Micro Data Analysis and Forecasting of Cryptocurrency Markets

Alfred Holmes

September 2019

1   Introduction

Cryptocurrencies are typically traded on order-driven markets in which market participants place limit and market orders. Limit orders can be bids or asks, bids being offers to buy at a particular price and asks offers to sell. Market orders are orders executed against the best available limit order prices (buying at the lowest ask or selling at the highest bid). If a limit order is placed that overlaps an existing limit order on the other side of the market, for example there is an offer to buy 1 BTC for 2 USD and someone offers to sell 1 BTC for 1 USD then the second limit order will be considered by the exchange as a market order and the trade will be priced at the original limit order. Order-driven markets are now the standard way that stocks, futures and currencies are traded. Two benefits of studying cryptocurrencies over traditional financial instruments is that the exchanges freely give out discretized order data and that the markets are always open, except for the odd update here and there.

There are two types of trades that occur on cryptocurrency markets, high frequency (risk free or low risk) trades and trades to facilitate longer term (risky) investments. Typically high frequency trading strategies can be completely hedged and therefore risk free, but some traders may use statistical models or gambling to try to increase profits by not completely hedging. Examples of high frequency trades include market making and arbitrage. Market making is where an individual provides liquidity by offering to both buy and sell an asset. They place limit orders in such a way as to always buy and sell the same amount. Market makers make money by having a difference in the prices at which they buy and sell in exactly the same way as a currency exchange in an airport, for example. When a trade occurs on an exchange there is a maker and a taker, the maker being the party that provided the market liquidity by placing a limit order and the taker the party that is removing (taking) liquidity from the market. The presence of market makers gives meaning to which side of the market the maker placed the limit order on as typically the taker will be from a longer term investment. If, for example, there is a higher rate of market orders to buy than sell then the price is likely to increase even though on the exchange exactly the same amount of the asset is being bought and sold. The mechanism for this price increase is that the market makers will either raise their prices to sell due to the increased demand or increase their buy prices to increase the chance that a seller will sell to them and so they will reduce their exposure. Arbitrage is a different high frequency technique where traders (using automated systems) look for discrepancies in the pricing of assets to make risk free profits. This can be achieved by making a trade and on a different exchange the reverse of that trade or by trading in a loop on one exchange that offers different quote assets. For example, if the price of 1 BTC was 1 USD, 1 ETH was 2 USD and there was an offer to buy 1 ETH for 1 BTC then the arbitrage trader would spend 1 USD on 1 BTC, trade the 1 BTC for 1 ETH and then sell the 1 ETH for 2 USD, making 1 USD profit. It is through this mechanism that the markets on different exchanges keep the same prices and the prices on one exchange make sense, unlike the example given.

Each market on each exchange is a complex system with external influences, primarily other markets and news relating to the traded assets. This makes the study and forecasting of cryptocurrency market dynamics rather complicated as there are many exchanges each offering many markets, with any trade on any one of the thousands of markets perhaps influencing the whole system. It is reasonable however to only focus on the markets with the highest trading volume and assume that arbitrage causes the less liquid markets to follow the higher trade volume markets for the simple reason that the lower volume market will be more easily moved by the arbitrage traders. In this sense arbitrage trades act as a proxy for market movements that happen on one market to simultaneously occur on other markets. Obviously it is the case that each market would behave differently without the presence of some other market but it makes sense to focus on markets with a high proportion of the traded volume as should be the markets driving the market movements.

2 Data

2.1 Orders

In this study data was acquired from the exchanges Binance and Coinbase, accounting for approximately 17% and 2.6% of the global spot volume respectively at the time of writing[3] although these numbers are likely to have been quite different in the time frames studied[1]. Both exchanges offer REST APIs[1][2] which give historical trade data. Each market order (that is either a market order executed on an exchange or a limit order that overlaps with the other side of the order book and is hence reported as a market order) results in a series of trades where liquidity is removed from the exchange. The exchanges record and publish the time, volume, price and the side of the order book that liquidity is being removed from for each of these interactions. The scripts[] in the github repository can be used to download this data although it takes some time due to rate limits. In order to analyse the orders, the raw data from the exchanges is downloaded and stored in a database and then orders are that occur at the same time and remove liquidity from the same side are grouped together[Github library].

