Centre for Computational Finance and Economic Agents

Working Paper Series

WP029-08

Wing Lon Ng

Modelling Dynamic Demand and Supply Curves of Electronic Markets

September 2008

University of Essex

CCFEA

www.ccfea.net

Modelling Dynamic Demand and Supply Curves of Electronic Markets

Wing Lon Ng
University of Essex
Centre For Computational Finance and Economic Agents (CCFEA)
Wivenhoe Park, Colchester C04 3SQ, Essex, United Kingdom
E-mail: wlng@essex.ac.uk

Abstract

This paper focuses on the dynamics of bid and ask curves of electronic stock markets. A major drawback of most existing models in the literature is that they focus on trade dimension and execution price, only showing the final result of the matching process of buyers and sellers. By contrast, the original bid and ask curves within the electronic market provide a deeper insight into the market's real microstructure. The issue addressed here concerns the entire time-varying order book status that traders are facing before entering the market. Since empirical stock prices and trading quantities have discrete values, the resulting bid and ask curves always represent step functions, that can be approximated via flexible gamma jump processes. Subsequently, a VAR(1) model was applied in order to analyze the dynamics of the market parameters, and a common linear ACD$(1,1)$ model was affixed to recover the temporal structure of the multidimensional stochastic process.

Key Words: Ultra high frequency transaction data, limit order book, gamma processes, ACD, VAR, supply & demand curves.

JEL Classifications: $C22, C32, C41$.

1 Introduction

The econometric investigation of supply and demand curve has been of great interest for a long time. Lacking of empirical data, the first studies were only concentrated on the theoretical aspects of market curves (see Shepherd (1936), Tintner (1938), Maverick (1940) and Ramsey (1972)). Based on the influential work by Scholes (1972), who questions the theory of (nearly) perfect elasticity of market curves as assumed by the hypothesis of efficient, arbitrage-free markets, many following studies find empirical evidence for downward sloping demand curves (see Shleifer (1986), Loderer, Cooney, and Van Drunen (1991), Kaul, Mehrotra, and Morck (2000), Levin and Wright (2001) and Wrugler and Zhuravskaya (2002)) and upward sloping supply curves (see Bagwell (1992), and also Kim, Lee, and Morck (2004), who provide a recent survey of this research field). Generally, the (latent) markets curves, or, their elasticity are estimated by running linear OLS-regressions on daily data of stock prices or index values.

In contrast to this wide stream of literature that is primarily interested in the slope direction of market curves (see also Hausman and Newey (1995)), this paper analyses the dynamic behavior of contemporaneous cross-sectional bid and ask curves by reconstructing the order book and approximating them with flexible jump models. In contrast to Cao, Hansch, and Wang (2004), who truncate the curves after the tenth price step, this study considers the full markets curves. Neglecting the entire order book means disregarding essential information of market activities that is captured in the supply and demand curves. Unlike Bowsher (2004), who only models the five minutes sampled average price by means of cubic splines with fixed knots and does not take the irregular spacing in time into account, this papers applies an ACD model to recover the temporal structure of the inequidistant time intervals between observations

