Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 123 (2013) 2603–2619

stochastic processes and their applications

www.elsevier.com/locate/spa

Estimating the efficient price from the order flow: A Brownian Cox process approach

Sylvain Delattre[a,*], Christian Y. Robert[b], Mathieu Rosenbaum[c]

[a] *LPMA, Université Paris Diderot (Paris 7), France*
[b] *SAF, Université Lyon 1, France*
[c] *LPMA, Université Pierre et Marie Curie (Paris 6), France*

Available online 11 April 2013

Abstract

At the ultra high frequency level, the notion of price of an asset is very ambiguous. Indeed, many different prices can be defined (last traded price, best bid price, mid price, etc.). Thus, in practice, market participants face the problem of choosing a price when implementing their strategies. In this work, we propose a notion of efficient price which seems relevant in practice. Furthermore, we provide a statistical methodology enabling to estimate this price from the order flow.

Keywords: Efficient price; Order flow; Response function; Market microstructure; Cox processes; Fractional part of Brownian motion; Non parametric estimation; Functional limit theorems

1. Introduction

1.1. What is the high frequency price?

The classical approach of mathematical finance is to consider that the prices of basic products (future, stock, etc.) are observed on the market. In particular, their values are used in order to price complex derivatives. Since options traders typically rebalance their portfolio once or a few times a day, such derivatives pricing problems typically occur at the daily scale.

* Corresponding author.
E-mail addresses: delattre@math.univ-paris-diderot.fr (S. Delattre), christian.robert@univ-lyon1.fr (C.Y. Robert), mathieu.rosenbaum@upmc.fr (M. Rosenbaum).

0304-4149/$ - see front matter
http://dx.doi.org/10.1016/j.spa.2013.04.012

When working at the ultra high frequency scale, even pricing a basic product, that is assigning a price to it, becomes a challenging issue. Indeed, one has access to trades and quotes in the order book so that at a given time, many different notions of price can be defined for the same asset: last traded price, best bid price, best ask price, mid price, volume weighted average price, etc. This multiplicity of prices is problematic for many market participants. For example, market making strategies or brokers optimal execution algorithms often require single prices of plain assets as inputs.

Choosing one definition or another for the price can sometimes lead to very significantly different outcomes for the strategies. This is for example the case when the tick value (the minimum price increment allowed on the market) is rather large. Indeed, this implies that the prices mentioned above differ in a non negligible way.

In practice, high frequency market participants are not looking for the "fair" economic value of the asset. What they need is rather a price whose value at some given time summarizes in a suitable way the opinions of market participants at this time. This price is called *efficient price*. Hence, this paper aims at providing a statistical procedure in order to estimate this efficient price.

1.2. Ideas for the estimation strategy

In this paper, we focus on the case of large tick assets. We define them as assets for which the bid-ask spread is almost always equal to one tick. Our goal is then to infer an efficient price for this type of asset. Naturally, it is reasonable to assume that the efficient price essentially lies inside the bid-ask spread but we wish to say more.

In order to retrieve the efficient price, the classical approach is to consider the imbalance of the order book, that is the difference between the available volumes at the best bid and best ask levels, see for example [4]. Indeed, it is often said by market participants that "the price is where the volume is not". Here we consider a dynamic version of this idea through the information available in the *order flow*. More precisely, we assume that the intensity of arrival of the limit order flow at the best bid or the best ask level depends on the distance between the efficient price and the considered level: if this distance is large, the intensity should be high and conversely. Thus, we assume the intensity can be written as an increasing deterministic function of this distance. This function is called the *order flow response function*. In our approach, a crucial step is to estimate the response function in a non parametric way. Then, this functional estimator is used in order to retrieve the efficient price.

Note that it is also possible to use the buy or sell market order flow. In that case, the intensity of the flow should be high when the distance is small. Indeed, in this situation, market takers are not loosing too much money (with respect to the efficient price) when crossing the spread.

1.3. Organization of the paper

The paper is organized as follows. The model and the assumptions are described in Section 2. Particular properties of the efficient price are given in Section 3 and the main statistical procedure is explained in Section 4. The theorems about the response function can be found in Section 5 and the limiting behavior of the estimator of the efficient price is given in Section 6. One numerical illustration can be found in Section 7 and a conclusion is given in Section 8. Finally the proofs are relegated to Section 9.