Modelling the Dynamics of Cross-Sectional Price Functions: an Econometric Analysis of the Bid and Ask Curves of an Automated Exchange

Clive G. Bowsher*
Nuffield College, University of Oxford, Oxford, OX1 1NF, U.K.
clive.bowsher@nuffield.ox.ac.uk

September 22, 2004

Abstract

Functional Signal plus Noise (FSN) time series models are introduced for the econometric analysis of the dynamics of a large cross-section of prices in which contemporaneous observations are functionally related. A semiparametric FSN model is developed in which a smooth, cubic spline signal function is used to approximate the price curve data. Estimation may then be performed using quasi-maximum likelihood methods based on the Kalman filter.

The model is used to provide one of the first studies of the dynamics of the bid and ask curves of an electronic limit order book and enables the comprehensive measurement of the dynamic determinants of traders' execution costs. It is found that the differences between the bid and ask curves and their intercepts (i.e. the immediate price impacts of market orders) are well described by covariance stationary processes. The in-sample, 1-step ahead point predictions for these curves perform well and motivate the development of parametric FSN models that take into account the monotonicity of the price curves and can be used to form predictive distributions.

Keywords: functional time series, bid and ask curves, liquidity, electronic limit order book, cubic spline, state space form, Kalman filter, quasi-maximum likelihood.

JEL classification: C14, C32, C33, C51, G12.

1 Introduction

The quoted prices available on a financial market at any given time can often usefully be viewed as a function of some feature of the asset traded such as the volume of the transaction or the maturity date of the asset in question. Prominent examples include the bid and ask curves of an automated exchange and the yield curve for pure discount bonds. A record of such asset prices is thus a panel dataset of high cross-sectional dimension, in which contemporaneous observations are related to one another in a functional manner. This paper develops a novel econometric framework for the coherent, joint modelling of both the cross-sectional and dynamic aspects of such asset price functions.

*Comments are welcome and should be directed to the above e-mail address.

We apply our approach in a study of the dynamic evolution of the bid and ask curves of a fully automated financial market or 'electronic limit order book.' Electronic limit order books have become an important means of trading equities, derivatives and bonds worldwide and their increasing use seems set to continue. For example, the London Stock Exchange's electronic order book SETS accounted on average for 59% of the daily total value traded on the exchange in 2001 and for 62% during the first nine months of 2002 (London Stock Exchange plc (2003b)). Analysis of the dynamics of the bid and ask curves of the limit order book is of crucial importance for the control of traders' execution costs and for an understanding of the dynamics of liquidity.

The main contributions of the paper may be summarised as follows. First, *Functional Signal plus Noise (FSN)* time series models are introduced for analysing the dynamics of a large cross-section of prices in which contemporaneous observations are functionally related. Applied to the average price (i.e. bid and ask) curves of an electronic limit order book market, these enable the comprehensive measurement of the determinants of traders' execution costs (i.e. 'liquidity') in a way that takes the dynamic aspects of the problem into account. The FSN models specify the evolution over time of stochastic functions, a problem that has received relatively little attention in the econometrics and statistics literature. Second, a semiparametric FSN model is developed in which a smooth, cubic spline signal function is used to approximate the price curve data. The use of cubic splines results in models that may be estimated using quasi-maximum likelihood (QML) methods based on the Kalman filter. Third, this FSN model is then used to provide one of the first studies of the dynamics of the bid and ask curves of a limit order book (LOB). It is found that the average price curves that give the immediate price impact of market orders are well described by covariance stationary processes which lack deterministic time-of-day effects. The in-sample, one-step ahead point predictions for these curves perform well. Finally, we consider the specification of fully parametric FSN models that take into account the monotonicity property of the bid and ask curves and that can potentially be used to form predictive distributions (or 'density forecasts') for the econometric solution of trading decision problems.

The bid and ask curves of the LOB at time t have a straightforward definition: they express the *average price* per unit volume of instantaneously selling and buying the security respectively (using a so-called 'market order'), as a function of the *volume*, v, of the security traded. They are denoted here by $p_t^b(v)$ and $p_t^a(v)$ respectively. To understand the need for a dynamic model of these curves in order to provide econometric solutions to trading decision problems, consider an agency trader acting on behalf of a client who is instructed to buy or sell a large volume of shares in a particular security by the end of the current trading day. Assuming that the trader restricts attention to market orders and wishes to buy, she must decide how to split up the volume over the course of the day in order to minimise her expected loss. The current ask curve at time t determines the cost of purchasing various volumes at time t, and is available from the electronic trading screen that the trader views. Note that the ask curve thus provides the *immediate price impact* of a market order to buy volume v, in excess of the best ask – i.e. the price difference $p_t^a(v) - p_t^a(1)$. This is an important aspect of most definitions of liquidity, sometimes known as 'depth' (see Kyle (1985)). In solving her decision problem, the trader will also be concerned with dynamic issues such as the degree of persistence of the slope of the ask curve ("Over what time horizon can low price impact now be expected to imply low price impact in the future?") and the effect of her market orders on the future evolution of the curve ("To what extent does the consumption of liquidity now affect future price impact and the future level of the best ask?"). Such questions form the motivation for the time series models of bid and ask curves developed here.

The work presented is also a contribution to applied *functional time series* analysis. Each average price curve may be regarded as a finite-dimensional vector, albeit of very high dimension. However, the standard approaches of multivariate time series or panel data econometrics are of little help in this setting, owing to the high dimensionality of the curves and the close, functional relationship between the observations lying on those curves. Indeed, the analysis of time series of stochastic functions such as $\{p_t^a(v)\}_t$ is in a state of relative infancy.[1] Previous work in the statistics literature includes the use of Functional AutoRegressive (FAR) models for forecasting entire smooth, continuous functions by Besse and Cardot (1996) and Besse, Cardot, and Stephenson (2000). We are aware of only two previous functional time series studies in the econometrics literature. In concurrent, independent work, Diebold and Li (2003) introduce a dynamic version of the Nelson and Siegel (1987) yield curve in which the three parameters or 'factors' describing the curve follow an autoregressive process. Their approach thus has features

[1] *Functional Data Analysis* by Ramsay and Silverman (1997) is a landmark in this area but does not consider the time series case in which the functions are dependent.

in common with our own, although we view the use of cubic spline signal functions as offering considerable flexibility in a broad range of empirical applications. Gouriéroux, Le Fol, and Meyer (1998) also analyse the dynamics of the bid and ask curves of a financial market, but in transaction time. A comparison with their two factor model is provided towards the end of the paper.

In the Functional Signal plus Noise (FSN) models developed here, the information about the functional, cross-sectional relationship between contemporaneous observations is captured by modelling the observed curves as the sum of a smooth signal function, $S(v)$, and a noise process. The signal function used is a cubic spline uniquely determined by its *knot-prices* – i.e. the prices that correspond to the volumes at which the knots of the spline are positioned. The state equation of the FSN model determines the stochastic evolution of the low-dimensional vector of knot-prices and hence of the entire spline signal. The choice of a cubic spline for the signal function is advantageous since the FSN model can then be written in linear state space form, thus allowing the use of the Kalman filter to perform both QML estimation and 1-step ahead linear point prediction. Harvey and Koopman (1993) were the first to describe a linear state space model with a state equation determining the stochastic evolution of a cubic spline function. There, as in Koopman and Ooms (2001), the stochastic spline is used to model the latent, time-varying periodic pattern of a scalar time series. This contrasts the present work in which the stochastic splines are used as a tool in functional time series analysis and assume the role of smooth approximations to the observed functional data.

The structure of the paper is as follows. Section 2 explains the basic features of an electronic LOB, defines its bid and ask curves in terms of the limit orders that make up the order book, and provides a description and graphical display of the SETS bid and ask curve dataset analysed later. Section 3 introduces the semiparametric FSN model and shows how it may be estimated by QML. Section 4 then presents the results of fitting this FSN model to the SETS dataset. An analysis of the one-step ahead prediction errors of the model motivates the development of fully parametric FSN models taking into account the monotonicity of the bid and ask curves in section 5. A comparison with previous research on the dynamics of liquidity and average price curves, and a discussion of possible extensions to this work are given in section 6, before concluding in section 7.

The Appendix provides the necessary mathematical details on cubic spline functions (see section 8.1). These are denoted here, as a function of v, by $S(v)$. The following additional notation is used throughout the paper. For a vector or matrix β, we write $\beta > 0$ (resp. $\beta \geq 0$) if and only if every element of β is positive (resp. non-negative). The Hadamard or 'element-by-element' product of two vectors x and y, both of dimension m, is written $x \odot y$, where $x \odot y := (x_1 y_1, ..., x_m y_m)'$. Finally, $Z \sim (\mu_z, \Omega_z)$ is used to mean that the random vector Z has finite mean μ_z and finite covariance matrix Ω_z.

2 Limit Order Book Data and Bid and Ask Curves

The following section briefly explains the roles of market and limit orders in the operation of an electronic LOB and provides a definition of its bid and ask curves in terms of the limit orders present on the book. A description and graphical display of the London Stock Exchange SETS (Stock-exchange Electronic Trading Service) limit order book dataset analysed later in section 4 is then provided.

