Journal of Empirical Finance 19 (2012) 610–625

Contents lists available at SciVerse ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Modelling and forecasting liquidity supply using semiparametric factor dynamics ☆

Wolfgang Karl Härdle, Nikolaus Hautsch, Andrija Mihoci *

Humboldt Universität zu Berlin and C.A.S.E. – Center for Applied Statistics and Economics, Spandauer Str. 1, D-10178 Berlin, Germany
Center for Financial Studies (CFS), Frankfurt, Germany

ARTICLE INFO

Article history:
Received 21 April 2010
Received in revised form 21 March 2012
Accepted 2 April 2012
Available online 10 April 2012

JEL classification:
C14
C32
C53
G11

Keywords:
Limit order book
Liquidity risk
Semiparametric model
Factor structure
Prediction

ABSTRACT

We model the dynamics of ask and bid curves in a limit order book market using a dynamic semiparametric factor model. The shape of the curves is captured by a factor structure which is estimated nonparametrically. Corresponding factor loadings are modelled jointly with best bid and best ask quotes using a vector error correction specification. Applying the framework to four stocks traded at the Australian Stock Exchange (ASX) in 2002, we show that the suggested model captures the spatial and temporal dependencies of the limit order book. We find spill-over effects between both sides of the market and provide evidence for short-term quote predictability. Relating the shape of the curves to variables reflecting the current state of the market, we show that the recent liquidity demand has the strongest impact. In an extensive forecasting analysis we show that the model is successful in forecasting the liquidity supply over various time horizons during a trading day. Moreover, it is shown that the model's forecasting power can be used to improve optimal order execution strategies.

1. Introduction

Due to technological progress in the organization of trading systems and exchanges, electronic limit order book trading has become the dominant trading form for equities. Open limit order books provide important information on the current liquidity supply as reflected by the offered price-quantity relationships on both sides of the market. These supply and demand schedules provide valuable information on traders' price expectations in the spirit of the seminal paper by Glosten (1994), reflect the current implied costs of trading as well as demand and supply elasticities. However, while the dynamic behavior of liquidity demand, as reflected by trading intensities and trade sizes, has been already studied in various papers (see, e.g., (Hautsch and Huang (2012) and Brownlees et al. (2009)), the stochastic properties of liquidity supply is still widely unknown. An obvious reason is that liquidity supply is reflected by high-dimensional bid and ask schedules which are not straightforwardly modelled in a dynamic setting. Consequently, it is a widely open question whether and to which extent liquidity supply might be predictable.

☆ This work was supported by the Deutsche Forschungsgemeinschaft via Collaborative Research Center 649 "Ökonomisches Risiko", Humboldt-Universitt zu Berlin, Germany.

* Corresponding author. Tel.: +49 30 2093 5623.
E-mail address: mihociax@cms.hu-berlin.de (A. Mihoci).

0927-5398/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2012.04.002

The paper's major idea is to capture the shape of high-dimensional ask and bid curves by a lower-dimensional factor structure which is estimated non-parametrically. We propose a dynamic semiparametric factor model where the shape of order schedules is captured by a non-parametric factor structure while the curves' dynamic behavior is driven by time-varying factor loadings. The latter are modelled parametrically employing a vector error correction model (VECM). We show that the model captures the dynamics of high-dimensional order curves very well and is sufficiently parsimonious to produce valuable out-of-sample predictions. Moreover, the schedule of market depth posted around best quotes reveals strong serial dependence and thus is predictable. This structure is induced by the inventory character of order volume which is strongly persistent over time.

By providing empirical evidence on the dynamics and predictability of order book schedules, this paper fills a gap in empirical literature and complements recent (mostly theoretical) work on order splitting and dynamic order submission strategies. For instance, the question of how to reduce the costs of trading by optimally splitting a large order over time (e.g., over the course of a trading day) is of high relevance in financial practice. Obizhaeva and Wang (2005) and Engle and Ferstenberg (2007) analyze optimal splitting strategies whose implementations ultimately require predictions of future liquidity demand and supply. Bertsimas and Lo (1998) and Almgren and Chriss (2000) derive optimal execution strategies by minimizing expected costs of executing, an order in the context of static price impact functions. Optimal execution in a limit order book market is analyzed by Alfonsi et al. (2010). They allow for general shapes of order book curves and derive explicit optimal execution strategies in discrete time. By providing insights into the actual form of order book curves and their dynamic behavior, our results can be used as valuable inputs in theoretical frameworks.

