Crypto Asset Valuation

In this paper, Melanion Capital presents a quantitative model that is able to predict the evolution of the price of a crypto asset by decomposing the return into a cyclical and a trend components and applying signal filters.

MELANION CAPITAL

Published on:

7/17/2019

Authors:

Bing Zhou

Ghali Laraqui

Crypto assets have gathered significant attention from the media, financial analysts, governments, regulatory institutions, and investors over the past two years and a half. Crypto is defined broadly as digital units of account in which cryptographic techniques are used to regulate the generation and distribution of units on a blockchain.

In practice, crypto means multiple things to different people: an investment asset class like commodities, a store of value like gold, a legitimate medium of exchange, a covert method of exchange, an immutable record of rights and ownership, or even an incentive mechanism like rewards points. In this paper, we use crypto to refer to all crypto assets.

Cryptocurrencies, security tokens, and utility coins are different types of crypto assets. Some of these terms may be used interchangeably, particularly where concepts are applicable broadly to all types of assets, tokens, and coins. Crypto assets have potential. But for them to realize this potential, institutionalization is needed. Institutionalization is the at-scale participation in the crypto market of banks, broker dealers, exchanges, payment providers, fintech's, and other entities in the global financial services ecosystem. We believe this is a necessary next step for crypto to create trust and scale. Given that the high volatility and short track records of these assets, they are difficult to predict with accuracy.

In this paper, Melanion Capital presents a quantitative model that is able to predict the evolution of the price of a crypto asset by decomposing the return into a cyclical and a trend components and applying signal filters.

Contents

1. CRYPTO ASSET VALUATION FRAMEWORKS

1.1. Token Velocity Thesis

Token transaction velocity is one of the key levers that determines long-term token value. Main Argument: Drawing from The Monetary Equation of Exchange (MV=PQ), which economists call The Quantity Theory of Money, velocity is a significant driver of token price, and the lower the velocity, the greater token price is via an appreciation of M on the left side of the identity. The implication of this thesis is that tokens with low velocity, i.e. those that sit longer in wallets for whatever reason (speculation, store of value, etc.), will see higher prices than other coins, all else equal.

A key implication of this idea is that protocols and projects should give users a good reason to hold some coins beyond what they will spend in the system. Motivations could include holding the coin as speculative investment or store of value. Alternatively, the protocol/project could design features that force velocity reductions, such as staking functions (seen in FunFair) or balanced burn-and-mint mechanics (seen in Factom). Generally speaking, staking features such as those in PoS protocols should help support low velocity.

This method has widely been recognized criticisms, such as:

- Velocity cannot be precisely defined or measured, whereas the model assumes that it can be defined/estimated and employed to model value.
- The other factors in the equation, M, P, and Q can also not be easily measured or estimated. In fact, economists would say that you need models to estimate any one of these variables along with their correlations with one another.
- When velocity changes, the choice to record the effect in M, P, or Q is arbitrary and yields different implications for token price. Further, V's relationship and correlation with these factors is dynamic, and assuming a steady relationship with P, Q, or M is again arbitrary and problematic.
- M itself is very difficult to measure in crypto land, as there can be locked up or un-mined currency that may or may not be reflected in the model's M value.

1.2. Network Value-to-Transaction Ratio (NVT)

NVT = network value / daily trading volume. NVT is a valuation ratio that compares the network value (equals the market cap) to the network's daily on-chain transaction volume.

Main argument: Similar to the popular equity P/E valuation ratio (either stock price / earnings per share, or market cap / total earnings), NVT may indicate whether a network token is under or overvalued by showing the market cap relative to the network's transaction volume, which represents the utility that users derive from the network. When the ratio becomes very high, it indicates potential token over-valuation.