2.1 Institutional background and general definitions

As was noted above, the bid (ask) curve of the limit order book at time t, $p_t(v)$, is a function of volume that gives the average price per unit volume of instantaneously selling (buying) the security using a market order. When a trader submits, say, a buy market order to the electronic system, it is matched against and results in the removal of an equal limit order volume from the sell side of the book. The sequence in which limit orders are matched against incoming market orders is determined by price, with higher (lower) priced limit buy (sell) orders having priority. As the name suggests, limit orders reside on the LOB awaiting execution. The operation of such 'order driven' markets is described in detail by Gouriéroux and Jasiak (2001, Ch.14) and Bauwens and Giot (2001).[2]

A *buy market order* with volume v is taken here to refer simply to an order to buy v shares that results in immediate execution at the best price available. A *buy limit order* with volume v and price P represents a firm offer to buy v shares at a price (less than or) equal to P. Sell orders are defined analogously. At any point in time the 'bid' or buy side of the LOB can be

[2]Further details concerning the SETS trading system in particular are given in London Stock Exchange plc (2003a).

viewed as the collection of all buy limit orders present on the book, grouped according to their price. Similarly, the 'ask' or sell side consists of the collection of all sell limit orders. The *best bid* is defined as the maximum price associated with any buy limit order, and the *best ask* as the minimum price associated with any sell limit order. These prices represent the best priced offers on the book and are denoted here by p_t^{b*} and p_t^{a*} respectively. Defining $m_t := [p_t^{a*} + p_t^{b*}]/2$ and $s_t := p_t^{a*} - p_t^{b*}$, m_t and s_t are referred to as the *mid-quote* and *bid-ask spread* respectively.

Let the 'tick size' be the minimum price difference between any two limit orders having different prices.[3] Then, normalising the tick size to be one and defining V_{it}^b $(i = 0, 1, ...)$ as the total volume of buy limit orders at time t whose price is exactly i ticks away from the best bid (i.e. whose price is $p_t^{b*} + i$), the bid side of the LOB can be written simply as $\{V_{it}^b; i = 0, 1, ...\}$. Similarly the ask side of the book is given by the collection of volumes $\{V_{it}^a; i = 0, 1, ...\}$. Now consider what happens when a sell market order with volume $v = (V_{0t}^b + V_{1t}^b)$, say, is submitted at t. The market order will execute against all of the buy limit orders present at or one tick away from the best bid, resulting in a trade with an average price of $[V_{0t}^b p_t^{b*} + V_{1t}^b(p_t^{b*} + 1)]v^{-1}$.

This is the basic idea behind the formal definition of the average price curves given below, the only complication being that market orders with a volume v that only partially exhausts the total volume of limit orders present at some price must also be considered.

Definition 1 ***Average Price Curves.*** *The* ***bid*** *(resp.* ***ask****)* ***curve*** *at time t expresses the average price per unit volume that would result from instantaneously selling (resp. buying) v shares by submitting a market order at t.*[4] *Let V_{it} be the total volume of buy (resp. sell) limit orders at time t whose price is exactly i ticks away from the best bid (resp. ask), $i = 0, 1, 2, ...$, and let p_t^* be the best bid (resp. ask). Then the average price curve is given as a function of v by*

$$p_t(v) = \frac{1}{v}\left[\left(v - \sum_{i=0}^{s} V_{it}\right)(p_t^* + s + 1) + \sum_{i=0}^{s}(p_t^* + i)V_{it}\right], \tag{1}$$

where $s := \sup\{k : \sum_{i=0}^{k} V_{it} \leq v, \quad k = 0, 1, 2, ...\}$ provided this set is not empty, $s := -1$ otherwise and empty sums are equal to zero. When the need arises to distinguish the bid and ask curves explicitly, the notation $p_t^b(v)$ and $p_t^a(v)$ is used respectively.

[3] For the dataset analysed in section 4 the tick size is equal to one penny.

[4] It is assumed that this price is known given observation of the LOB at time t, thus abstracting from the possibility of time lags between submission and execution of a market order.

Equation (1) can readily be deduced by noting that a market order with volume v fully exhausts the total volume of all limit orders whose price is within s ticks of the best priced limit order, and then is matched against some but not all of the limit order volume present at $(s+1)$ ticks from the best price. It follows from (1) that the bid (ask) curve is non-increasing (non-decreasing).[5]

To avoid confusion, it is noted that the domain of the average price function or curve at time t is the finite set $\{1, 2, ..., v_{n(t)}\}$, or for practical purposes some subset of that set, where $v_{n(t)}$ is the aggregate volume of limit orders present on the relevant side of the book at time t. The on-screen information viewed by the trader includes both the bid and ask curves with domain $\{1, 2, ..., v_{n(t)}\}$. The average price curves may thus always be represented by finite dimensional vectors, which is the practice adopted here. Furthermore, the domain of the price curve will henceforth be written as the vector $\nu = (v_1, v_2, ..., v_n)'$, where $v_{n(t)}$ is now assumed not to depend on t for the sake of simplicity, and the price curve itself will be denoted for short by $p_t(\nu)$.

2.2 SETS data

The dataset analysed in section 4 was extracted from the London Stock Exchange's Transaction Data Service (TDS) database and covers the period from 2 January 2002 to 31 January 2002 inclusive for the heavily-traded FTSE100 stock AstraZeneca. The TDS database allows the reconstruction of the SETS limit order book and its associated average price curves at any given time (for further details of which see section 8.3 of the Appendix). The dataset consists for each trading day of the bid and ask curves observed at 1 minute intervals during the 8.25 hour period from 08:15 GMT to 16:30 GMT inclusive, giving 496 pairs of curves per day. This 8.25 hour period includes almost all of the period of continuous trading on SETS each day.[6]

A graphical display of the ask and bid curves at 5 minute intervals for the entirety of one randomly selected day of the dataset, namely 30 January 2002, is given in Figure 1. Time is shown on the horizontal axis, and increases from right to left. Particularly noticeable are the high degree of variability of the average price that is observed right across the volume range and the lack of any obvious relationship between the bid and ask sides. As in the Paris CAC

[5] Considering v as a continuous variable, equation (1) also implies that $p_t(v)$ is continuous, not differentiable everywhere and for a bid (ask) curve piecewise convex (concave).

[6] The first 15 minutes of the standard period of continuous trading are omitted in order to guarantee that any extensions to the opening auction have been completed and continuous trading has begun.

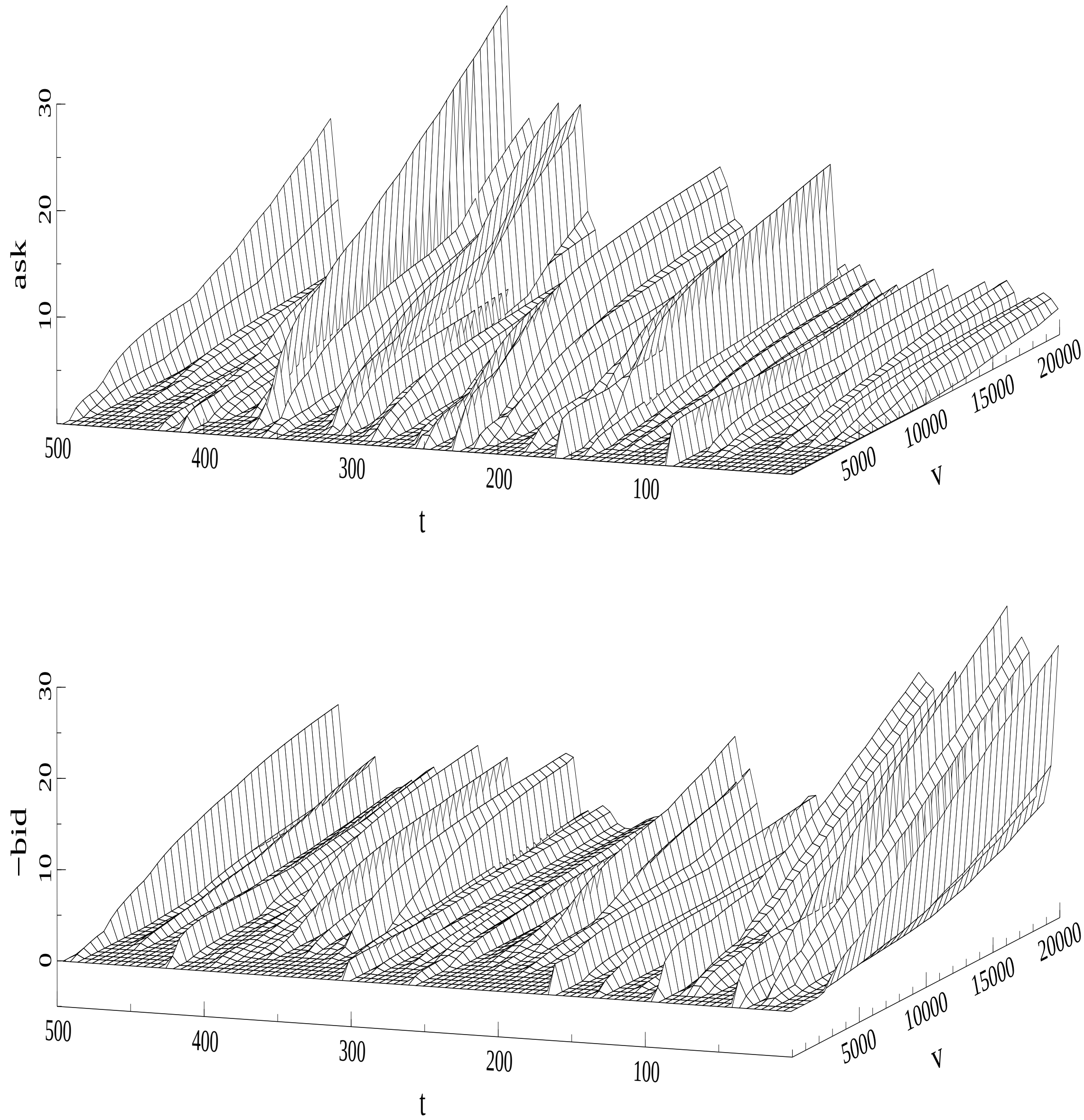

Figure 1: The ask and bid curves plotted at 5 minute intervals for AstraZeneca on 30 January 2002. The curves are in pence relative to the best bid and ask respectively; minus the bid curve is shown for visual clarity; time, t, is in minutes after 08:15 GMT and volume, v, is in number of shares.

