BID/ASK Methodology

Version 1.0

BLOCKSIZE produces the BID/ASK Data Feed as part of its product BLOCKSIZE CONNECT, a collection of rates quoted in U.S. dollars, euros, and various other currencies for an expanding set of cryptocurrencies and other digital asset classes.

The BID/ASK PRICES are designed to serve as a transparent and independent pricing source that promotes the functioning of efficient markets, reduces information asymmetries among participants, facilitates trading, and accelerates the adoption of cryptocurrencies as an asset class with the highest standards.

The BID/ASK PRICES Data Feed is calculated using a robust and resilient methodology that is resistant to manipulation and adheres to international best practices for financial benchmarks.

This methodology is governed by the BLOCKSIZE Data Committee and is part of the BLOCKSIZE CONNECT Manifest (see Appendix A.1) that aims to ensure that the BID/ASK PRICES Data Feeds serve as an accurate source of transparent and reliable pricing.

Contact email: connect@blocksize-capital.com

1 Calculation of Volume-Weighted Average Bid/Ask

The calculation algorithm of the BID/ASK PRICES Data Feed is as follows:

1. Calculate the volume-weighted sum of bid (superscript b) prices $\sum_j p_j^b \times v_j^b$ and, respectively, ask (superscript a) prices $\sum_j p_j^a \times v_j^a$ denominated in units of the underlying asset from observable order book updates that occurred for each of the accepted markets (subscripts j)
2. Calculate the aggregate volume for each of the accepted markets by adding the bid and ask size of events across all accepted exchange markets, respectively, $\sum_j v_j^b$ and $\sum_j v_j^a$. The resulting figure is referred to as the bid or ask volume weight.
3. Divide the volume-weighted sum of prices, $\sum_j p_j^x \times v_j^x$, by the total volume weight, $\sum_j v_j^x$, for bids ($x \equiv b$) and asks ($x \equiv a$) to obtain aggregated BID/ASK PRICES π^x:

$$\pi^x = \frac{\sum_j p_j^x \times v_j^x}{\sum_j v_j^x}$$

2 Data Contingency Rules

The following contingency rules are followed to address situations where data is delayed, missing, or unavailable due to periods of low liquidity such as extraordinary market circumstances or outside factors beyond the control of BLOCKSIZE.

1. If observable order book updates from an accepted market are unable to be collected due to technical problems specific to the accepted market's exchange during the calculation of the BID/ASK PRICES Data Feed (such as malformed data), the observable order book updates are excluded from the calculation of the specific instance of the given Price Data Feed.
2. If no observable order book updates from an accepted market exist during the current timeframe, the value of the BID/ASK PRICES Data Feed will rely on the on various other accepted markets for its calculation.
3. If none of the accepted exchanges in Appendix A.3 report observable order book updates, the BID/ASK PRICES Data Feed will not report a calculated price. In the highly unlikely event that a user establishes a connection to a BID/ASK PRICES Data Feed in a timeframe where none of the accepted exchanges (Appendix A.3) is reporting observable order book updates, the price will be computed from the last observed order book updates.

3 BID/ASK PRICES Data Quality Assurance

All observable order book updates from accepted markets are evaluated using algorithmic anomaly detection methods. If potential errors or anomalies in the transaction data are detected, it is not included in the calculation of the BID/ASK PRICES Data Feed. The anomaly detection methods are applied to transaction data collected within liquidity-adjusted timeframes of a maximal duration of one minute.

Complementary to the algorithmic anomaly detection approaches, the BLOCKSIZE Data Committee weighs in its expert judgment to maintain and improve the data quality of the BID/ASK PRICES Data Feeds. The Data Committee may decide to include the transaction data from new sources that are vetted for data integrity and quality. Alternatively, the Data Committee may decide to exclude previously accepted markets that started to deliver erroneous or anomalous data. Any exercise of such expert judgment must be approved by staff members.

3.1 Overview: Data Processing and Anomaly Detection Methods

BLOCKSIZE uses mathematical statistics to automatically perform the task of ensuring high data quality for the BID/ASK PRICES. The anomaly detection and data processing uses liquidity-adjusted timeframes to ensure the accuracy and contemporariness of confidence intervals as well as the computed BID/ASK PRICES.

The approach to create adaptive, liquidity-adjusted timeframes is specified in Subsection 3.2. . In Subsection 3.3. the used Price-Based anomaly detection approach is introduced. This filter is based on the computation of statistical estimates for confidence intervals for order book updates. Since this price-based method establishes a filter for individual events with anomalous prices we label it “Price-Based Filter”.
For the BID/ASK PRICES Data Feeds, BLOCKSIZE considers order-book data timeframe of at most 30 seconds and applies the Price-Based Filter in a liquidity-adjusted window of at most one minute in length.

3.2. Adaptive Liquidity-Adjusted Timeframes

To ensure that anomaly detection and computational methods can be applied reliably across all of the pairs in BLOCKSIZE’s REAL-TIME-PRICES product, the timeframes considered are dynamically adapted to the number and volume of order book updates that are observed for a given pair.