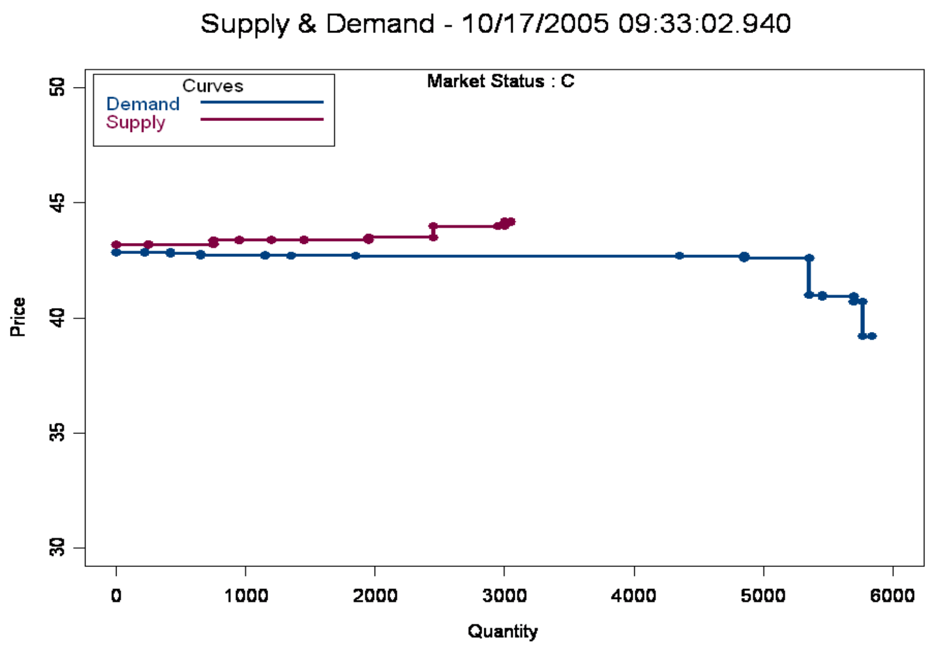

Figure 1: Modelling discrete supply and demand curves with step functions of jump processes

The step functions of both supply and demand curves at each instant of time are modeled via special compound Poisson processes driven by quantity, not by time (see Figure 1). The main objective is to explain the multidimensional transaction process and, thus, the resulting price elasticity implied by the slope of the supply and demand curve. The steps of both curves are captured by a Gamma random variable describing the stochastic consumer or producer surplus of both market sides. In order to study the dynamics of the market curves, this paper extends the bivariate modeling framework proposed by Russell and Engle (2005), applying an VAR framework as an alternative to complex multivariate intensity approaches, which is easier to understand, due to its simple autoregressive structure. Therefore, the final model yields a three-step estimation approach, where (a) the market curves at each instant of time are captured by a so-called Gamma process, (b) whose parameters are dynamically modelled by a VAR model, which is (c) combined with a common ACD model, taking the irregular spacing in time into account. Similar to intensity approaches, this concept has the advantage that no aggregation or censoring over time is needed, since all order arrivals between consecutive trades are considered.

The outline of the paper is structured as follows: Section 2 briefly illustrates the economic motivation. In Section 3, the model will be introduced. Section 4 describes the three-step estimation procedure. In Section 5, the data and results are presented. Section 6 concludes.

2 Economic Motivation

For decades, the literature on financial econometrics and quantitative finance has been used to focus only on the stochastic process of daily prices or returns on assets and their volatility. Generally, most studies only pays little attention to other essential variables in financial markets. In fact, since the "market price" is economically defined as the intersection of the supply and demand, prices are generated by matching the trading willingness of buyers and sellers – not by past (autoregressive) prices. With the appearance of ultra-high frequency data and the increased interest in empirical market microstructure, researchers have paid more attention on the detailed evolution of the "price-generating" transaction process of modern electronic trading systems.

Since the new real-time transaction data arrive at irregular time intervals similar to point processes, Engle and Russell (1997) modelled these time stamped data with the (univariate) *Autoregressive Conditional Duration* (*ACD*) framework. Based on their seminal work, many studies have concentrated on its further development considering additional key variables than the duration (such as price, return, volume, etc.) in order to analyze multivariate settings and describe limit order book activities more accurately. However, a disadvantage of most existing models in the literature is that they are restricted to transaction data of executed orders, which only show the final result of the matching process of buyers and sellers. Therefore, an appropriate econometric model for describing electronic markets should concentrate on the complete supply and demand side of the market, rather than only on the execution price at the intersection.

Of course, the study of bid and ask orders and quotes and their timing is not new. Recently, the literature is especially focusing on the trader's order choice and submission strategy in order to detect specific patterns revealing the order aggressiveness of market participants (see, for example, Ranaldo (2004) or Hall and Hautsch (2006)). Unfortunately, most former studies are only concentrated on the arrival process of incoming orders and differentiate between the market side in order to investigate the order flow of buyers and sellers and, thus, the asymmetry of the market (see, for example, Hedvall and Niemeyer (1997), Hall, Hautsch, and MacCulloch (2003), Obizhaeva and Wang (2006), Hall and Hautsch (2004) and Escribano and Pascual (2006)). Extending these approaches, Ellul, Holden, Jain, and Jennings (2007) model the order submission process with special focus on the market side of the order, the limit type, the execution method and the price aggressiveness.