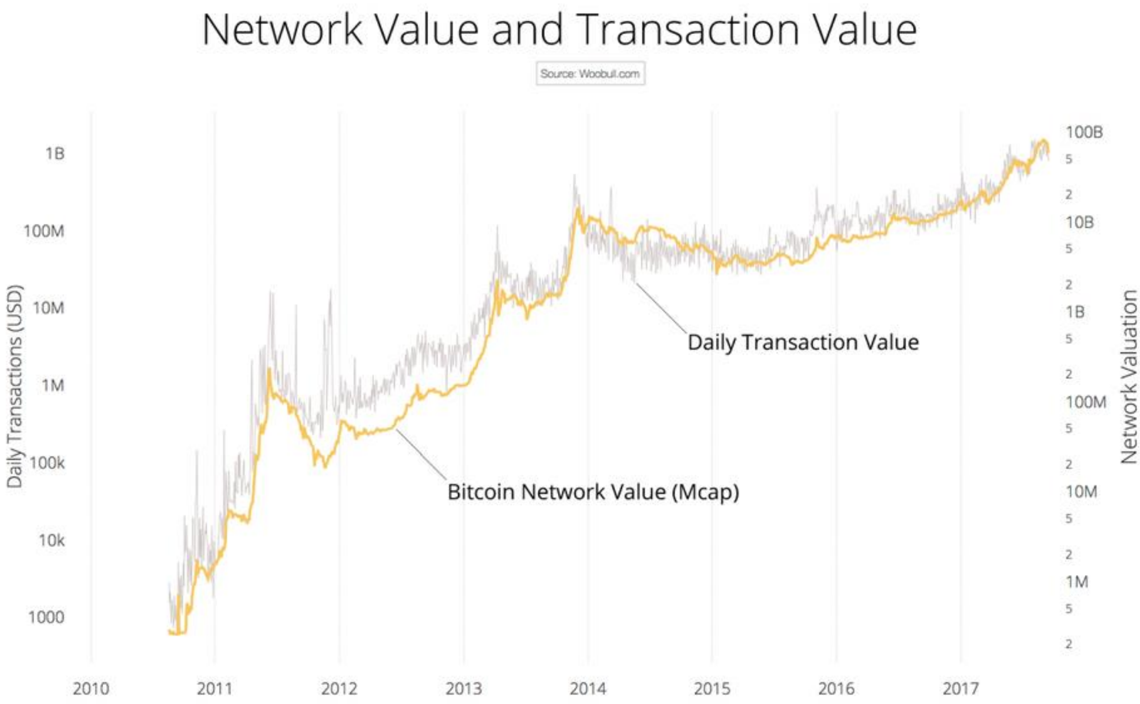

Figure 1: Evolution of Bitcoin transaction volume and Network Value

The ratio best applies to assets whose on-chain transaction volume closely represents utility to users. For instance, bitcoin's on-chain transaction volume represents the utility it provides to users to send money internationally for very low fees and a degree of anonymity. For networks with high levels of transaction detail privacy such as Monero and Zcash, the ratio is undefined. For networks with staking rewards such as Dash, transaction activity resulting from staking would inflate the denominator, inadvertently causing the ratio to be underestimated. This effect could be corrected for by subtracting staking activity from transaction volume.

Criticisms:

- Transaction volumes tend to follow changes in price, so the two variables have an endogenous and "reflexive" relationship, weakening the indicative power of the ratio.
- Thought leaders have experimented with the time frame used to measure daily transaction volume. See the references below for some of this analysis.

2. BUILDING A VALUATION MODEL

As an investable asset, the current value of the cryptocurrencies can be expressed as the expectation of the future value.

$$V_{t1} = E(V_{t2}|F_{t1}) \qquad (1)$$

Where t1 < t2, F_{t1} represent all the public information which impact the cryptocurrency valuation at time t1. F_{t1} can be further divided into two categories: qualitative information and quantitative information. Our Model will focus only on quantitative information.

2.1. Qualitative information

Qualitative information is a general term, which include all the information hard to be quantified. For instance, if the government initiate a ban or a regulation rule which may have a negative impact on the crypto liquidity, capacity, or future applications. The crypto value will generally be decreased. If a new technique or a new potential penetration in the existing payment system has been revealed publicly, the crypto value will go up due to that favor news.

