1}^{K} \beta_{ik} f_{k,t} + \epsilon_{i,t}, \\ f_{k,t} &= \lambda_k f_{k,t-1} + \eta_{k,t}, \; k = 1, 2, \dots, K \\ \epsilon_{i,t} &= \phi_{i,1}\epsilon_{i,t-1} + \phi_{i,2}\epsilon_{i,t-2} + \dots + \phi_{i,p_i}\epsilon_{i,t-p_i} + u_{i,t}. \end{aligned} \tag{1}$$

where $p_{i,t}$ is the price at time t of the cryptocurrency i and $f_{k,t}$, for $k = 1, 2, \dots, K$ are the common factors.

Dynamic factor model

The presence of three cointegrating vectors implies that our four cryptocurrencies share one common integrated I(1) factor.

In order to detect the correct number of common factors:

- we estimate the model in (1) with just one factor f_1;
- we compute the covariance matrix of residuals $\epsilon_{i,t}$ and corresponding eigenvalues;
- Since the eigenvalues are $(1184763.89, 6946.14, 185.03, 14.32)$, we add a second relevant factor.

Dynamic factor model

Based on the previous analysis we assume $k = 2$ and the considered model specification is finally given by:

$$\begin{aligned} p_{i,t} &= \alpha_i + \beta_{i1} f_{1,t} + \beta_{i2} f_{2,t} + \epsilon_{i,t}, \\ f_{1,t} &= \lambda_1 f_{1,t-1} + \eta_{1,t}, \\ f_{2,t} &= \lambda_2 f_{2,t-1} + \eta_{2,t}, \\ \epsilon_{i,t} &= \phi_{i,1}\epsilon_{i,t-1} + \phi_{i,2}\epsilon_{i,t-2} + \ldots + \phi_{i,p_i}\epsilon_{i,t-p_i} + u_{i,t}. \end{aligned} \quad (2)$$

By cointegration analysis we expect $\lambda_1 = 1$ and $f_{1,t} = f_{1,t-1} + \eta_{1,t}$.

We apply the technique of dynamic factor analysis using the Matlab® software for Bayesian VAR models provided by Koop and Korobilis (2009).

Dynamic factor model

Table: Parameter estimates from January, 2016 to December, 2018

	BTC ($i = 1$)	ETH ($i = 2$)	LTC ($i = 3$)	XMR ($i = 4$)
α_i	4045.7660	238.8579	53.1539	82.4727
β_{i1}	0.9911	0.0700	0.0170	0.0260
β_{i2}	0.1406	5.2255	0.0351	-0.0542

The first factor essentially emulates the dynamics of Bitcoin and the second factor is strongly related to Ethereum.

Market neutral strategy

The integrated factor is an unforecastable random walk while the stationary factor is a forecastable I(0) process.

We want create a long-short portfolio that does not depend on the unforecastable I(1) factor but only on stationary forecastable factors.

Since we are interested in building an investment strategy to take advantage of the dynamic model defined in (2), we can, without loss of generality, scale the price equations with the corresponding β_{i1} coefficients. Hence, we get:

$$p^*_{i,t} = \frac{p_{i,t}}{\beta_{i1}} = \frac{\alpha_i}{\beta_{i1}} + f_{1,t} + \frac{\beta_{i2}}{\beta_{i1}} f_{2,t} + \epsilon_{i,t}$$

Market neutral strategy

The scaled prices are equal to the integrated factor $f_{1,t}$ plus a stationary process $f_{2,t}$ and a constant term.

The gain from a simple long-short strategy with any pair of assets is only driven by the stationary factor f_2.

The difference (spread) between the prices of cryptocurrencies $i, k \in \{1, 2, \ldots, I\}$ is given by the following equation:

$$d_{ik,t} = \frac{p_{i,t}}{\beta_{i1}} - \frac{p_{k,t}}{\beta_{k1}} = \left(\frac{\alpha_i}{\beta_{i1}} - \frac{\alpha_k}{\beta_{k1}}\right) + \left(\frac{\beta_{i2}}{\beta_{i1}} - \frac{\beta_{k2}}{\beta_{k1}}\right) f_{2,t} + \epsilon_{ik,t}.$$

Market neutral strategy

Assume that we are at time τ and that the price of our basket of cryptocurrencies is described by the model in (2).