However, the shape of supply and demand curves, which summarizes the most essential characteristics of the market, was not studied. Although nowadays these (observable) structures can be easily recovered by using detailed order book data, only a few authors have seriously paid attention to the market curves of electronic trading systems (see Cao, Hansch, and Wang (2004) and Bowsher (2004)). Comprehending the general market conditions, under which traders either submit or change or cancel their orders, will lead to a better understanding of the complex trading process. The issue addressed here concerns the time-varying composition of the order book that can be effectively expressed via supply and demand curves. In contrast to transaction data, the original bid and ask curves of the electronic market provide a deeper insight into the market's microstructure. The ultimate goal is explain how market prices and their volatility are influenced by the interaction of a large number of anonymous traders, who arrive at the market at random times, can choose whether to trade immediately or to wait, and can behave strategically by changing their orders at any time.

3 Methodology

Limit orders are price-contingent orders to buy (sell) if the price falls below (rises above) a predetermined price, and the limit order book is the collection of all outstanding limit orders. A sell limit order is also called an ask, whereas a buy limit order is called a bid. The lowest offer is called the best ask price, respectively, the highest bid is called the best bid price. Generally, all market participants can either submit orders in the limit order book and wait until the orders get executed or, alternatively, trade immediately by placing a market order against the best limit order of the opposite market side.

Assume arrival times of orders $t_0, t_1, t_2, ...$ with $t_i \in \mathbb{R}_{\geq 0} \forall i$ as random variables distributed in time by a point process. Their filtration $\mathcal{F}_i$ is defined as

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq ... \subseteq \mathcal{F}_{i-1} \subseteq \mathcal{F}_i = \sigma\left(t_0, t_1, ..., t_{i-1}, t_i\right) \tag{1}$$

with $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Obviously, $\mathcal{F}_i$ is the σ-field generated by all time random variables observed *until* t_i. Now, let $\mathcal{F}_i^*$ denote the extended σ-field containing the marks $\mathbf{M}_i$ of the point process, a vector of all relevant variables,

i. e. the complete information about every incoming orders until time t_i

$$\mathcal{F}_i^* = \sigma\left(\mathcal{F}_i, \mathbf{M}_i, \mathbf{M}_{i-1}, ..., \mathbf{M}_0\right) \quad . \tag{2}$$

With the additional information about the status of every order that is stored in $\mathbf{M}_i$, one can find out, which orders are submitted, still queued, changed, deleted or totally or partially executed at a certain time t_i and, thus, recover the sub-sets of all remaining ask and bid orders *at* time t_i. In other words, these sets represents the time-varying "snapshots" of the order book *at* time t_i. With this information at hand, one is able to recover the status of the order book and to reconstruct the demand and supply curves by means of the order disposition based on the ranks of order limits.

In the following, I will (a) first introduce the Gamma processes for describing the market curves for each instant of time, then (b) model the dynamics of these parameters with a VAR model, and finally (c) combine this framework with a linear ACD model in order to account for the irregular spacing in time.

3.1 The Market Curves

Consider the additive Lévy process $P(t)$ denoting the sum of all inputs in the interval $(0, t)$ with $t \in (0; \infty)$. Under assumption of independence of stationary increments in disjoint time intervals, the sample paths of the non-decreasing Lévy process increase only by jumps. Further, for $\varepsilon > 0$, the number of jumps with size greater than ε in the finite time interval $[0, t]$ has Poisson distribution with finite mean $tv(\varepsilon, \infty)$, where $v(.)$ defines the Lévy measure of the process. Following Basawa and Brockwell (1978), the distribution of $P(t)$ is related to $v(.)$ by

$$E\left(\exp\left(-\lambda P(t)\right)\right) = \exp\left(-t\int_0^\infty \left(1 - e^{-\lambda u}\right) v(du)\right) \quad .$$

Basically, stock prices and trading quantity have discrete values. Since the empirical bid and ask curves always represent a monotonously increasing or decreasing step functions, it seems to be natural to approximate these functions with jump processes. It is assumed that the market curve at time t_i is a realization of a "quasi" Lévy process that is driven by the cumulated quantity τ, not by time t. In this study, I set $\varepsilon = 0.01$ Euro-Cent, which is the smallest possible tick size in most European electronic trading systems, meaning that $v(\varepsilon, \infty)$ is finite. For modeling the price jumps of the supply and demand curve, a Gamma process is suggested, where $v(0, \infty)$ is infinite. That means that there will be, almost surely, infinitely many jumps in each finite time interval, but the number of jumps greater than ε is finite. Since a demand curve definitionally represents a decreasing function, I simply construct a mirrored analogon that has positive jumps. Except the two parameters α and β, all other variables and parameters introduced in this section are assumed to be time-independent (for reason of simplicity, I dropped the subindex i that indicates the i-th arrival time t_i).