Those kinds of event may impact on certain group of cryptos, or Just impact on a specific crypto. Statistically, the cryptocurrencies are positive corelated, therefore the news influence will diffuser over a very short, delayed time cross-sectionally. Although those impact are hardly quantified, someone could use the technique such as Support Vector machine SVM to build a math model to predict impact (positive or negative) of the public written news. We will extend this topic in the discussion section.

2.2. Quantitative information

Here, quantitative information means the historical numerical data. Such as the historical prices (bid, ask and close), trading volumes of the cryptocurrencies. There are many math methods which can be used to predict future value of the cryptocurrencies.

We have tried different features or combinations of the features. Here is an example our findings to predict the future movement of the cryptocurrencies.

3. MELANION VALUATION MODEL

Based on historical behavior of cryptocurrencies and their dynamics, Melanion built a model that decomposes the movement of the price into a long term component and a cyclical one.

3.1. Cryptocurrency Price Dynamic

According to the historical behavior, the current value of the crypto is the future price based on current information. Therefore

$$dP_t = \tau_t P_t dt + \sigma_t P_t dW_t \qquad (2)$$

Where P_t is the price of the crypto, W_t is a Wiener Process. σ_t is the volatility varied τas a function of time. τ_t is the price trend. According to the historical prices of cryptocurrency, the Price long term trend τ_t is much bigger than the short term price volatility for the last two years.

3.2. Long term trend

We can further decompose the long-term trend into:

$$\tau_t = \varphi\pi \qquad (3)$$

Where, φ is the trend amplitude (normalized by Price), $\pi \in \{-1,1\}$ is the trend directions. Here we note π_- represent down trend, π_+ represent up trend.

If we just take long term trend factor into account for the price movement, the expected price movement can be expressed below

$$E(\Delta P_{t+\Delta t}) = E(\tau_t)\, P_t\Delta t = E(\varphi)E(\pi)P_t\Delta t \qquad (4)$$

Here, we impose that coming trend amplitude and coming trend direction should be independent. The expected coming trend direction can be calculated below

$$E(\pi) = P(\pi_+) - P(\pi_-) \qquad (5)$$

By definition,

$$P(\pi_+) + P(\pi_-) = 1 \qquad (6)$$

By predicting the up or down trend probability, the coming trend direction can be estimated. We propose the trend direction probability can define as the following expression:

$$P(\pi_\pm) = 0.5 + k_\pm 1_{S_t > f(\sigma_s)} \qquad (7)$$

$k_\pm$ is the empirical factors, which can be calibrated with the historical data. S_t is the week stationary time series which constructed via crypto price series; $f(\sigma_s)$ is a utility function based on volatility σ_s.

$1_{S_t > f(\sigma_s)}$ is the indicator function, when $S_t > f(\sigma_s)$, it equal to 1, else is zero

We have found an interesting S_t time series, which exhibits a high predictability for the trend direction of certain Cryptocurrency.

$$S_t = \frac{Average\left(P_{t_{60D}}\right) - Median\left(P_{t_{60D}}\right)}{\sigma_{P_{t_{60D}}}} \qquad (8)$$

$Average\left(P_{t_{60D}}\right)$ is the Average Price for the last 60 days (including current day); $Median\left(P_{t_{60D}}\right)$ is the median Price for the last 60 days; $\sigma_{P_{t_{60D}}}$ is the standard deviation of the last 60 days Price.

4. FILTERS IMPLEMENTATION

In this part, we will explore different ways to refine the model in order to improve its predictive power. To do so, we will apply to the signal different filters with the goal to capture a trend and a cyclical component.

4.1. Linear Filter

We denote by y_t the ordered sequences of observations of the process Y_t.Let x_t be trend process. y_t process can be treated as the sum of two unobserved parts:

$$y_t = x_t + \varepsilon_t \qquad (9)$$

ε_t is non trend process (or noise). Generally speaking, there is no precise definition for the trend process, but it is accepted to be a smooth function representing long movements. It means that changed in the trend x_t must be smaller than those of the process y_t. From statistical standpoint,

$$\sigma(y_t) \gg \sigma(x_t) \qquad (10)$$

Since the process y_t also contains the non trend component ε_t . Technically it is quite difficult to estimate x_t . Here we introduce filtering concept. To simplify the discussion, we take ε_t as a white noise. This is also correspondent to the dynamic of the crypto assets.