(Cotation Assistée en Continu) dataset analysed by Gouriéroux, Le Fol, and Meyer (1998), the bid and ask curves are far from being always symmetric about the volume axis.

3 Functional Signal plus Noise Models

We now present the new Functional Signal plus Noise (FSN) model for analysing time series of cross-sectional price functions. The exposition is given for the bid and ask curves of a LOB. With minor modification, however, the FSN framework is potentially broadly applicable to the empirical analysis of time series of price functions due to the flexibility of cubic splines as approximating functions. In this section a semiparametric FSN model is considered in which a smooth, cubic spline signal function is used to approximate the price curve data. The parameters of the model determine completely the mean and autocovariance structure of the time series of price functions and so this type of model can be estimated by QML. It is shown below how to write the FSN model in linear state space form, thus allowing the use of the Kalman filter to compute both the Gaussian quasi-likelihood function and one-step ahead, linear point predictions.[7] A discussion of the specification of fully parametric FSN models is deferred until section 5.

3.1 Semiparametric FSN model

FSN models are presented here for the *difference* between the bid (ask) curve as given by Definition 1 and the best bid (ask) price, that is for $\{p_t(v) - p_t^*\}_t$. Furthermore, the models are for a time series of *either* bid *or* ask curves.[8] The extension to joint models incorporating the best quoted prices and dependence between the bid and ask curves is discussed in section 6.2. Henceforth, except where indicated otherwise, the notation $p_t(\nu)$ is thus used for the difference between the bid (ask) curve as given by Definition 1 and the best bid (ask), that is $p_t(\nu) := (p_t(v_1) - p_t^*, ..., p_t(v_n) - p_t^*)'$, where $\nu = (v_1, ..., v_n)'$ is the vector of volumes under consideration. Observation of the curves is assumed to take place at $t = 1, 2, ..., T$.[9]

[7] The term *Kalman filter* should always be taken here to refer to the recursions as they are conveniently stated in Koopman, Shephard, and Doornik (1999, Section 4.3, pp. 122-123). For a textbook exposition of the Kalman filtering procedure, see Harvey (1989, Ch. 3).

[8] Note that the primary object of interest from the perspective of the agency trading decision problem discussed in section 1 is either the bid or ask curve, depending on whether the trader is instructed to sell or buy the security. Agency traders instructed to sell, say, are not permitted by clients to buy the security as part of an overall trading strategy.

[9] In the empirical analysis of section 4 the price curves corresponding to the dth day, $(p_1^{(d)}(\nu)', ..., p_T^{(d)}(\nu)')'$ are modelled as being independent and identically distributed across days.

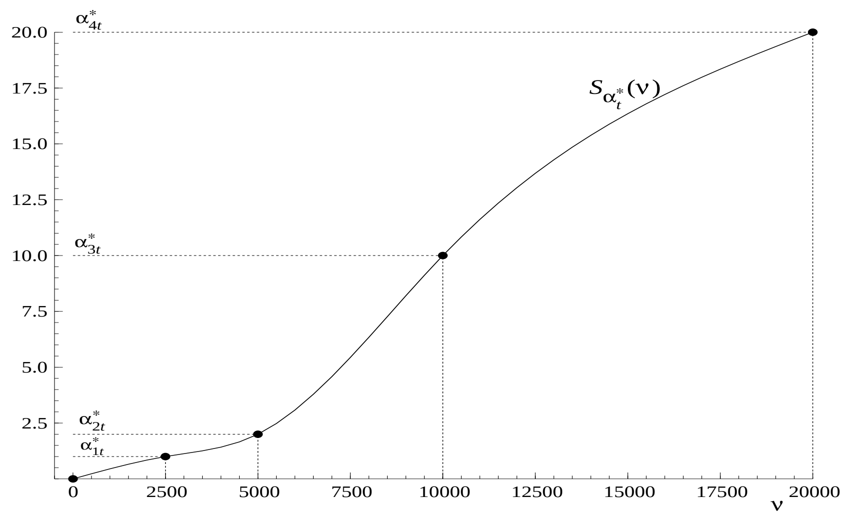

Figure 2: An illustrative spline signal function, $S_{\alpha_t^*}(\nu)$, for an ask curve. The prices at the knots, α_{jt}^*, are plotted using filled circles; the knots are given by $k = (1, 2500, 5000, 10000, 20000)'$; the states, α_{jt} $(j = 2, 3, 4)$, are the vertical distances between the horizontal dashed lines and α_{1t} is the vertical distance between the abscissa and the first horizontal dashed line.

The semiparametric FSN model consists of a stochastic, smooth, cubic spline signal function, denoted here by $S_{\alpha_t^*}(\nu) = (S_{\alpha_t^*}(v_1), ..., S_{\alpha_t^*}(v_n))'$, plus a noise process. A cubic spline is essentially a piecewise cubic function with pieces that join together to form a smooth function overall. The essential aspects of cubic spline theory and related definitions needed here are described in the Appendix (see section 8.1). The spline signal $S_{\alpha_t^*}(\nu)$ has knots positioned at the volumes $k = (1, k_1, ..., k_m)'$ – these are deterministic and fixed over time. The notation $S_{\alpha_t^*}(\nu)$ is used to imply that the spline interpolates to $\alpha_t^* = (0, \alpha_{1t}^*, ..., \alpha_{mt}^*)'$ – i.e. $S_{\alpha_t^*}(k_j) = \alpha_{jt}^*$ for $j = 0, ..., m$. The α_{jt}^* can thus be thought of as the 'latent prices at the knots' or more concisely as the *knot-prices* of the spline.[10] The state equation of the FSN model then determines the stochastic evolution of these knot prices over time. An illustrative spline signal function for an ask curve at time t is shown in Figure 2.

A formal definition of the semiparametric FSN model now follows.

[10]Note that the spline signal function is constrained here to pass through the point $(1, 0)$ for all t. Fixed intercepts other than zero create no difficulty. The extension to the case where the spline function has a stochastic intercept is also straightforward.

Definition 2 ***Semiparametric Functional Signal plus Noise (FSN) Model.*** *The model for the observable n-vectors, $p_1(\nu), ..., p_T(\nu)$, is given by*

$$\begin{aligned} p_t(\nu) &= S_{\alpha_t^*}(\nu) + \epsilon_t \\ &= Z\alpha_t + \epsilon_t, \qquad (2) \\ \alpha_{t+1} &= \delta + \Phi\alpha_t + \eta_t, \qquad (3) \end{aligned}$$

for $t = 1, ..., T$, where $S_{\alpha_t^*}(\nu)$ is a natural cubic spline on $(k; \alpha_t^*)$ and Z is determined uniquely by $Z\alpha_t = W\alpha_t^*$, where the $n \times (m+1)$ matrix W is given by Theorem 2 of the Appendix. For all t, the state vector $\alpha_t = (\alpha_{jt}^* - \alpha_{j-1,t}^*)_{j=1}^m$, with $\alpha_{0t}^* = 0$. The initial condition $\alpha_1 \sim (a_1, \Omega_1)$ applies. The series $\{u_t := (\epsilon_t', \eta_t')'\}$ satisfies

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega_\epsilon & 0 \\ 0 & \Omega_\eta \end{pmatrix} \right\} \quad \forall t, \qquad (4)$$

$\mathsf{E}[u_t u_s'] = 0 \ \forall s \neq t$, and $\mathsf{E}[u_t \alpha_1'] = 0 \ \forall t$. Note that $\{u_t\}$ is a vector white noise process.

The ***Gaussian FSN Model*** *is the semiparametric FSN model with the additional condition imposed that both u_t and α_1 have multivariate Normal distributions.*

The state vector of the FSN model is given by $\alpha_t = (\alpha_{jt}^* - \alpha_{j-1,t}^*)_{j=1}^m$ for $j = 1, ..., m$. The jth state is thus the *price differential* between the jth and $(j-1)$th knot-prices. It is important to note from Theorem 2 of the Appendix that the spline signal $S_{\alpha_t^*}(\nu)$ can be written as a linear function $W\alpha_t^*$ of the vector of knot-prices, α_t^*, and hence as a linear function $Z\alpha_t$ of the state vector (where Z follows directly from the fact that $\alpha_{jt} = \alpha_{jt}^* - \alpha_{j-1,t}^*$). Importantly, the deterministic matrices W and Z depend only on the vector ν and the knots k, and do not depend on the price differentials α_t. The consequence is that the FSN model has the linear state space form, as described by Harvey (1989, pp. 100-104).

The choice of a cubic spline as the signal function in the FSN model thus has two advantageous features. First, the stochastic evolution over time of the signal function is determined completely by the time series properties of the m-dimensional state vector α_t, where m is relatively small (e.g. $m = 4$), thus allowing the construction of a parsimonious model even when the dimension of $p_t(\nu)$ is very much larger. Second, since the model has the linear state space form, the Kalman filter may be used to perform both QMLE and 1-step ahead, linear point prediction. These two features, together with the flexibility of cubic splines as approximating functions, make the cubic spline framework adopted here a particularly attractive one.