The one-day ahead forecasted prices is:

$$\hat{p}_{i,\tau+1} = \mathbb{E}_\tau\left(p_{i,\tau+1}\right) = \hat{\alpha}_i + \hat{\beta}_{i1}\mathbb{E}_\tau\left(f_{1,\tau+1}\right) + \hat{\beta}_{i2}\mathbb{E}_\tau\left(f_{2,\tau+1}\right)$$

where

$$\mathbb{E}_\tau\left(f_{1,\tau+1}\right) = \hat{\lambda}_1 f_{1,\tau} \quad , \quad \mathbb{E}_\tau\left(f_{2,\tau+1}\right) = \hat{\lambda}_2 f_{2,\tau}.$$

The one-day-ahead forecasted difference in $\tau+1$ for any pair $i,k \in \{1,2,\ldots,I\}$ of cryptocurrencies is given by

$$\begin{aligned} \hat{d}_{ik,\tau+1} &= \left(\frac{\hat{\alpha}_i}{\hat{\beta}_{i1}} - \frac{\hat{\alpha}_k}{\hat{\beta}_{k1}}\right) + \left(\frac{\hat{\beta}_{i2}}{\hat{\beta}_{i1}} - \frac{\hat{\beta}_{k2}}{\hat{\beta}_{k1}}\right)\mathbb{E}_\tau\left(f_{2,\tau+1}\right) \\ &= \left(\frac{\hat{\alpha}_i}{\hat{\beta}_{i1}} - \frac{\hat{\alpha}_k}{\hat{\beta}_{k1}}\right) + \left(\frac{\hat{\beta}_{i2}}{\hat{\beta}_{i1}} - \frac{\hat{\beta}_{k2}}{\hat{\beta}_{k1}}\right)\hat{\lambda}_2 f_{2,\tau} \end{aligned}$$

Market neutral strategy

If the difference is strictly positive, a future revenue can be obtained by applying a long-short investment in the pair i, k.

We maximize the revenue in $\tau + 1$ by investing on several pairs $i, k \in \{1, 2, \ldots, I\}$ which display a non negative forecasted difference.

Denote with $\hat{p}^{(i)}_{\tau+1}, i = 1, 2 \ldots, I$ the ordinal statistics (in decreasing order) of the scaled forecasted prices for time $\tau + 1$, the multiple pair trading consists essentially of short positions on the first half of cryptocurrencies (with higher forecasted prices) and long positions in the second half of cryptocurrencies (with lower forecasted prices).

Market neutral strategy

Denote with v_τ the value at time τ of the above investment portfolio. The one-day-ahead expected value of the strategy, computed at time τ, is given by

$$g_{\tau+1} = \mathbb{E}_\tau\left[v_{\tau+1}\right] = \sum_{i=1}^{\lfloor I/2 \rfloor} \hat{p}^{(i)}_{\tau+1} - \sum_{i=\lceil I/2 \rceil+1}^{I} \hat{p}^{(i)}_{\tau+1}$$

where $\lfloor I/2 \rfloor, \lceil I/2 \rceil$ are, respectively, the floor and ceil rounding of $I/2$.

Market neutral strategy

We compare the forecasted gain in $\tau + 1$ with the corresponding gain in τ and we apply the following strategy:

- if $g_{\tau+1} > v_\tau + c\sigma^v_\tau$, go long with the multiple pair trading,
- if $g_{\tau+1} < v_\tau - c\sigma^v_\tau$, go short with the multiple pair trading,
- if $v_\tau - c\sigma^v_\tau \leq g_{\tau+1} \geq v_\tau + c\sigma^v_\tau$, hold the current positions (no trade),

where c is an arbitrary chosen constant and σ^v_τ is the standard deviation of the trading position value corresponding to the basket price time series observed up to time τ ($\pm c\sigma^v_\tau$ is used to avoid huge transaction costs).

Every long-short portfolio strategy will be liquidated at the following date.

Market neutral strategy

If the above trading strategy is repeated for m consecutive days $\{\tau+1, \tau+2, \dots \tau+m\}$ then the expected cumulative gain in $\tau+m$ is given by

$$G_{\tau+m} = \sum_{l=\tau}^{\tau+m-1} [g_{l+1} - v_l] \, \mathbb{1}_{trade(l)} \; ,$$

where the indicator function $\mathbb{1}_{trade(l)}$ is defined by

$$\mathbb{1}_{trade(l)} = \begin{cases} 1 \text{ if there is trade at time } l \\ 0 \text{ if there is no trade at time } l. \end{cases}$$

Empirical results

We apply the market neutral strategy to the daily prices of Bitcoin, Ethereum, Litecoin and Monero, from January 1, 2019 to November 30, 2019, i.e. a total of $M = 334$ days.

Each day the dynamic factor model is estimated on a moving window of daily observations available for the previous three years ($T = 1096$ observations).

Once parameters are estimated on each moving window, the prices forecast for the four cryptocurrencies in the basket are computed as

$$\hat{p}_{i,\tau+1} = \mathbb{E}_\tau\left(p_{i,\tau+1}\right) = \hat{\alpha}_i + \hat{\beta}_{i1}\mathbb{E}_\tau\left(f_{1,\tau+1}\right) + \hat{\beta}_{i2}\mathbb{E}\left(f_{2,\tau+1}\right).$$

We compute the one-day-ahead forecasted difference and we apply the proposed market neutral strategy.