First of all, define $U_k(\varepsilon), k = 1, ..., N$ as the sequence of (limit ordered) jumps with size greater than ε. Further, let $N^{Ask}(\tau, \varepsilon)$ and $N^{Bid}(\tau, \varepsilon)$ denote the number of ask and bid order queuing in the order book. The random variables $N(\tau, \varepsilon), U_1(\varepsilon), U_2(\varepsilon), ...$ are independent, where N has Poisson distribution with mean

$$\tau \int_{\varepsilon}^{\infty} \frac{\alpha}{u} \exp(-\beta u)\, du$$

and U follow a Gamma density

$$f(u) = \beta^{\alpha} u^{\alpha-1} \exp(-\beta u) / \Gamma(\alpha)$$

for $u \geq \varepsilon$ (Basawa and Brockwell (1978)). Economically, these jumps describe the relative consumer and producer surplus of market participants. Trading in electronic markets is anonymous. Hence, the assumption of Poisson "arrival" of the price jumps is reasonable, because I assume that all (contemporaneous) trading quantities at t_i are are independent. Suppose that both supply and demand curve are built up by successive aggregation of all N (observed) jumps, then the respective price function $P(\tau)$ determining the market curve is a realization of a Gamma process. The price function dependent on τ represents the aggregation over the entire consumer or producer surplus of all active market participants with respect on their planned trading quantity and has the probability density

$$f(p(\tau)) = \beta^{\alpha} p^{\alpha\tau-1} \exp(-\beta p) / \Gamma(\alpha\tau) \tag{3}$$

and Lévy measure

$$v(du) = \frac{\alpha}{u} \exp(-\beta u)\, du$$

with $\alpha, \beta > 0$ and $\Gamma(.)$ denoting the Gamma function

$$\Gamma(c) = \int_0^{\infty} x^{c-1} e^{-x} dx \qquad . \tag{4}$$

The parameter α controls the shape and β the scale of the distribution and, thus, the jumps and the slope of the market curve. The density in equation (3) represents a convolution and results from the fact that

$$\sum_{k=1}^{N} U_k \rightarrow P(\tau) \text{ as } \varepsilon \rightarrow 0 \qquad . \tag{5}$$

3.2 The Dynamics

Now, it is assumed that the parameters of the bid and ask curves at time t_i, denoted as α_i and β_i, are realizations of the stochastic process $A(t_i) \equiv A_i$ and $B(t_i) \equiv B_i$, both controlling the dynamic behavior of the market curves (the superindices of the variables in eq. (6) indicate the respective market side of the curve). Since this paper focus on the relationships between both curves, I apply a VAR model in order to study the multivariate intertemporal dependence structure of all variables market parameters characterizing the supply and demand. I restrict the attention to a lag order of one in order to avoid computational burdens, since the number of parameters will grow quadratically with every additional explanatory variable. Supposing that the intercept of both demand and supply curve have no impact on the trader's submission strategy, I only consider the spread $Spr\,(t_i)$ (defined as the difference between the lowest ask limit and the highest bid limit) in the calculations, yielding

$$\underbrace{\begin{pmatrix} Spr_i \\ A_i^{Ask} \\ B_i^{Ask} \\ A_i^{Bid} \\ B_i^{Bid} \end{pmatrix}}_{\mathbf{Y}_i} = \boldsymbol{\delta} \underbrace{\begin{pmatrix} Spr_{i-1} \\ A_{i-1}^{Ask} \\ B_{i-1}^{Ask} \\ A_{i-1}^{Bid} \\ B_{i-1}^{Bid} \end{pmatrix}}_{\mathbf{Y}_{i-1}} + \mathbf{Z}_i \tag{6}$$

with $\mathbf{Y}$ as the (mean-adjusted) the vector of all variables, $\boldsymbol{\delta}$ as the 5×5 coefficient matrix and $\mathbf{Z} \sim \mathbf{N}\,(\mathbf{0}, \boldsymbol{\Sigma})$ as the multivariate normally distributed residual vector of the VAR(1) model.