It should be noted that the dynamics of α_t^* implied by the state equation of the FSN model differ greatly from the corresponding specification in Harvey and Koopman (1993). There the 'knot-ordinates', α_t^*, follow a multivariate random walk whereas in (3) their dynamics are determined by a first order vector autoregression (VAR) in the spline 'differentials', α_t. It is found in the empirical analysis of bid and ask curves in section 4 that this focus on the differentials between the knot prices works well in practice. More generally, finite order VARs in the knot-prices or spline differentials can be accommodated within the linear state space FSN framework.

3.2 QMLE and prediction using the Kalman filter

The FSN model is a semiparametric model whose parameters determine completely the mean and autocovariance structure of the time series of price curves. Dunsmuir (1979) has shown how to estimate such vector linear time series models by QML, the quasi-likelihood function being derived under the assumption that the time series has a multivariate Gaussian distribution. Such a quasi-likelihood function is obtained here by computing the likelihood of the Gaussian FSN model of Definition 2 using the Kalman filter. Under the Gaussian FSN model, the price curves $\{p_1(\nu), ..., p_T(\nu)\}$ have a nT-variate Gaussian density (likelihood) which may be evaluated for a particular set of observed curves using the prediction error decomposition and $f(p_t(\nu)|p_{t-1}(\nu), ..., p_1(\nu))$ from the Kalman filter for $t = 1, ..., T$.

Dunsmuir (1979, pp.499-503) considers the QMLE of models for a vector signal observed with noise of the form

$$p_t = s_t + \epsilon_t, \tag{5}$$

where both the signal s_t and noise ϵ_t are strictly stationary multivariate ARMA processes, and the signal and noise are 'incoherent', i.e. $\mathsf{E}[s_t \epsilon_r'] = 0 \ \forall r, t$. He establishes the strong consistency and asymptotic normality of the QMLEs for such a model obtained by maximising frequency domain approximations to the Gaussian likelihood. Under the additional restriction that the processes $\{\alpha_t\}$ and $\{\epsilon_t\}$ are strictly stationary, the semiparametric FSN model of Definition 2 has exactly this form. Since $\{\alpha_t\}$ is a VAR process, it is necessary but not sufficient for the strict stationarity of $\{\alpha_t\}$ that the roots of the matrix polynomial $h(z) = I_m - \Phi z$ all lie outside the unit circle (a condition that is satisfied by the point estimate of Φ obtained in the empirical analysis of section 4.1). Also, it would not be overly restrictive to take $\{\epsilon_t\}$ to be *i.i.d.* in the semiparametric FSN model, which would obviously satisfy the strict stationarity requirement. The asymptotic results of Dunsmuir (1979) thus provide a theoretical justification for the QML method of inference employed here for the semiparametric FSN model.[11]

Finally it is noted that since the semiparametric FSN model in Definition 2 is in the linear state space form, the one-step ahead point predictor given by the Kalman filter, $[\hat{p}_t(\nu)| p_{t-1}(\nu), ..., p_1(\nu)]_{KF}$, is a minimum mean square error predictor in the class of predictors that are a linear function of the past observations $(p_{t-1}(\nu), ..., p_1(\nu))$.[12] Furthermore, the one-step ahead prediction errors $e_t := p_t(\nu) - [\hat{p}_t(\nu)|p_{t-1}(\nu), ..., p_1(\nu)]_{KF}$ satisfy

$$\mathsf{E}[e_t] = 0,\ \mathsf{Var}[e_t] = F_t,\ \mathsf{E}[e_t e_s'] = 0\ \forall t \neq s, \tag{6}$$

where $F_t = ZP_tZ' + \Omega_\epsilon$ is constant for all $t > 1$. The matrix P_t is the unconditional covariance matrix of the error of the minimum mean square linear estimator of α_t and is given by the output of the Kalman filter. These results hold exactly when the parameters of the model are known (for details see Harvey (1989, pp. 110-112)). Use is made of the properties in (6) when evaluating the in-sample predictive performance of the semiparametric FSN model in section 4.2 below.

4 Empirical Analysis of the SETS Average Price Curves

The growing importance of electronic LOBs as a means of trading financial securities worldwide has already been noted. However, empirical work analysing the data produced by such LOBs is relatively sparse compared, for example, with the volume of literature analysing NYSE equity data. We partially address this lacuna by presenting an analysis of the London Stock Exchange's SETS electronic limit order book for the heavily-traded FTSE100 stock AstraZeneca. Furthermore, this is one of the first studies to focus on the times series properties of the bid and ask *curves*, as opposed to those of the mid-quote or bid and ask prices.

[11]Note that in a semiparametric FSN model of ask curves, say, $\{\alpha_t\}$ is not thought to have zero mean since on average the ask curves are strictly increasing. Dunsmuir (1979) assumes that $\mathsf{E}[s_t] = 0$ but it is conjectured here that his results will continue to hold in the case of a constant, non-zero mean.

[12]Note that $[\hat{p}_t(\nu)|p_{t-1}(\nu), ..., p_1(\nu)]_{KF}$ is only guaranteed to equal $\mathsf{E}[p_t(\nu)|p_{t-1}(\nu), ..., p_1(\nu)]$ for the Gaussian FSN model.

The dataset covers the period from 2 January 2002 to 31 January 2002 inclusive and was described previously in section 2.2. In the empirical analysis, each price curve is taken to be the vector of dimension 40 consisting of the prices corresponding to the volumes $\nu = (500, 1000, ..., 19\ 500, 20\ 000)'$. The maximum volume of 20 000 was chosen so as to be less than the total volume of limit orders on the relevant side of the book for the majority of trading days but very rarely exceeded by the size of an individual market order.[13]

There follows a discussion of the QML estimates obtained using a restricted semiparametric FSN model and then an evaluation of the in-sample, 1-step ahead predictions given by that model.

4.1 QML estimates of the FSN model

Recall that the semiparametric FSN model discussed in section 3 is applied here as a marginal model for a time series of either bid or ask curves. The primary relevance of one or the other to an agency trader has already been noted (see footnote 8). The results of fitting a restricted semiparametric FSN model to both the AstraZeneca bid and ask curve datasets in turn are presented below. In each case, the average price curves corresponding to the dth day, $(p_1^{(d)}(\nu)', ..., p_T^{(d)}(\nu)')'$, are modelled as being independent and identically distributed across days. The extension to a joint model of both bid and ask curves that nests the marginal models is discussed in section 6.2. It is noted that no assumption of independence between the bid and ask curves is being invoked here.

The semiparametric FSN model used imposes the following additional restrictions to those in Definition 2: $\Phi = diag(\phi_1, ..., \phi_m)'$, $\Omega_\epsilon = \sigma_\epsilon^2 I_n$, and $\alpha_1 \sim (a_1, \Omega_1)$ with a_1 and Ω_1 chosen so that $\{\alpha_t\}_{t=1}^T$ is covariance stationary when $|\phi_j| < 1\ \forall j$. The knots of the cubic spline are positioned at the volumes $k = (1, 2500, 5000, 10\ 000, 20\ 000)'$. Thus the dimension of the state vector, $m = 4$, is considerably smaller than that of the observations $p_t^{(d)}(\nu)$ at $n = 40$. The quasi log-likelihood may then be computed by applying the Kalman filter to the Gaussian FSN model with the same additional restrictions imposed.

Numerical optimisation of the log-likelihood was performed using the MaxBFGS algorithm

[13] Trading days on which the total depth on the relevant side of the market fell below 20 000 at some time were excluded from the analysis. This resulted in there being 21 and 19 days in the final datasets for the bid and ask curves respectively.

δ'	-0.267	-0.243	-0.371	-0.486	σ^2_ϵ	0.018			
ϕ'	0.363	0.431	0.595	0.774					
Ω_η	0.598	0.123	0.054	0.051	$Corr_\eta$	1.000	0.298	0.087	0.054
	-	0.286	0.213	0.046		-	1.000	0.495	0.070
	-	-	0.649	0.473		-	-	1.000	0.480
	-	-	-	1.490		-	-	-	1.000

Table 1: QMLEs of the semiparametric FSN model for the bid curves of AstraZeneca in January 2002. Results based on 21 independent trading days; $\phi = (\phi_1, ..., \phi_m)'$; $Corr_\eta$ is the correlation matrix corresponding to Ω_η.

with numerical derivatives in Ox (see Doornik (2001)).[14] In order to provide confidence in the numerical robustness of the results, the quasi log-likelihood resulting from a slight variant of the Gaussian FSN model just described was optimised using a method incorporating an EM algorithm step. The variant model differs only in the initialisation used, namely $\alpha_1 \sim \mathsf{N}(a_1, \Omega_\eta)$ where $a_1 = (\delta_j/[1-\phi_j])_{j=1}^m$ as before, so that $\{\alpha_t\}_{t=1}^T$ is still mean stationary but not covariance stationary. This model is useful because it permits the analytical derivation of an EM step for the estimation of the parameters of the two covariance matrices. The method gave very similar QMLEs (not reported here) and is described in detail in the Appendix (see section 8.2).

Tables 1 and 2 show the QMLEs for the bid and ask curves respectively. The results exhibit a considerable degree of symmetry between the bid and ask sides of the market and so are discussed together. First recall that the knots of the spline are not equally spaced, and that the inter-knot volume differentials, $\Delta k_j := k_j - k_{j-1}$, satisfy $\Delta k_1 = \Delta k_2 < \Delta k_3 < \Delta k_4$. The covariance stationary AR(1) processes for the inter-knot price differentials, $\{\alpha_{jt}\}$, obtained by setting the parameter values equal to their QMLEs have the following properties. The mean is negative (positive) for the bid (ask) curves and increases in absolute magnitude as j increases. The variance of α_{jt} is also increasing in j (except for the decrease on going from $j = 1$ to $j = 2$), as is the autoregressive parameter ϕ_j. The increase in persistence can be partially explained in terms of limit orders being added to and removed from the book (due to order placement, execution and cancellation) less frequently at prices that are further from the best ask and bid.