2.2 Order Book Data

Typically cryptocurrency exchanges do not offer historic order book data through their APIs. Some websites[] allow the purchase of this data but the quality and completeness cannot be guaranteed and in the authors view their products are overpriced. Most exchanges do however offer Web Socket Streams of live exchange data so it is possible to record the order book through time. An example script doing this can be found in the github repository[Web Socket Script]. Decentralized exchanges in which orders are written to a blockchain, like Bitshares[Bitshares], clearly allow the access of historical order book data, but these have not been explored due to the low volume and the lack of high frequency trading on these exchanges, making them fundamentally different from the main high volume price driving exchanges studied.

3 Empirical Analysis

3.1 Market Movement Time Series

From each order, it is possible to calculate the price changed caused by that order. That is, if the order occurs at time t and $S_{t_{+(-)}}$ is the last traded price just after (before) time t then the price change caused by order i, R_i, is given by $R_i = S_{t_+} - S_{t_-}$. Let $r_i = \log R_i$

3.2 Order Rates

As explained below, a key statistic when modelling an order-driven market is the order rate.

3.3 Price Change Variance

4 Diffusion Models

A simple model for a market, taking inspiration from physical Brownian motion in 1 dimension is a diffusion model[Economic Fluctuations and Diffusion]. The idea is that the price of the asset is modelled as the position of a pollen grain floating on water and each trade represents the collision of a water molecule and the pollen. Although this model is simple, it is used to base the understanding of the fundamental market mechanics assumed in this study. Suppose that orders occur at a constant rate λ and change the price by ϵ_i, which is some random variable with mean 0 and variance σ^2, assuming independence of the ϵ_i the price change $S_{\Delta t}$ in a given interval Δt is given by

$$S_{\Delta t} = \sum_{i=0}^{\lambda \Delta t} \epsilon_i \tag{1}$$

By the central limit theorem,

$$\frac{S_{\Delta t}}{\sqrt{\lambda \Delta t}} \sim \mathcal{N}(0, \sigma^2) \tag{2}$$

Using this to produce forecasts requires the prediction of the number of trades and the variance of each price movement produced by those trades. Some obvious flaws with this approach are that negative prices occur with a non 0 probability and the order rate and price change variance of the orders are not constant, so it is better to

---

[1]The website coinmarketcap.com can be used to track exchange volume through time, although the reported volumes may be inaccurate with some exchanges reporting fake volume. We haven't perused the tracking of the studied exchanges' volume through time

consider $S_{\Delta t}$ as the change in the logarithm of the price and try to predict the rate and variance based on the current market state. Hence we have that

$$S_{t_2} - S_{t_1} = \sum_{i=0}^{\int_{t_1}^{t_2} \lambda(t)dt} \epsilon_i \sigma_i \tag{3}$$

4.1 Altcoin-Bitcoin Correlations

The diffusion model offers an explanation to the observed correlations in the prices of cryptocurrencies against traditional currencies[Correlations of cryptocurrencies]. Typically higher volume (more individual trades) are observed between altcoins and Bitcoin than altcoins and traditional currencies. The diffusion model would suggest a higher volatility in the market with more orders being placed, so one should expect to see the price of an altcoin being driven by the Bitcoin market and then the altcoin's traditional currency markets responding through arbitrage trades. This would lead to a high correlation between the Bitcoin price and the altcoin price against the traditional currencies and indeed, this is what is observed.

4.2 Arbitrage Trading Between Exchanges

An issue with studying individual orders to forecast the price of an asset is that different exchanges generally keep the same prices through arbitrage trading, so a lot of information about the global market is missing by focusing on one or two exchanges. This is a fair criticism and the effects of arbitrage trading on the individual exchanges' markets needs to be studied further, as well as the effects of futures trading on the underlying asset price. Since Binance holds the largest share of Bitcoin spot volume we make the assumption that arbitrage occurs on other exchanges moving the price towards Binance's in a similar way to the altcoin markets described above and so the aribtrage effects on Binance's market are negligible.

References

[1] *Binance REST API*. URL: https://github.com/binance-exchange/binance-official-api-docs/blob/master/rest-api.md. (Accessed: 12-08-2019).

[2] *Coinbase Pro REST API*. https://docs.pro.coinbase.com/.

[3] *FTX Volume Monitor*. https://ftx.com/volume-monitor. Accessed: 12-08-2019.