$\hat{x}_t$ is the estimator of the underlying trend process x_t. L denotes a linear filter for the process y_t , therefore

$$\hat{x}_t = Ly_t \qquad (11)$$

Since the filter has to be causal, equation (11) can be written as below

$$\hat{x}_t = \sum_{i=0}^{n-1} L_i \, y_{t-i} \qquad (12)$$

With this notation, a linear filter is characterized by a window kernel L_i and its support. The kernel defines the type of filtering, whereas the support defines the range of the filter for instance, if we take a square window on the compact support [0, T] with

$$T = n * \Delta t \tag{13}$$

T being the width of the window.

4.2. Moving average filters

Here we introduce the well known moving average filter:

$$L_i = \frac{1}{n}\mathbf{1}\{i < n\} \tag{14}$$

In this case, only the window support n need to be calibrated. If the window T is very large, after the filtering, the sequence will exhibit is long term trend. If the window T is very small comparing to the observed sequence, the treated sequence will be very close to the observed sequence. In this case, the filter become un-necessary.

For T > 0, if ε_t and x_t are independent,

$$\hat{x}_t = 1/n \sum_{i=0}^{n-1} x_{t-i} \tag{15}$$

If the trend μ_t is constant during certain period, then

$$\mu_t = d\hat{x}_t/dt \tag{16}$$

4.3. Moving average crossovers

In practice, we use crossover technique to extract the trend μ_t. This trend is estimated form the difference between two moving averages over two different time windows n_1 and n_2 . Supposing that $n_1 > n_2$, the trend μ_t can be estimated from

$$\hat{x}_{t,n1} = 1/n_1 \sum_{i=0}^{n_1-1} x_{t-i} \tag{17}$$

$$\hat{x}_{t,n2} = 1/n_2 \sum_{i=0}^{n_2-1} x_{t-i} \tag{18}$$

$$\mu_t = 2(\hat{x}_{t,n2} - \hat{x}_{t,n1})/(n_1 - n_2) \tag{19}$$

In particular, the estimated trend is positive, if the short term moving average is higher than the long term moving average. Therefore, the sign of the trend changes when short term moving average cross the long term moving average.

The main advantage of using a moving average filter is the reduction of the noise due to the central limit theorem. For the limit case $n \rightarrow \infty$, the signal is completely denoised, but it corresponds to the average value of the trend. In practice, the estimator is also biased (the historical characteristic does not guarantee the future trend). In trend filtering, we also face a trade off between denoising maximization and bias minimization.

4.4. Kernel for Time window

Equation (14) shows that the windows is uniform. In practice, the different Kernel function can be applied for the time window according to the various requirement. For instance, one possibility is to take an **asymmetric window** function with a triangular form:

$$L_i = 2(n-i)/n^2 \, \mathbf{1}\{i < n\} \tag{20}$$

There are other kernel functions shown in Figure 1. The asymmetric window should more adapted to the reality.

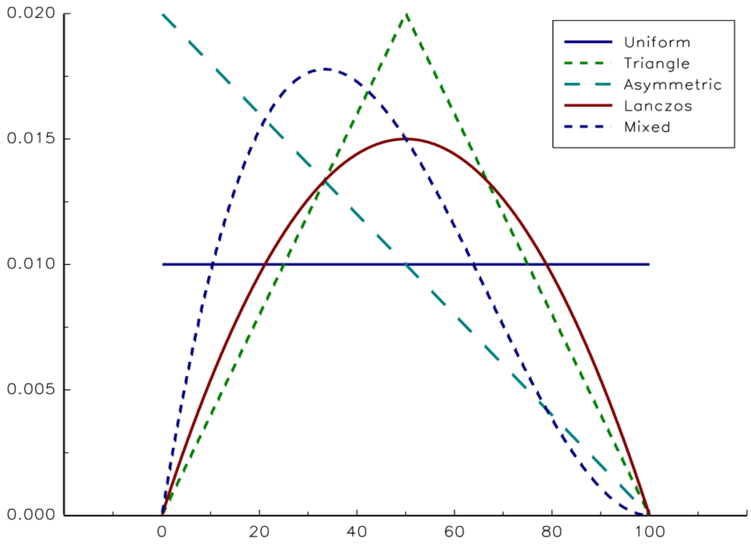

Figure 2: Kernel Functions for Time Window (n=100)

