

A new approach to Level 3 market data

Putting the ‘buy’ into ‘build versus buy’

When Optiver undertook a new quantitative trading project several years ago, we quickly ran into a problem: Where to source the data? Quant strategies rely heavily on historical prices to detect patterns and identify and execute trades, making solid market data a must. But in Europe, where liquidity is fragmented across more than two dozen countries and exchanges, this data can be particularly challenging to collect.

High-quality historical data is something financial institutions increasingly can’t do without – and market makers like Optiver are no exception. By allowing us to spot patterns, test hypotheses and refine our algorithmic models, data is critical for helping Optiver provide quick and accurate pricing across thousands of financial instruments. Extensive datasets allow us to backtest strategies across a vast range of market conditions, ensuring they’re sufficiently robust before we deploy them in live trading.

But historical data at this level of granularity isn’t easy to manage. One vendor we approached early on provided us with normalised and un-normalised datasets, but the former wasn’t up to Optiver standards, and we found the un-normalised product difficult to deal with and costly to add to.

“High-quality historical data is something financial institutions increasingly can’t do without”

This is the story of how we became, first, customers of BMLL’s Level 3 historical data and, later on, investors. It’s also the story of how companies like BMLL, with the help of Optiver, are finally putting the ‘buy’ into the age-old ‘build versus buy’ question when it comes to historical market data.

Levels 1, 2 and 3

For most investors, Level 1 historical data – the best bid/ask price and last trade price in an instrument – is usually all that’s required. Many institutional investors find Level 2 data more suitable for their needs, which in addition to what’s available in Level 1, offers a deeper look into the order book.

But firms like market makers and quantitative hedge funds require access to the most granular and comprehensive market data, known as Level 3.



An easy way to think about Level 3 data is that it includes the full depth of the order book. This goes beyond the best bid and ask prices for a particular instrument to include multiple levels of buy and sell orders. It also includes all outstanding bid and ask orders across a range of price levels – the so-called price ladder. Level 3 data also displays the total number of shares, contracts, or lots available to buy or sell at each specific price point in the order book.

Level 3 data shows message traffic through which participants send cancellations or modifications of their orders. Some exchanges capture every trading intention (such as update, cancel, remove) in their data, providing an additional layer of information to users. Moreover, a user of Level 3 data can understand whether a trade was a buy or sell, without requiring inference, due to clearly identified aggressor sides.

Depending on the exchange, other key features of Level 3 data can include:

- Sequence numbers (which provide the true ordering, rather than just timestamps);
- Auction imbalance messages; and
- Exchange specific fields

The challenges

Trading firms looking to gather, clean and normalise this data face enormous challenges. Redundancies, for one. Many data vendors, including exchanges, provide new ‘flags’ without retiring old versions, meaning the data can contain 30+ auction flags that mean very similar things. There’s also a wealth of inconsistent information to sift through. Trading venues use their own data formats based on their matching engine and

multicast system versions. The best bid price, for instance, can be called ‘best bid’, ‘bid1’ or ‘top bid,’ depending on the venue.

In our early experience, the vendors who offered to help us gather and clean the data weren’t much better. Existing data vendors either cleaned and normalised the data in a way that wasn’t up to Optiver’s standards, or else made it onerous for us to add new countries and exchanges to their existing dataset.

BMLL’s solution

In late 2023, after several months of exploring the market for data vendors, we approached BMLL. Founded at the University of Cambridge, BMLL is an independent provider of Level 3 data and analytics from financial markets across the world. What immediately set BMLL apart from any other data vendor we encountered was the way it normalised Level 3 data. It was a robust process, which made it easy for us to query and analyse.

BMLL builds its data products directly from raw captures and exchange data, which it then stores. Next, BMLL builds a global data model that’s normalised down to Level 3.

What this means is that a small number of values are used for each data field, which remain consistent over time. For example, some data vendors have 31 or more auction types in their data (including both 'Unknown' and 'Undefined' as possible values). By contrast, BMLL's normalised datasets include just five types of auctions.

The data is then made available in two formats, depending on what problem the user is trying to solve:

- Normalised data, built from BMLL's global data model, which includes a small number of fields per venue; and
- Harmonised data, which makes it easy to rebuild and replay the order book consistently across multiple markets, while preserving the raw fields that come from the exchange.

From customer to investor

This was the first time at Optiver that we'd encountered an off-the-shelf product that met our expectations. BMLL's normalisation process was similar to how we would do it ourselves, and it resulted in high-quality data that was easy to use and scale to new markets. When we decided to expand this

quantitative strategy to the US, it took weeks, not months. And BMLL's data is delivered via AWS, which made it simple for us to consume.

“BMLL's normalisation process was similar to how we would do it ourselves”

After starting off as a customer, Optiver became an investor in BMLL last year. As a strategic investor, Optiver can offer our partners lessons learned from decades of trading and deploying tech in the world's most competitive markets. And as a customer of BMLL's, we're able to assist in product development and strategic direction. Today, Optiver and BMLL are working to finally put a compelling 'buy' into the 'build versus buy' choice for market data.

Learn more about BMLL's harmonised Level 3 data here: www.bmlitech.com