For this purpose, the timeframe for each pair, $\rho = [t - \tau, t]$, is adjusted dynamically, $\rho \equiv [t - \tau(t), t]$, attempting to fulfill the following conditions at all points in time:

1. There are $\mathrm{N} \geq 10$ events in the considered window ρ

2. The aggregate volume of the bids ($x \equiv b$) and asks ($x \equiv a$) events $\sum_j v_j^x$ in the window ρ is at least USD 1000 or an equivalent amount, if the quote currency is not USD
3. All included events in the window $\rho = [t - \tau(t), t]$ are less than 60 seconds old, i.e., $\tau(t) \equiv 1min. \forall\, t \geq 1min.$
4. There are at most 1000 bid / ask updates in the timeframe ρ

Conditions 1. and 2. are considered necessary the accuracy of statistical estimates used for the computation of the confidence interval for valid BID/ASK PRICES that are used for the anomaly detection approach described in the following Subsection 3.3. .
Conditions 3. and 4. exclude data that is not sufficiently recent from the consideration for the anomaly detection of the BID/ASK PRICES Data Feeds.

It is worthwhile to note that a timeframe $\rho = [t - \tau(t), t]$, due to the application of Conditions 3. and 4., may not fulfill the prerequisite conditions 1. and 2. for the application of anomaly detection methods. This unlikely Edge case is discussed together with other edge cases in Subsection 3.4. .

3.3. Price-Based Filter Method

For the construction of the Price-Based Filter, the deviation of the price of each new order book update observation from the mean of the prices of all other non-anomalous order book updates in a liquidity-adjusted timeframe. This comparison is done by tracking the standard deviation of all prices in the tracked in the adaptive Price-Based Filter timeframe.

The Price-Based Filter approach is applied in precisely the same way for the bid and ask update observations. Thus, for notational convenience, the superscript denoting bid prices or ask prices is dropped from the equations in the remainder of this Subsection.

Consider the order book updates with prices $\theta = \{p_1, p_2, \dots, p_N\}$ in the window $\rho = [t - \tau, t]$, i.e., $\theta = \{p_1, p_2, \dots, p_N\}$ with $t_i \in \rho$.

To decide if a new price $p_{N+1}$is anomalous or can be included, the standard deviation σ_θ and the average $\bar{p}_\theta$ of prices in θ are computed:

$$\theta = \{p_1, p_2, \dots, p_N\}$$

$$\bar{p}_\theta = \sum_{i=1,\dots,N} \frac{p_i}{N}$$

The confidence interval Ψ_θ for the inclusion of a new price p_{N+1} into the set θ of prices that are considered for the computation of BID/ASK data is obtained by adding/subtracting a multiple ξ of the standard deviation σ_θ from the average $\bar{p}_\theta$ of prices that are already in θ:

$$\Psi_\theta = [\, \bar{p}_\theta - \xi\sigma_\theta, \bar{p}_\theta + \xi\sigma_\theta]$$

If the new price event $p_{N+1}$is within the confidence interval Ψ_θ, it is accepted as valid and otherwise discarded as anomalous.

The value of the constant ξ determines the size of price deviations to be considered anomalous; BLOCKSIZE BID/ASK PRICE data feeds found $\xi = 3.0$ a good choice to maintain exceptional quality standards.

3.4 Edge Cases and Applicability of Anomaly Detection Methods

The Price-Based Filter method is applicable only in the case that the confidence intervals they use for filtering anomalies are accurate. BLOCKSIZE empirically found that the fulfillment of condition 1. -- $N \geq 10$ observed events -- and condition 2. -- more than USD 1000 or equivalent in observed volume (cf. Subsection 3.2.) renders the performance of the Price-Based Anomaly Detection approach in Subsections 3.3. to be sufficiently robust.

In the rare case that the conditions 1. and 2. are not fulfilled, market conditions are considered anomalous as reliable confidence intervals for the validity of order book event observations cannot be established. For such anomalous situations, the Price-Based Filter is not applied and all available non-erroneous order book update observations are included for the computation of the BID/ASK PRICES Data Feed.

Following rare abrupt price jumps to prices outside the confidence interval of the Price-Based Filter can trap the approach into excluding all following trade event observations as anomalous. For rectifying this flaw of the method to fail to update its confidence interval, a detector for price jumps was added to the Price-Based Filter approach. This detector checks if $N \geq 4$ outliers are detected in sequence above or in sequence below the confidence interval with $\geq 500\ USD$ or equivalent of aggregated volume. If these price jump conditions are met, the detector triggers a re-initialization of the confidence interval of the Price-Based Filter approach. The confidence interval is recomputed including the sequence of $N \geq 4$ order book updates representing the sudden price action that have erroneously been classified as outliers. The re-initilization maintains the construction rules for the Adaptive Liquidity-Adjusted timeframe ρ as specified in Subsection 3.2. .

Appendix: Current Documents

The BID/ASK PRICES Methodology described in here is part of the BLOCKSIZE CONNECT Manifest, which is describing how BLOCKSIZE is dealing with aspects of its BLOCKSIZE CONNECT suite of data subscriptions.

A.1 Current version of the BLOCKSIZE CONNECT Manifest:
https://www.blocksize.info/blocksize-connect/manifest/

A.2 Current list of supported instruments:
https://www.blocksize.info/blocksize-connect/instruments-realtime/

A.3 Current list of supported markets:
https://www.blocksize.info/blocksize-connect/markets-overview/

A.4 Current version of the BID/ASK PRICES Methodology:
https://www.blocksize.info/blocksize-connect/manifest/real-time-prices-methodology/

Date: 2023-11-21