

# Fight for the Price at the Middle of the Book

## Research Project

Charles-Albert Lehalle, Sasha Stoïkov

Printed April 16, 2015, version 0.1

It seems to be a consensus on several stylized facts on orderbooks dynamics:

- the imbalance between the two first queue has a predictive power (see Stoïkov);
- even modelling more finely the orderbook than with imbalance (see Huang-Lehalle-Rosenbaum) some volatility seems to be missing; the price is too mean reverting, mostly because conditionally to the full consumption of one first limit, the “discovered one” is really huge;
- this volatility may be retrieved if we implement some “reset” of the orderbook state around the price;
- there is long memory in the signs of transactions.

We would like to understand how these stylized facts intricate. We expect to find around the stopping times at which a quote is fully consumed enough trace of a “fight for the price” among metaorders to offer a consistent understanding of the upper list of stylized facts.

We will restrict our analysis to “large tick assets” in the sense of  $\eta$  being small (i.e. lower than  $1/3$ ). Reminder,  $\eta = N_c/(2N_a)$ , on a daily scale (see Rosenbaum-Robert).

**Descriptive statistics.** The first step is to try misc conditioning between what happens just after the full consumption of a quote and the future price changes. These plain statistics will allow us to understand the most probable configurations after a full consumption. For instance:

- the imbalance in consuming flows (i.e. trades) at the new limit,
- the frequency of large price changes conditionally to the size of the new first limit,
- the difference between the cancellation rates on the new first limits just after a full consumption.

**A first step towards an explanation.** More structurally, we can think about a model with two regions:

- one region, close to small size of at least one of the two queues (say  $Q^a$  or  $Q^b$  is smaller than a threshold  $\theta$ ), in which the *imbalance of transaction flows* dominates.
- the remaining  $(Q^a, Q^b)$  quadrant, where the *imbalance of queues* dominates.

It should not be that hard to compute the value of  $\theta$  associated to the best predictability of the two regions models, and to compared this best predictability to some raw ones.

We may have to put aside the price movement without any transaction consuming one of the two queues.