

Modelling and Forecasting Liquidity Supply Using Semiparametric Factor Dynamics*

Wolfgang Karl Härdle[†] and Nikolaus Hautsch[‡] and Andrija Mihoci[§]

Abstract

The limit order book is the source of information regarding liquidity supply and hence for quantifying liquidity risk. In modelling and forecasting of the limit order book we use a Dynamic Semiparametric Factor Model. Using high frequency data from the Australian Stock Exchange we are able to provide limit order book estimates with interpretable factors and parametric models for the loading series reflecting the dynamic behaviour. Moreover, we obtain forecasting results and thereby enable risk management applications such as optimal splitting strategies for large transactions. This work presents the first attempt to model the entire limit order book in a directly data driven fashion.

JEL classification: C14, C32, C53, G11

Keywords: limit order book, liquidity risk, semiparametric models, factor models

*We are very grateful to Anthony Hall for providing us with data. Moreover, financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Ökonomisches Risiko", Humboldt-Universität zu Berlin is gratefully acknowledged. Furthermore, we are grateful to Szymon Borak for helping us with the implementation of the Dynamic Semiparametric Factor Model (DSFM) in MATLAB.

[†]Humboldt-Universität zu Berlin, CASE - Center for Applied Statistics and Economics, Spandauer Str. 1, 10178 Berlin, Germany

[‡]School of Business and Economics as well as CASE - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Quantitative Products Laboratory (QPL), Berlin, and Center for Financial Studies (CFS), Frankfurt. Address: Spandauer Str. 1, 10178 Berlin, Germany, tel: +49 (0)30 2093 5711, email: nikolaus.hautsch@wiwi.hu-berlin.de

[§]Humboldt-Universität zu Berlin, CASE - Center for Applied Statistics and Economics, Spandauer Str. 1, 10178 Berlin, Germany

On financial markets one observes that the classical ways of trading (floor, over-the-counter) are replaced by electronic trading and that the importance of the limit order book markets emerges. Our research questions include:

- (i) modelling and forecasting the demand and supply curves for shares under a flexible statistical framework, and
- (ii) analysis of the dynamics and predictability of the equity value, i.e. the value for selling and purchasing shares.

Our paper is studying the behaviour of the limit order book, and consequently of the liquidity supply. In particular, the important factors and variables that have significant impact on the movements of the limit order book over time are accessed. Our contribution lies in a semiparametric modelling and forecasting of the limit order book.

In economics it is a challenging task to model and forecast demand and supply curves. In our paper, demand and supply curves for stocks located above the market's equilibrium quantity are analysed. A typical snapshot of a limit order book for two stocks, traded at the Australian Stock Exchange (ASX) in 2002, is given in Figure 1.

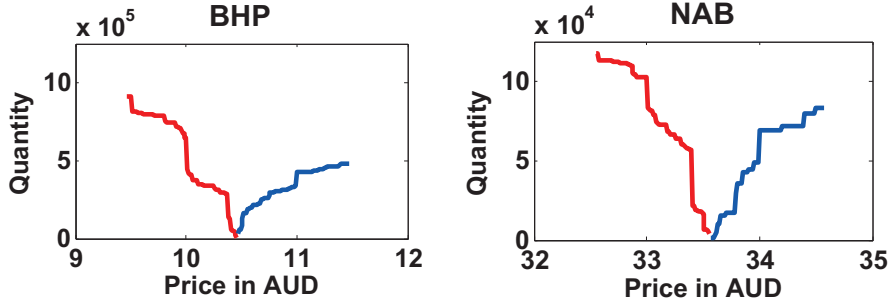


Figure 1: Limit order book for Broken Hill Proprietary Ltd. (BHP) and National Australia Bank Ltd. (NAB) at the ASX on 8. July 2002 at 10:15. The red line shows the demand (bid) curve, and the blue line shows the supply (ask) curve.

Similarly, in statistics it is challenging to model the limit order book, the object of our interest, because it is high-dimensional and it changes rapidly over time. Our objective is to provide a flexible framework for modelling and forecasting the limit order book. By using the Dynamic Semiparametric Factor Model (DSFM), stipulated under the philosophy '*smooth in space and parametric in time*', we can achieve our goal. The estimated factors and the time series of corresponding factor loadings are used for predicting skewed demand and supply curves over time, and thereby accomplishing our goal.

The dynamics and the predictability of equity value play an essential role in finance and practice. Since we are modelling and forecasting the volume for shares demanded

and supplied by market participants at certain price levels through time, we are able to analyse the dynamics of the equity value. Different equity values for two stocks traded at the Australian Stock Exchange (ASX), at two different time points, are illustrated in Figure 2. By including liquidity demand and realized volatility as additional explanatory variables, it is possible to predict the equity value through time. In this sense, it is possible to help, among others, large institutional investors to rebalance their portfolio values, banks to adjust their capital adequacy ratio and companies to model their capital structure. In general, investors can improve their trading strategies.

