

Predicting MEV as a Continuous Variable

Team Code w7eo5guuwmvr67mzm8zw7eo59aakjv2g

September 2022

Abstract

In this study, we predict maximal extracted value (MEV) as a continuous variable. We use an autoregressive integrated moving average (ARIMA) model to analyze how previous blocks' MEV affects future blocks' MEV. Because the magnitude of MEV in a block may be influenced by off-chain events, such as market activity on centralized exchanges, we investigate the effect of off-chain exchange activity on MEV.

1 Data

We first pulled data from high-volume centralized exchanges such as Binance and FTX. We gathered every trade from January 1, 2021 to July 1, 2022. We chose to begin on January 1, 2021 because Flashbots released in January 2021. We decided it was prudent to gather every trade during that time because we are able to similarly inspect transactions with great granularity on the blockchain. The asset pairs we decided to research were ETH/USD, BTC/USD, and USDT/USD. These pairs were chosen because they have high daily volume on both centralized and decentralized exchanges. We believe that because those pairs are traded at much higher volumes on centralized exchanges (CEX), price discovery takes place in those venues, and the prices on decentralized exchanges follow after a short time. Our CEX data included the price, timestamp, and volume of each trade, as those factors are to be considered in our model. Our on-chain data was a time series of the magnitude of MEV, using timestamps instead of block heights.

2 Analysis

Our model is an autoregressive integrated moving average (ARIMA) model. We were not able to get to this stage of the study due to the teams' time constraints. However, we have the following recommendations. We will not be able to auto-fit an ARIMA model as that would create some statistical problems. What we recommend is to compare a smoothed series of both datasets and determine

if there are any lag effects in either of the time series. We can then simulate various times by slicing the dataset and see how much of a difference we get in the sum of squares metrics for both MEV and off-chain price data.

3 Conclusion

Unfortunately we were not able to complete this study, thus we do not yet have a conclusion. However, we will continue on this project under less pressure and reach a conclusion. The results of this study could possibly be implemented to prevent on-chain adverse selection. If there are signals that MEV (in other words, inefficiencies) in on-chain markets, market participants may be able to take steps to correct their positions before being adversely selected.