4.5. Hodrick-Prescott filter to capture the trend

Before, we decompose the time sequence into the sum of trend part and noise part. Here we introduce another scenario. We could decompose the observed signal $y_t$into two components: one cycle part c_t (which is short-term) and one trend part x_t (which is medium-to-long term).

$$y_t = x_t + c_t \tag{21}$$

To estimate trend, we introduce so-called Hodrick-Prescott Filter. Hodrick Prescott filter is a bandpass filter where it tries to decompose the time-series signal into a trend (mid-term growth) and a cyclical component (recurring and seasonal signal).

The loss function that it tries to minimize is the following:

$$\min_{x_t} \sum_{t=1}^{n} (y_t - x_t)\,(y_t - x_t)^2 + \lambda \sum_{t=1}^{n} (x_{t-1} - 2x_t \; + \; x_{t+1})^2 \qquad (22)$$

The first term is the square of difference of observed signal and trend signal (cyclical component). In this function λ is the regularization parameter which controls the competition between the smoothness of x_t and the cylical component $(y_t - x_t)$. We may rewrite the objective function in the vectorial form:

$$\min \parallel y - x \parallel_2^2 + \lambda \parallel Dx \parallel_2^2 \qquad (23)$$

Where

$$\mathrm{y} = (y_1, \dots y_n) \quad , \; \mathrm{x} = (x_1, \dots x_n) \qquad (24)$$

and D operator is the (n-2) * n matrix

$$D = \begin{bmatrix} d & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & d \end{bmatrix} \qquad (25)$$

d is the constant operational vector (1, -2, 1).

The estimator is then given by the following solution:

$$x = (I + 2\,\lambda\, D^T D)^{-1}\, y \qquad (26)$$

Change the smoothing parameter λ, you could actually change what type of effects you may want to include or capture (if you want to capture some variation and volatility in short-term signal, then you may want to use a smaller smoothing parameter so that you have less smooth signal. If you want to also capture only a long-term range signal, the smoothing parameter could be chosen arbitrarily large.

However, to get some changes, we need to not to choose very large smoothing optimization parameter.

However, according the equation (22), the standard Hodrick-Prescott filter is non-causal filter. To adapt the reality (Never Use the future information to predict the history), we need to modify the smoothness of x_t into $x_{t-2} - 2x_{t-1} \; + \; x_t$. which is one time step backward. Here one need to be clarified, when we choose the larger smoothing parameter λ, we actually try to extract the longer-term trend. The residual cycle component will be more volatile. While we choose the smaller smoothing parameter λ, the extracted trend become shorter, the cycle component become less volatile. In this case, the process will be more sensitive to the smoothness. one need to more careful to use one step back smoothness.

5. DISCUSSION

As of today, there is no cryptocurrency pricing model that is widely accepted by every actors of the crypto ecosystem. Indeed, given the lack of long term historical data and the unpredictability of this kind of asset, it is not easy to evaluate or predict the future movements of a cryptocurrency. Moreover, the existing models, such as the Token Velocity Thesis and the Network Value-to-Transaction models have lots of drawbacks.

We propose to use the historical market data to predict the value of the crypto and believe this may pave a new way to fully assess the value and to understand the dynamic of the crypto's value.

Currently, we are still at the beginning stage for this project. We just show some promising results. There are lots of the related projects can be explored, for instance using Natural Language Processing to transfer the non-quantified information into the simple impact indicator for the crypto valuation could be an interesting project and add value to this methodology.

Melanion Capital
17 Avenue Georges V
75008 Paris

For info:
Contact@melanion.com