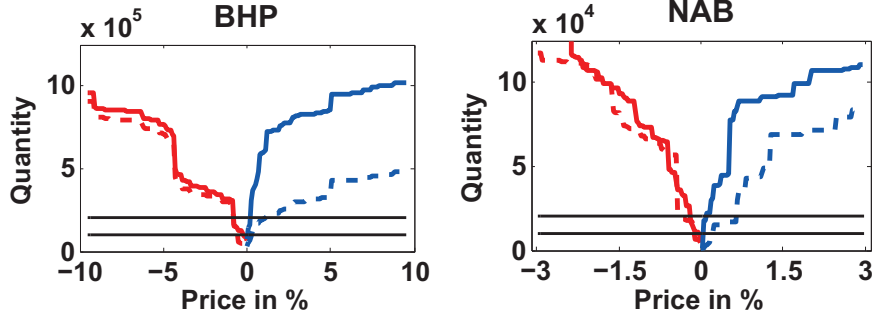


Figure 2: Limit order book for BHP and NAB at the ASX, 8. July 2002, 10:15 (dashed) and 15:44 (solid) indicating different equity values for selling/buying given number of stocks denoted by two horizontal lines

The datasets we are working with are in their nature high-dimensional and they are high-frequently observed. For modelling and forecasting liquidity supply, the following two stocks traded at the Australian Stock Exchange (ASX) in 2002 were selected:

- (i) Broken Hill Proprietary Ltd. (BHP), and
- (ii) National Australia Bank Ltd. (NAB).

The seasonal factor for a given trading day for both sides of the market is plotted in Figure 3.

The DSFM is an orthogonal L -factor model of an observable J -dimensional random vector:

$$Y_{t,j} = m_{0,j} + Z_{t,1}m_{1,j} + \cdots Z_{t,L}m_{L,j} + \varepsilon_{t,j}, \quad (1)$$

where $m(\cdot) = (m_0, m_1, \dots, m_L)^\top$ denotes time-invariant factors, a tuple of functions with the property $m_l : R^d \rightarrow R$, and $Z_t = (1, Z_{t,1}, \dots, Z_{t,L})^\top$ denotes the time series of factor loadings. The time index is denoted by $t = 1, \dots, T$, and the cross-sectional index

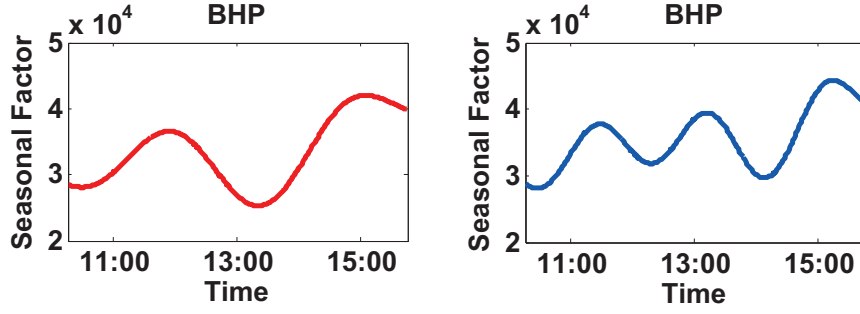


Figure 3: Intraday seasonal factor for the quantities offered at the best bid prices (left panel, $M = 2$), as well as for the quantities supplied at the best ask prices (right panel, $M = 3$) for BHP at any trading day

by $j = 1, \dots, J$. For estimation purposes, the number of underlying factors L should not exceed the dimension of the object, J . This inequality is preserved in our case.

It is notable from Figure 4 that the first estimated factor captures the slope of the limit order book, whereas the second estimated factor is accessing the slope of the curve, as expected. This fact is supported by the size of both factor loadings time series.

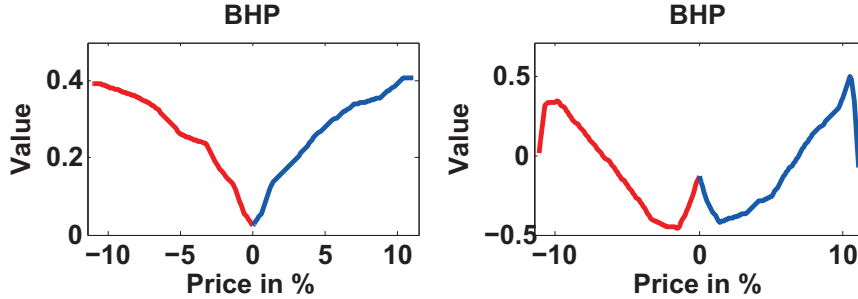


Figure 4: Factors for both sides of the limit order book for BHP from 8. July 2002 to 16. August 2002 using the DSFM-Separated approach: 1st factors (left panel) and 2nd factors (right panel). The relative prices are observed on 16. August 2002 at 15:44.

Based on further analysis, we conclude that the Dynamic Semiparametric Factor Model is suitable for modelling and forecasting the limit order book by reducing its dimensionality, while providing a compromise between the spatial and the time structure.