

# 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. This is the standard way that stocks, futures and currencies are now traded. One benefit of studying cryptocurrencies over traditional financial instruments is that they freely give out discretized order data so that the markets can be studied and eventually modeled on an order by order basis. This study examines some key statistical properties of the available data and proposes novel methods for the modeling of order-driven markets.

## 2 Data

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[FTX] although these numbers are likely to have been quite different in the time frames studied. Both exchanges offer REST APIs[Binance API][Coinbase API] 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 of each of these interactions. The scripts[] 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].

## 3 Diffusion Models

A simple model for a market, taking inspiration from physical Brownian motion in 1 dimension is the diffusion model[Economic Fluctuations and Diffusion]. The idea is that the price of the asset 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  $\lambda$  and change the price by  $\epsilon_i$ , which is some random variable with mean 0 and variance  $\sigma^2$ , assuming independence of the  $\epsilon_i$  the price change  $S_{\Delta t}$  in a given interval  $\Delta t$  is given by

$$S_{\Delta t} = \sum_{i=0}^{\lambda \Delta t} \epsilon_i \quad (1)$$

By the central limit theorem,

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

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 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.