[14] Approximate standard errors could be computed here using a sandwich-type estimate incorporating an 'outer product of gradients' (or OPG). However, evaluation of the OPG estimate for the semiparametric FSN model is computationally very intensive when the number of time series observations is large, as is the case here. Future work will instead focus directly on the impact of parameter uncertainty on the out-of-sample, point forecasts and predictive distributions of the FSN model.

δ'	0.276	0.246	0.361	0.452	σ_ϵ^2	0.026			
ϕ'	0.424	0.499	0.651	0.80					
Ω_η	0.774	0.125	0.008	0.027	$Corr_\eta$	1.000	0.238	0.010	0.023
	-	0.356	0.201	0.015		-	1.000	0.393	0.019
	-	-	0.731	0.475		-	-	1.000	0.418
	-	-	-	1.771		-	-	-	1.000

Table 2: QMLEs of the semiparametric FSN model for the ask curves of AstraZeneca in January 2002. Results based on 19 independent trading days; $\phi = (\phi_1, ..., \phi_m)'$; $Corr_\eta$ is the correlation matrix corresponding to Ω_η.

For $j = 4$ there is a high degree of persistence.

The inter-knot price differentials are also contemporaneously positively correlated across j, with the highest correlations occurring for immediately neighbouring parts of the spline and with other correlations being relatively small. This is an interesting feature of the results – for example, a 'steeper' than average tail end of the price curve does not tend to be strongly associated with a 'steeper' than average curve for smaller volumes (i.e. those less than or equal to 2500 shares). Another important feature is that since the state vector α_t appears to be covariance stationary, so then are the spline signal, $S_{\alpha_t^*}(\nu) = Z\alpha_t$, and the average price curve, $p_t(\nu)$. Furthermore, the evaluation of the in-sample prediction errors of the model presented below strongly suggests that it is not necessary to incorporate a deterministic, time-of-day 'level effect' in the model specification. This is quite surprising and suggests that while bid-ask spreads exhibit well-known intradaily deterministic patterns, the differences between the price curves and their intercepts (i.e. the immediate price impacts of market orders) are covariance stationary.[15]

4.2 In-sample predictive evaluation of the FSN model

The specification of the semiparametric FSN model may be evaluated by analysing the in-sample, 1-step ahead point predictions, $[\hat{p}_t^{(d)}(\nu)|\ p_{t-1}^{(d)}(\nu), ..., p_1^{(d)}(\nu)]_{KF}$, given by the Kalman filter for each trading day d. It was noted previously that, since the model has the linear state space form, this one-step ahead predictor is a minimum mean square linear predictor. The analysis below of the 1-step ahead prediction errors, $e_t^{(d)} := p_t^{(d)}(\nu) -\ [\hat{p}_t^{(d)}(\nu)|\ p_{t-1}^{(d)}(\nu), ..., p_1^{(d)}(\nu)]_{KF}$, reveals that the semiparametric FSN model performs well in an in-sample predictive evaluation.

[15] Coppejans, Domowitz, and Madhavan (2003) also report that the considerable variation over time in aggregate volume on the book at various (price) distances from the mid-quote is not readily explicable in terms of simple time-of-day effects.

$volume\ (v)$	$Mean^a$	Var^a	$MAPE^a$	$RMSE^a$	$Mean^b$	Var^b	$MAPE^b$	$RMSE^b$
2000	0.007	1.034	0.457	0.791	-0.003	1.023	0.411	0.686
4000	-0.007	0.953	0.660	1.065	0.006	0.963	0.599	0.954
6000	-0.005	0.984	0.808	1.273	0.003	0.990	0.731	1.166
8000	0.001	1.005	0.929	1.445	-0.002	1.006	0.853	1.356
10 000	0.003	0.998	1.034	1.600	-0.002	0.999	0.962	1.529
12 000	0.001	0.993	1.122	1.754	0.001	0.994	1.057	1.689
14 000	-0.006	1.008	1.208	1.904	0.004	1.004	1.149	1.836
16 000	-0.009	1.013	1.287	2.036	0.007	1.010	1.229	1.969
18 000	-0.005	0.997	1.357	2.159	0.004	1.003	1.300	2.093
20 000	0.004	0.985	1.432	2.311	-0.002	0.980	1.368	2.219

Table 3: Properties of in-sample, 1-step ahead prediction errors from the Kalman filter for the AstraZeneca dataset in January 2002. A superscript 'a' indicates the ask side, and a 'b' the bid side. The mean and variance (Var) are for the standardised errors, and the RMSE and MAPE are for the errors without standardisation.

Let $e_t^{(d)}(v)$ be the element of the $e_t^{(d)}$ vector corresponding to volume v, and similarly denote by $F_t^{(d)}(v)$ the element of the diagonal of the $F_t^{(d)}$ matrix that gives the variance of $e_t^{(d)}(v)$ under the assumption that the semiparametric FSN model is correctly specified (see equation (6)). We may then consider the standardised prediction errors, $\tilde{e}_t^{(d)}(v) := e_t^{(d)}(v)/(F_t^{(d)}(v))^{1/2}$. It follows from the discussion in section 3.2 that, under correct specification, $\mathsf{E}[\tilde{e}_t^{(d)}(v)] = 0$ and $\mathsf{Var}[\tilde{e}_t^{(d)}(v)] = 1\ \forall(v,t,d)$, and $\mathsf{E}[\tilde{e}_s^{(c)}(v)\tilde{e}_t^{(d)}(v)] = 0\ \forall(v,s,t,c,d)$ such that $s \neq t$ if $c = d$. Clearly each series, $\{\tilde{e}_1^{(d)}(v), ..., \tilde{e}_T^{(d)}(v)\}_{d=1,2,...}$, formed by concatenating the individual series corresponding to each trading day is white noise. The in-sample predictive performance of the semiparametric FSN model may thus be assessed by examining the sample mean, variance and autocorrelations of the standardised 1-step ahead prediction errors for a range of volumes, v.[16]

Table 3 reports, using both the bid and ask datasets, the sample mean and sample variance of the standardised errors for $v \in \{2000, 4000, ..., 20\ 000\}$, together with the root mean square error (RMSE) and mean absolute prediction error (MAPE) corresponding to the non-standardised errors, $\{e_t^{(d)}(v)\}_{t,d}$. Figures 3 and 4 plot, for the bid and ask curves respectively, the autocorrelation functions of the series $\{\tilde{e}_1^{(d)}(v), ..., \tilde{e}_T^{(d)}(v)\}_d$ for lags up to a time period of 30 minutes and compare them to those for the data, $\{p_1^{(d)}(v), ..., p_T^{(d)}(v)\}_d$. The means and variances of the standardised errors are all close to zero and one respectively, revealing no model misspecification. Furthermore, the model is rather successful at 'whitening' the autocorrelation functions of the original data (particularly at higher volumes v), although there appears to be

[16]For this purpose, the parameter values are set equal to the QMLEs obtained above before applying the Kalman filter.

some residual positive autocorrelation of the prediction errors.[17] Overall, the point predictions are unbiased and the in-sample predictive performance of the restricted semiparametric FSN model estimated in section 4.1 above is pleasing. Finally, note that the RMSEs and MAPEs are similar when comparing the bid and ask sides, and are increasing in the volume, v.

With the Gaussian FSN model as the DGP, the standardised errors $\{\tilde{e}_t^{(d)}(v)\}_{t,d}$ are *i.i.d.* $\mathsf{N}(0,1)$. A graphical analysis of these errors (not shown here) immediately reveals that they are far from Gaussian and strongly asymmetric. The histograms of the standardised errors for the ask curves, for example, are skewed to the left with long right tails. Intuitively, this can be explained in terms of the Gaussian model attaching positive probability mass to negative inter-knot price differentials for the ask splines despite the fact that downward sloping ask curves are never observed in practice. This then results in fewer large over-predictions than are expected under Gaussianity. Similarly, the histograms of the standardised errors for the bid curves are skewed to the right with long left tails. In both cases the extent of the observed asymmetry decreases as v increases. Clearly these observations do not invalidate the use of the Gaussian FSN model as the *auxiliary* model in the QML estimation procedure. They do however emphasise the need to move away from Gaussianity in order to develop fully parametric FSN models that appropriately address the monotonicity property of the average price curves. This is the subject matter of the following section.

5 Towards Fully Parametric FSN Models

The motivation in terms of developing econometric models that can be used to provide decision theoretic solutions to trading decision problems has already been emphasised. In general, this requires that the econometric model deliver 1-step ahead, conditional predictive *distributions* for the average price curves. More generally, such a predictive distribution can be seen as the most complete solution to the forecasting problem. This section therefore develops fully parametric, non-Gaussian FSN models contained in the semiparametric FSN model presented above that take account of the monotonicity of the average price curves.

[17] Indeed, formal Box-Ljung tests reject the null of zero autocorrelation at the 1% level. Further development of the model to accommodate this feature is left to future work.