While to the best of our knowledge our study is the first which models the shapes and dynamics of a complete (high-dimensional) order book, there is a substantial body of empirical literature on the dynamics of limit order books and the analysis of traders' order submission strategies, such as, e.g., Biais et al. (1995), Griffiths et al. (2000), Ahn et al. (2001), Ranaldo (2004), Hollifield et al. (2004), Bloomfield et al. (2005), Degryse et al. (2005), Hall and Hautsch (2006, 2007), Large (2007), Hasbrouck and Saar (2009) or Cao et al. (2009).

An important aspect in this literature is to analyze the question of how to optimally balance risks and gains of a trader's decision whether to post a market order or a limit order. As recently illustrated by Chacko et al. (2008), a limit order can be ultimately seen as an American option and transaction costs are rents that a monopolistic market maker extracts from impatient investors who trade via aggressive limit orders or market orders. Consequently, the analysis of liquidity risks (see, e.g., Johnson, (2008), Liu (2009), Garvey and Wu (2009), Goyenko et al. (2009)) and transaction costs (see, e.g. Chacko et al. (2008), Hasbrouck (2009)) are in the central focus of recent literature.

Given the objective to capture not only the volume around the best quotes but also pending quantities more deeply in the book, the underlying problem becomes inherently high-dimensional. A typical graphical snapshot of ask and bid curves for four stocks traded at the Australian Securities Exchange (ASX) in 2002, is given by Fig. 1. The curse of dimensionality applies immediately as soon as time variations of the order curve shapes have to be taken into account. As shown by Fig. 1 and as illustrated in more detail in the sequel of the paper, order volume is not necessarily only concentrated around the best quotes but can be substantially dispersed over a wider range of price levels. This is a typical scenario for moderately liquid markets as that of the ASX. In such a context, the dynamic modelling of all volume levels individually becomes complicate and intractable.

We suggest reducing the high dimensionality of the order book by means of a factor decomposition using the so-called Dynamic Semiparametric Factor Model (DSFM) proposed by Fengler et al. (2007), Brüggemann et al. (2008), Park et al. (2009) and Cao et al. (2009). Accordingly, we model the shape of the book in terms of underlying latent factors which are defined on a grid space around the best ask or bid quotes and can depend on additional explanatory variables capturing, e.g., the state of the market. In order to avoid specific functional forms for the shape of the curves, the factors as well as the corresponding loadings are estimated nonparametrically using B-splines. Then, in a second step, we model the multivariate dynamics of the factor loadings together with the best bid and the best ask price using a VEC specification.

Using this framework we aim answering the following research questions: (i) How many factors are required to model order book curves reasonably well? (ii) What does the shape of the factors look like? (iii) What do the dynamics of the estimated factor loadings look like? (iv) Does there exist evidence for a strong cross-dependence between both sides of the order book? (v) Can quotes be predictable in the short run? (vi) Does the shape of the order book curves depend on past price movements, past trading volume as well past volatility? (vii) How successful is the model in predicting future liquidity supply and can it be used to improve order execution strategies?

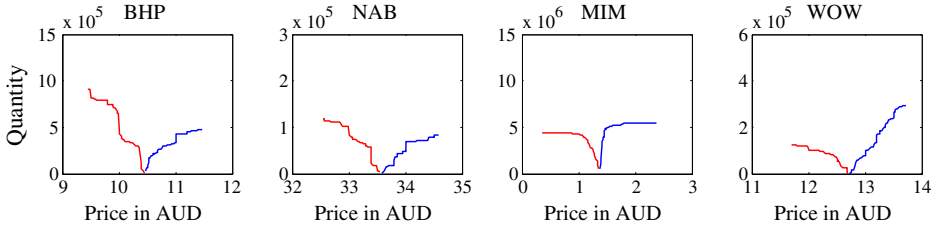

Fig. 1. Limit order books for selected stocks traded at the ASX on July 8, 2002 at 10:15. Red: bid curve, blue: ask curve.