Empirical results

Table: Summary Statistics of cumulative gain $G_{\tau+M}$ and net cumulative gain $G^*_{\tau+M}$

	$c = 0.00$	$c = 0.10$	$c = 0.20$	$c = 0.30$	$c = 0.50$	$c = 1.00$
Panel a: $G_{\tau+M}$						
Trade n.	252.00	232.00	222.00	200.00	172.00	72.00
Mean	**7625.65**	7620.59	7623.34	7615.08	7554.18	7467.75
St. Dev.	5474.23	5472.13	5474.02	5467.97	5477.52	5580.96
Panel b: $G^*_{\tau+M}$, taking into account transaction fees						
Trade n.	252.00	232.00	222.00	200.00	172.00	72.00
Mean	3031.17	3028.80	**3032.97**	3027.38	2973.89	2883.38
St. Dev.	2596.05	2594.68	2598.39	2593.16	2608.96	2720.23

The net gain is computed under the assumption that transaction fees are given by the 0.10% of the investment, which corresponds to the *maker fee* of Coinbase for the pricing tier from $100k to $1m of USD trading volume over the trailing 30 day period.

Empirical results

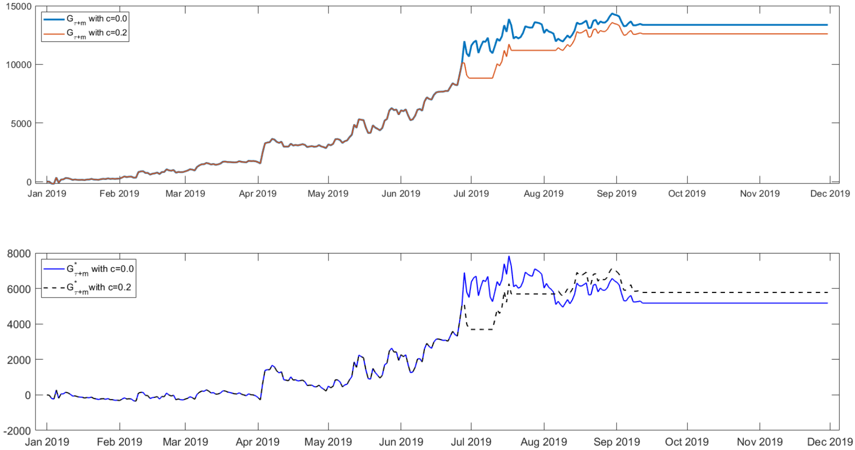


Figure: Cumulative gain $G_{\tau+m}$ (top) and net cumulative gain $G^*_{\tau+m}$ (bottom) with parameter $c = 0.00$ and $c = 0.20$ from $\tau + 1$=January 1, 2019 to $\tau + M$=November 30, 2019.

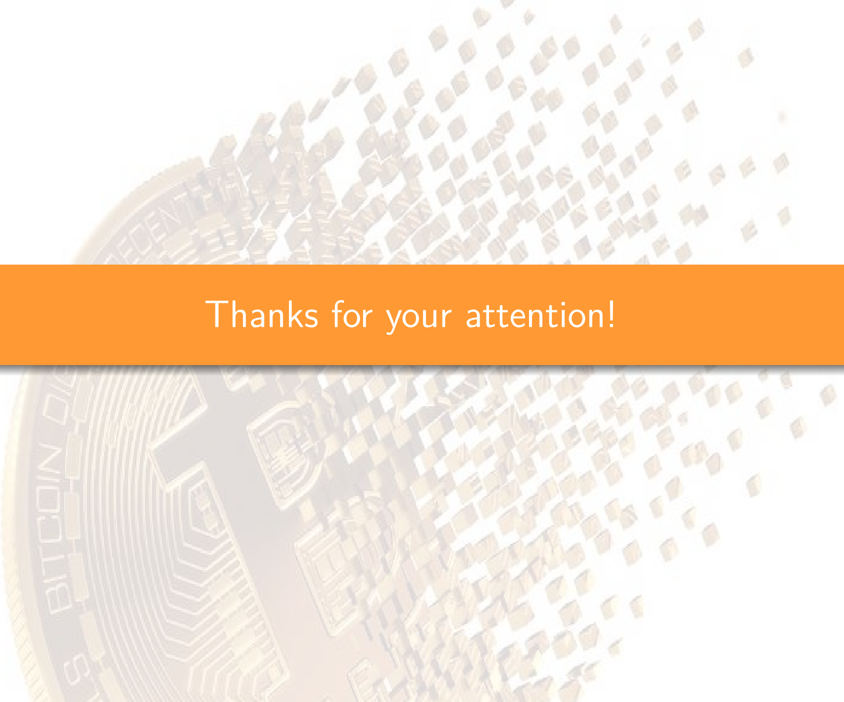
Concluding remarks

In this paper we argue that it is possible to form market neutral strategies in the cryptocurrencies world, where gains and losses depend only on the relative behavior of assets.

We use the technique of dynamic factor analysis to disentangle the integrated and stationary factors.

We propose a basic market neutral strategy to create a long-short portfolio that does not depend on the unforecastable I(1) factor.

We apply our strategy to empirical data and we obtain good results in terms of gain.



Thanks for your attention!