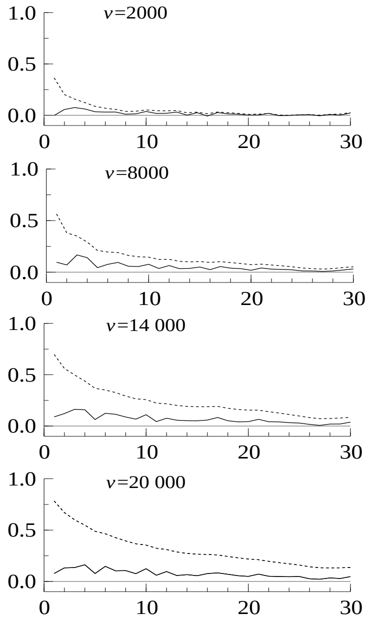

Figure 3: Autocorrelation functions of the 1-step ahead prediction errors from the Kalman filter for the AstraZeneca bid curves in January 2002. Autocorrelation function for the observed data shown as a broken line and that for the prediction errors as a continuous line.

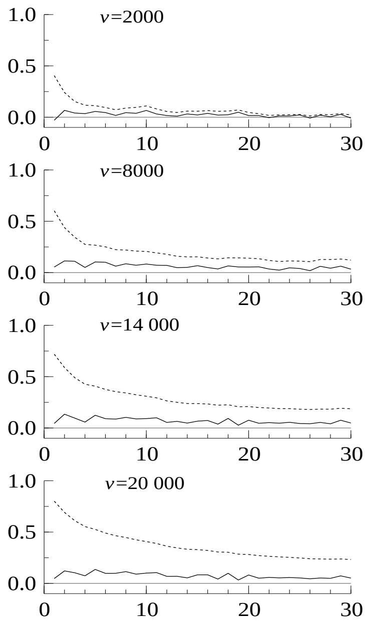

Figure 4: Autocorrelation functions of the 1-step ahead prediction errors from the Kalman filter for the AstraZeneca ask curves in January 2002. Autocorrelation function for the observed data shown as a broken line and that for the prediction errors as a continuous line.

5.1 The FSN-MEM model

The FSN models presented in this section possess latent splines that are *quasi-monotonic* (in the sense of Definition 5 of the Appendix) – i.e. the differentials between the knot-prices are *non-positive* in the case of a bid curve and *non-negative* for an ask curve. This is achieved by adopting a 'multiplicative error' specification for the state equation in which the multiplicative error, ζ_t, has non-negative elements with probability one.[18] Engle (2002, pp. 428-434) has stressed the potential usefulness of such multiplicative error models (MEMs) for non-negative processes in financial econometrics. Non-positive processes may also be modelled using MEMs, as is demonstrated below.

There now follows a formal definition of the so-called FSN-MEM model.

Definition 3 ***FSN-MEM Model.*** *Let $I_{side} = -1$ if the model is for a bid curve and $I_{side} = +1$ for an ask curve. The FSN-MEM model is given by*

$$\begin{aligned} p_t(\nu) &= S_{\alpha_t^*}(\nu) + \epsilon_t \\ &= Z\alpha_t + \epsilon_t, && (7) \\ \alpha_{t+1} &= (\delta + \Phi\alpha_t) \odot \zeta_t, && (8) \end{aligned}$$

for $t = 1, ..., T$, where $S_{\alpha_t^*}(\nu)$ is a natural cubic spline on $(k; \alpha_t^*)$ and Z, α_t^* are defined as in Definition 2; the random variables $(\alpha_1, \zeta_1, ..., \zeta_{T-1}, \epsilon_1, ..., \epsilon_T)$ are independent; $\epsilon_t \sim (0, \Omega_\epsilon)$ $\forall t$; $\zeta_t \sim (\mathbf{1}_m, \Omega_\zeta)$ and has support $[0, \infty)^m := [0, \infty) \times ... \times [0, \infty)$ $\forall t$; $\alpha_1 \sim (a_1, \Omega_1)$ and has support $(-\infty, 0]$ if $I_{side} = -1$ and support $[0, \infty)$ if $I_{side} = +1$; and $0 \leq \phi_j$ $\forall j, \delta I_{side} > 0$.

The model is still semiparametric at this stage and it is shown below in Proposition 1 that it is in fact a restricted version of the semiparametric FSN model of Definition 2. The most important restrictions are those ensuring that, with probability one, $I_{side}\alpha_t \geq 0$ – i.e. those ensuring the quasi-monotonicity of the spline signal function (see Definition 5). It also follows from Proposition 1 and standard ARMA theory that $\{\alpha_t\}$ is covariance stationary when $0 < \phi_j < 1$ $\forall j$ (and a_1, Ω_1 are chosen appropriately).

[18]The approach taken here offers substantial gains in terms of the simplicity and parsimony of the resulting model. We conjecture that this quasi-monotonic approach will also prove successful in empirical forecasting since non-monotonic signal functions are assigned a small probability under the model for empirically reasonable parameter estimates.

Proposition 1 *The FSN-MEM model in Definition 3 is a restricted version of the semiparametric FSN model of Definition 2.*

Proof. Equation (8) may be written as

$$\alpha_{t+1} = \delta + \Phi\alpha_t + \eta_t, \quad t = 1, ..., T, \tag{9}$$

where the random variable $\eta_t := \alpha_{t+1} - (\delta + \Phi\alpha_t)$. Let $\mathcal{F}_t^\alpha := \sigma(\alpha_1, ..., \alpha_{t+1})$, $t = 0, ..., T-1$. Then $\{\eta_t\}$ is a vector martingale difference sequence (MDS) with respect to $\{\mathcal{F}_t^\alpha\}$ since ζ_t is independent of $\mathcal{F}_t^\alpha$. Since η_t is a MDS, $\mathsf{E}[\eta_t] = 0$ and $\mathsf{E}[\eta_t\eta_s'] = 0 \ \forall s \neq t$.

Since ζ_t and α_t are independent, and $\mathsf{E}|\zeta_t| < \infty, \mathsf{E}|\alpha_1| < \infty$, it follows by induction that $\mathsf{E}|\alpha_t| < \infty$ and $\mathsf{E}|\alpha_t\alpha_t'| < \infty$ $(t = 2, ..., T)$.[19] Similarly, since

$$\eta_{jt} = (\zeta_{jt} - 1)(\delta_j + \phi_j\alpha_{jt}), \quad \forall j, t, \tag{10}$$

it holds that $\mathsf{E}|\eta_t\eta_t'| < \infty$. It has now been established that $\{\eta_t\}$ is white noise. Note that the sequence $\{\epsilon_t\}$ is *i.i.d.* and hence also white noise.

It remains to show that $\mathsf{E}[\eta_t\epsilon_s'] = 0$, $\mathsf{E}[\eta_t a_1'] = 0 \ \forall t, s$. The first again follows from (10) since the elements of the matrix $\eta_t\epsilon_s'$ are then the product of 3 independent random variables with finite mean and $\mathsf{E}[\zeta_{jt} - 1] = 0$. The second holds since ζ_t is independent of $\sigma(\alpha_1, \zeta_1, ..., \zeta_{t-1})$, and by induction $\mathsf{E}|\alpha_t\alpha_1'| < \infty \ \forall t$. ■

The importance of Proposition 1 is that, since the semiparametric FSN model has been shown to be useful in the empirical analysis of bid and ask curves, it is desirable to develop fully parametric models taking into account the monotonicity of the average price curves that are consistent with that semiparametric model. Such parametric FSN-MEM models may be specified by taking Definition 3 and specifying distributions for ϵ_t and the non-negative random vectors ζ_t and $I_{side}\alpha_1$. Retaining the Gaussian distribution of ϵ_t would seem to be a reasonable first approach, perhaps paying greater attention to the potential cross-sectional dependence amongst the elements of ϵ_t. The specification of the non-negative random vectors is less straightforward.

Although there is a literature on distributions for such random vectors with marginal distributions that all belong to the same parametric family (e.g. multivariate gamma distributions) – see Kotz, Balakrishnan, and Johnson (2000) – none of these seem well-suited to an analysis

[19] The following fact is used repeatedly here: if the r.v.'s X and Y are independent and $\mathsf{E}|X| < \infty, \mathsf{E}|Y| < \infty$ then $\mathsf{E}|XY| < \infty$ and $\mathsf{E}[XY] = \mathsf{E}[X]\mathsf{E}[Y]$.

in which 1-step ahead conditional densities play an important role. We are examining various alternative approaches in ongoing research. The resultant parametric FSN-MEM models may be estimated using either a variant of the QMLE approach described in section 3.2 or more efficiently by Markov Chain Monte Carlo methods. An evaluation of the empirical forecasting performance of the 1-step ahead predictive distributions given by these fully parametric models is in progress.

6 Comparison with Previous Research and Possible Extensions

6.1 Previous literature on the dynamics of liquidity and price curves

Previous research on the dynamics of liquidity has concentrated on the dynamics of scalar summary measures. For example, Coppejans, Domowitz, and Madhavan (2003) focus on the time series properties of the aggregate volume on the book within six ticks of the mid-quote. Engle and Lange (2001) seek to explain the determinants of VNET, their measure of the excess of buyer- or seller-initiated volume occurring during the time taken for the mid-quote to move by some predetermined, constant amount. Neither approach explicitly considers the bid and ask curves, nor lends itself to the prediction of those curves in the future.

Coppejans, Domowitz, and Madhavan (2003) also report mean price impact functions on the Swedish electronic LOB for stock index futures, but do not consider their dynamics. Indeed, the paper by Gouriéroux, Le Fol, and Meyer (1998) is ostensibly the only previous study that examines the dynamics of the bid and ask curves of an electronic LOB. It differs from the modelling framework presented here in an important respect, namely that the model is set in transaction time – i.e. the tth average price curve is now the curve observed at the time of the tth trade. Letting $p_t^b(\nu)$ and $p_t^a(\nu)$ now denote the tth bid and ask curves proper (as defined by Definition 1), with $\nu = (50, 100, ..., 5000)'$, their two factor model for the logarithm of the curves is given by

$$\left.\begin{array}{l} \log p_t^a(\nu) = a_0(\nu) + a_1(\nu)\log m_t(50) + a_2(\nu)\log s_t(2000) + u_t^a(\nu) \\ \log p_t^b(\nu) = b_0(\nu) + b_1(\nu)\log m_t(50) + b_2(\nu)\log s_t(2000) + u_t^b(\nu) \end{array}\right\}, \tag{11}$$

where the 'mid-price' and 'spread'-type factors are given by $m_t(50) := [p_t^a(50) + p_t^b(50)]/2$ and $s_t(2000) := p_t^a(2000) - p_t^b(2000)$ respectively, and $(u_t^a(\nu)', u_t^b(\nu)')'$ is white noise and independent of both factors at time t.[20] The coefficient functions $\{a_r(\nu), b_r(\nu)\}_{r=0}^2$ are approximated

[20] Note that $\log p_t^a(\nu) := (\log p_t^a(50), ..., \log p_t^a(5000))'$.

by piecewise linear functions. Nonlinear autoregressive models for the conditional mean of $(\log m_t(50) - \log m_{t-1}(50))$ and $\log s_t(2000)$ then allow the 1-step ahead point prediction of $\log p_t^a(\nu)$ and $\log p_t^b(\nu)$ given the curves observed prior to time t.

In common with our FSN model, the two factor model specifies the dynamics of a low dimensional vector which determines the price curve in a manner that is fixed over time. However, in the FSN model this low dimensional vector contains latent price differentials (or 'states') rather than observable factors. Other noteworthy differences are that the bid curve at 'time' $(t-1)$ is *necessarily* informative about the ask curve at 'time' t in (11) and that the two factor model imposes no conditions that take into account the monotonicity of the price curves.[21] Gouriéroux, Le Fol, and Meyer (1998) do not discuss fully parametric specifications using their two factor framework.

6.2 Suggestions for further research

The estimation of fully parametric FSN models and the empirical evaluation of their associated predictive distributions has already been discussed. One further direction for future research will be indicated here.

The FSN models can readily be expanded both to incorporate the dynamics of the best bid and ask (i.e. the intercepts of the curves) and to allow dependence between the bid and ask curves. For example, letting $p_t^b(\nu)$ and $p_t^a(\nu)$ again denote the tth bid and ask curves proper (as defined by Definition 1), a possible model in linear state space form is:

$$\left.\begin{array}{c} \begin{pmatrix} p_t^a(\nu) \\ p_t^b(\nu) \end{pmatrix} = \begin{pmatrix} m_t + \frac{s_t}{2} \\ m_t - \frac{s_t}{2} \end{pmatrix} + Z \begin{pmatrix} \alpha_t^a \\ \alpha_t^b \end{pmatrix} + \begin{pmatrix} \epsilon_t^a \\ \epsilon_t^b \end{pmatrix} \\ \begin{pmatrix} \alpha_{t+1}^a \\ \alpha_{t+1}^b \end{pmatrix} = \delta + \begin{pmatrix} \Phi^a & 0 \\ 0 & \Phi^b \end{pmatrix} \begin{pmatrix} \alpha_t^a \\ \alpha_t^b \end{pmatrix} + \begin{pmatrix} \eta_t^a \\ \eta_t^b \end{pmatrix} \\ m_{t+1} = m_t + g(\alpha_t^b, \alpha_t^a) + \eta_t^m \\ s_{t+1} = \delta^s + \phi s_t + \eta_t^s \end{array}\right\}, \tag{12}$$

where $m_t := [p_t^{b*} + p_t^{a*}]/2$ and $s_t := p_t^{a*} - p_t^{b*}$ are the mid-quote and bid-ask spread at time t, and g is a linear function of the price differentials. The measurement equation is as before except that an intercept term is now present and $p_t^a(\nu)$ and $p_t^b(\nu)$ are modelled jointly. The random variables m_t and s_t are considered as state variables, possibly observed without measurement error. Note that the marginal models for $\{p_t^a(\nu) - p_t^{a*}\}$ and $\{p_t^b(\nu) - p_t^{b*}\}$ implied by equation

[21] Although equation (11) appears to include a specification of the dynamics of the best bid and ask prices, empirical results presented in the paper do not consider volumes less than 50 shares.

(12) are, under the appropriate additional assumptions, the same as the semiparametric FSN model of Definition 2 that was used for the empirical analysis of exactly these time series in section 4.

The above formulation is intended to address two economic questions of interest. First, the extent to which information about the entire bid and ask curves helps in predicting movements in the mid-quote, as captured by the function g. For example, one possible specification for the price pressure resulting from a limit order imbalance between the two sides of the market is $g(\alpha_t^b, \alpha_t^a) = \sum_{j=1}^m w_j(\alpha_{jt}^a - |\alpha_{jt}^b|)$, $w_j \geq 0 \ \forall j$. Second, the extent to which 'liquidity shocks' on the bid and ask sides of the market are related – i.e. the degree of dependence between the random vectors η_t^a and η_t^b. Previous work tentatively suggests that these shocks may be uncorrelated or independent (see Daníelsson and Payne (2001)).

7 Conclusion

This paper has developed a novel econometric framework for modelling a panel of prices of very high cross-sectional dimension, in which contemporaneous observations are functionally related. Given the growing importance of fully automated financial markets, the new Functional Signal plus Noise (FSN) models are applied in an analysis of the bid and ask curves of an electronic limit order book. This enables the comprehensive measurement of the determinants of traders' execution costs (i.e. 'liquidity') in a way that takes the dynamic aspects of the problem into account. The models should also prove useful in the solution of the decision problems faced by traders in practice, either as the basis for automated trading algorithms or for trading 'decision support' tools used by human traders.

A semiparametric FSN model is developed in which a smooth, cubic spline signal function is used to approximate the price curve data. The state equation of the FSN model determines the stochastic evolution of a low-dimensional vector of prices corresponding to the knots of the spline, and hence the dynamics of the entire spline signal. The choice of a cubic spline for the signal function results in models that may be written in linear state space form, thus allowing the use of the Kalman filter to compute both 1-step ahead linear point predictions and the Gaussian quasi-likelihood function for estimation. The flexibility of cubic splines as approximating functions gives the FSN framework potentially broad applicability in modelling the dynamics of cross-sectional price functions in other settings.

The semiparametric FSN model is used here to provide one of the first studies of the dynamics of the bid and ask (or 'average price') curves of a limit order book, namely the London Stock Exchange's fully electronic SETS order book. The differences between the average price curves and their intercepts give the immediate price impact of a market order, as a function of the volume of the order. It is found that, in contrast to the widely reported intraday behaviour of the bid-ask spread, these price impact functions are well described by covariance stationary processes which lack deterministic time-of-day effects. The price differentials between the knots of the spline signal function are modelled by covariance stationary AR(1) processes whose persistence is found to increase on moving towards higher volumes. Contemporaneous correlations between the price differentials of the spline are positive and much larger for neighbouring parts of the spline with the implication that, for example, a steeper than average tail end of the price curve does not tend to be strongly associated with a steeper than average curve for smaller volumes.

The in-sample, one-step ahead point predictions given by the semi-parametric FSN model perform well. The prediction errors are unbiased, have low levels of auto-correlation and are able to match the observed average price curves quite closely. The prediction errors are also found to be highly non-Gaussian, thus motivating the development of parametric FSN models contained in the original semiparametric model that explicitly take into account the monotonicity of the average price curves. The proposed FSN-MEM model is based on a Multiplicative Error Model for the state equation that ensures that the price differentials of the ask (bid) spline signal are non-negative (non-positive). Future work will focus on the conditional predictive distributions given by fully parametric FSN-MEM models and their empirical usefulness in deriving optimal order submission and trading strategies.

Acknowledgments

I would like to thank David Hendry, Siem Jan Koopman, Roland Meeks, Bent Nielsen, Mark Salmon, and Neil Shephard for their comments. Any remaining errors are of course my own. Computational work was performed using Version 3.2 of Ox (Doornik 2001) and SsfPack 2.2 (Koopman, Shephard and Doornik 1999). Financial support from the ESRC under awards R42200034061 and R00023839 is gratefully acknowledged.

8 Appendix

8.1 A primer on cubic spline theory

A cubic spline is essentially a piecewise cubic function in which the pieces join together to form a smooth function overall.

Definition 4 ***Cubic spline on*** $(k; y)$. *Consider an interval of the real line* $[a, b]$*, subdivided by a vector,* k*, of points*

$$k = (k_j)_{j=0}^{m},$$

where $k_0 = a, k_m = b, k_{j+1} > k_j$ *for* $j = 0, ..., m-1$. *Each point* k_j *is referred to as a* ***knot.*** *Denote by* y *a vector of real-valued ordinates* $(y_j)_{j=0}^{m}$. *A function* $S(x)$ *with domain* $[a, b]$ *is a* ***cubic spline interpolating to*** y ***with knots*** k*, or more concisely a* ***cubic spline on*** $(k; y)$*, if and only if:*

(i) $S(k_j) = y_j$ $(j = 0, ..., m)$;

(ii) $S(x)$ *coincides with a polynomial of degree at most three on the sub-intervals* $[k_j, k_{j+1}]$ $(j = 0, ..., m-1)$; *and*

(iii) $S(x)$ *is twice continuously differentiable on* $[a, b]$.

If, in addition, the second derivatives at the end knots, $S''(k_0)$ *and* $S''(k_m)$*, are both zero,* $S(x)$ *is said to be a* ***natural cubic spline on*** $(k; y)$.

In this context, since the knots are positioned at deterministic volumes that are fixed throughout the analysis, whereas the states y to which the spline interpolates are stochastic, a cubic spline $S(x)$ on $(k; y)$ is denoted here by $S_y(x)$.

Definition 5 *Let* $S_y(x)$ *be a cubic spline on* $(k; y)$. *Then* $S_y(x)$ *is said to be* ***quasi-monotonic*** *if and only if* $(y_{j-1} \leq y_j \ \forall\ j = 1, ..., m)$ *or* $(y_{j-1} \geq y_j \ \forall\ j = 1, ..., m)$.

Note that quasi-monotonicity is necessary but not sufficient for the monotonicity of the cubic spline $S_y(x)$ on $[a, b]$.

The object of interest here is usually the restriction of $S_y(x)$ to a finite vector of points $\nu = (v_1, ..., v_n)'$. The cubic spline is then written as the finite dimensional vector $S_y(\nu) := (S_y(v_1), ..., S_y(v_n))'$. A well known result that arises by combining conditions on $S''(k_0)$ and $S''(k_m)$ with the 'continuity restrictions' implied by conditions (i), (ii) and (iii) of Definition 4 is that $S_y(\nu)$ is a linear function of the ordinate vector y. This result, stated for the case of a natural cubic spline in the theorem below, allows the FSN models to be written in linear state space form.

Theorem 2 *Let $S_y(x)$ be a natural cubic spline on $(k; y)$, with k and y vectors of dimension $m+1$. Also let $\nu = (v_1, ..., v_n)'$ be a finite vector of points in $[a, b]$, and $S_y(\nu) := (S_y(v_1), ..., S_y(v_n))'$. Then*

$$S_y(\nu) = Wy,$$

where the $n \times (m+1)$ weight matrix W depends only on ν and the knot positions k.[22] *Details of how to compute W may be found in equations (2.5),(2.6),(2.11),(2.12) and (2.14) of Poirier (1973, pp. 517-518), where π_0 and π_k should be set to zero for the case of a natural cubic spline.*

Proof. See, for example, Poirier (1973, pp. 517-518). ■

8.2 Computation of QMLEs using the EM algorithm

As was noted in section 4, the quasi log-likelihood resulting from a slight variant of the Gaussian FSN model described in that section may be optimised using a method involving the EM algorithm. The variant model differs only in the initialisation used, namely $\alpha_1 \sim \mathsf{N}(a_1, \Omega_\eta)$ where $a_1 = (\delta_j/[1-\phi_j])_{j=1}^m$ as before, so that $\{\alpha_t\}_{t=1}^T$ is still mean stationary but no longer covariance stationary. This alternative initialisation allows the analytical derivation of an EM step for the covariance matrix Ω_η whose implementation relies on the moment and state smoothers for state space models (see Koopman (1993) and Koopman, Shephard, and Doornik (1999)).

The optimisation method used involves two steps: an EM step in which the log-likelihood of the 'variant Gaussian FSN' model is maximised with respect to the parameters of the covariance matrices Ω_η and Ω_ϵ (holding all other parameters fixed); and a numerical optimisation step in which the log-likelihood is maximised with respect to these other parameters (holding only Ω_η and Ω_ϵ fixed). The method then consists of iterating between the two steps until convergence is achieved. For the numerical optimisation step, the MaxBFGS algorithm with numerical derivatives in Ox was used.

[22]In particular, it is implicit in the statement that '$S_y(\nu)$ is a linear function of y' that W does not depend on y.

The EM step comprises two updating equations which give a new set of covariance matrices, $\Omega_\eta^\dagger$ and $\Omega_\epsilon^\dagger = (\sigma_\epsilon^2)^\dagger I_n$, which always result in a non-decrease in the log-likelihood. The previous values of the covariance matrices, prior to the application of the EM updating equations, are denoted here by Ω_η^* and $\Omega_\epsilon^* = (\sigma_\epsilon^2)^* I_n$. The two equations, stated for the simplest case of a time series of bid or ask curves observed during a single trading day at times $t = 1, ..., T$, are given by

$$\Omega_\eta^\dagger = \frac{T}{T+1}\Omega_\eta^* + \Omega_\eta^* \left\{ \frac{1}{T+1} \left[\sum_{t=1}^{T} r_t r_t' - N_t \right] \right\} \Omega_\eta^* + \frac{1}{T+1}(V_1 + W_1 W_1'), \tag{13}$$

where $W_1 = \hat{\alpha}_1 - (I_m - \Phi)^{-1}\delta$, and

$$(\sigma_\epsilon^2)^\dagger = (\sigma_\epsilon^2)^* + n^{-1} tr \left\{ (\sigma_\epsilon^2)^* \left[T^{-1} \sum_{t=1}^{T} e_t e_t' - D_t \right] (\sigma_\epsilon^2)^* \right\}. \tag{14}$$

The smoothed quantities (e_t, D_t, r_t, N_t) and $(\hat{\alpha}_1, V_1)$ are given by equations (20) and (30) of Koopman, Shephard, and Doornik (1999) respectively, and are computed by the moment and state smoothers using the parameters Ω_η^* and $(\sigma_\epsilon^2)^*$. Equation (13) is a modified version of equation (3.6) of Koopman (1993, p. 123), now including an additional term due to the fact that the covariance of α_1 depends here on the unknown parameters Ω_η. The proof of the EM updating equations is along the lines of that given in Koopman (1993, pp. 125-126) and so is omitted. Indeed, the point of interest here is rather that the chosen initialisation, $\alpha_1 \sim \mathsf{N}(a_1, \Omega_\eta)$, allows the use of this sort of analytical EM step whilst leaving the stationary mean of α_{t+1} and conditional distribution of α_{t+1} given α_t implied by the Gaussian FSN model unchanged. The extension to several, independent trading days is straightforward.

8.3 Details on the construction of the SETS dataset

The algorithm used to perform the reconstruction of the limit order book here correctly takes account of the following features: (a) the existence of limit orders that are still 'active' after the end of the day on which the order was first placed; and (b) the possibility of unexecuted, so-called 'market' orders from the opening auction remaining on the book after the commencement of continuous trading. Note that the TDS dataset for the year 2002 does not provide complete details of limit orders that were entered *prior* to the first day of the dataset analysed in section 4 (i.e. 2 January 2002) but were still active on that date. Such orders only become apparent when they are cancelled or executed against. However, although cancellations prior to 08:00

GMT on 02.01.02 were observed, no further such executions or cancellations occurred during the rest of the entire month, strongly suggesting that no limit orders entered prior to 02.01.02 were still active after 08:00 GMT on that date.

References

Bauwens, L. and P. Giot (2001). *Econometric Modelling of Stock Market Intraday Activity*, Volume 38 of *Advanced Studies in Theoretical and Applied Econometrics*. Boston: Kluwer Academic.

Besse, P. C. and H. Cardot (1996). Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1. *Canadian Journal of Statistics 24*, 467–487.

Besse, P. C., H. Cardot, and D. B. Stephenson (2000). Autoregressive forecasting of some functional climatic variations. *Scandinavian Journal of Statistics 27*, 673–687.

Coppejans, M., I. Domowitz, and A. Madhavan (2003). Resiliency in an automated auction. Mimeo, Department of Economics, Duke University.

Daníelsson, J. and R. Payne (2001). Measuring and explaining liquidity on an electronic limit order book: evidence from Reuters D2002-2. Mimeo, Financial Markets Group, London School of Economics.

Diebold, F. X. and C. Li (2003). Forecasting the term structure of government bond yields. NBER Working Paper W10048.

Doornik, J. A. (2001). *Ox 3.0 - An Object-Oriented Matrix Programming Language*. London: Timberlake Consultants Ltd.

Dunsmuir, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. *The Annals of Statistics 7*, 490–506.

Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics 17*, 425–446.

Engle, R. F. and J. Lange (2001). Measuring, forecasting and explaining time varying liquidity in the stock market. *Journal of Financial Markets 4*, 113–142.

Gouriéroux, C. and J. Jasiak (2001). *Financial Econometrics.* Princeton and Oxford: Princeton University Press.

Gouriéroux, C., G. Le Fol, and B. Meyer (1998). Etude du carnet d'ordres. *Banque et Marchés 36*, 5–20.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Harvey, A. C. and S. J. Koopman (1993). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association 88*, 1228–1236.

Koopman, S. J. (1993). Disturbance smoother for state space models. *Biometrika 80*, 117–126.

Koopman, S. J. and M. Ooms (2001). Time series modelling of daily tax revenues. Tinbergen Institute Discussion Paper, TI2001-032/4.

Koopman, S. J., N. Shephard, and J. A. Doornik (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal 2*, 107–160.

Kotz, S., N. Balakrishnan, and N. L. Johnson (2000). *Continuous Multivariate Distributions.* New York: Wiley.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica 53*, 1315–1335.

London Stock Exchange plc (2003a). *Guide to Trading Services Version 2.0.* London Stock Exchange plc.

London Stock Exchange plc (2003b). *SETS five years on.* London Stock Exchange plc.

Nelson, C. and A. Siegel (1987). Parsimonious modeling of yield curves. *Journal of Business 60*, 473–489.

Poirier, D. J. (1973). Piecewise regression using cubic spline. *Journal of the American Statistical Association 68*, 515–524.

Ramsay, J. O. and B. W. Silverman (1997). *Functional Data Analysis.* New York: Springer.