3.3 The Irregular Spacing in Time

Since market event occur asynchronously, the common time series models with equidistant spacing are not appropriate for capturing real-time market activities. That means that researchers must not only consider the markets variables themselves (S, A, B), but also their stochastic time difference. Generally, a duration framework is the common approach to deal with point processes. Let

$$X_i = t_i - t_{i-1} \tag{7}$$

denote the i^{th} order duration between the i^{th} and $(i-1)^{th}$ incoming order. Furthermore, define $\Psi_i \equiv E\left(X_i|\mathcal{F}_{i-1}^{*}\right)$, that is the conditional expectation of duration given all information available at t_{i-1}. In the original ACD model proposed by Engle and Russell (1997), the conditional mean is written as a linear function of past durations

$$\Psi_i \equiv E\left(X_i|\mathcal{F}^*_{i-1}\right) = \omega + \sum_{j=1}^{p} a_j X_{i-j} + \sum_{k=1}^{q} b_k \Psi_{i-k} \quad , \tag{8}$$

whereby the parameters must satisfy the following constraints

$$\omega > 0$$

$$a_j, b_k \geq 0 \qquad \forall j, k \tag{9}$$

$$\sum_{j=1}^{p} a_j + \sum_{k=1}^{q} b_k < 1 \qquad , \tag{10}$$

to ensure stationarity of the process, similar to GARCH models. It is now assumed that

$$X_i = \Psi_i \cdot \varepsilon_i \tag{11}$$

with $\varepsilon_i \equiv \frac{X_i}{\Psi_i}$ as i.i.d.-innovations. Their density function $f(.)$ has normalization $E(\varepsilon_i) = 1$ by construction and a non-negative support to avoid negative durations. Engle and Russell (1997) used an Exponential and a Weibull distribution, whereas other authors favoured more flexible alternatives (see, for example Bauwens, Giot, Grammig, and Veredas (2004)). In this study, the generalized Gamma distribution with the density

$$f(x_i) = \frac{\gamma}{x_i \cdot \Gamma(\lambda)} \left(\frac{x_i}{\Psi_i} \cdot \frac{\Gamma\left(\lambda + \frac{1}{\gamma}\right)}{\Gamma(\lambda)} \right)^{\gamma\lambda} \cdot \exp\left(- \left(\frac{x_i}{\Psi_i} \cdot \frac{\Gamma\left(\lambda + \frac{1}{\gamma}\right)}{\Gamma(\lambda)} \right)^{\gamma} \right) \tag{12}$$

is chosen to allow for extreme durations (see Kleiber and Kotz (2003)).

Since the assumption of linearity is often too restrictive to capture the duration process (see equation (8)), several modified models for transaction data have been developed with other dependence structures of the conditional mean to account for nonlinear impacts. Generally, eq. (8) can be replaced by different types of ACD models by varying the functional form

$g(.)$ of Ψ_i in the model's equation

$$X_i = g(\Psi_i) \cdot \varepsilon_i$$

(for a survey, see Hautsch (2004), Bauwens, Giot, Grammig, and Veredas (2004), Fernandes and Grammig (2006) or Bauwens and Hautsch (2006)).

3.4 The ACD-VAR Model

To describe the parameters of the market curve and the order duration jointly, it is necessary to combine the *ACD* model with the *VAR* model, similarly to the *ACM-ACD* approach proposed by Russell and Engle (2005). The aim is to model the stochastics of the duration process of all time-stamped orders in the order book, as well as the imbedded dynamics of the market curves. Assuming that the joint conditional distribution of the order duration X_i and the vector of all market parameters $\mathbf{Y}_i$ are influenced by the common filtration $\mathcal{F}^*$, the joint conditional density of $(X_i, \mathbf{Y}_i)_{i\in\mathbb{N}}$ can be decomposed into

$$f_{X_i,\mathbf{Y}_i|\mathcal{F}^*_{i-1}}(x_i, y_i) = \underbrace{f_{X_i|\mathcal{F}^*_{i-1}}(x_i)}_{f_{ACD}} \cdot \underbrace{f_{\mathbf{Y}_i|X_i,\mathcal{F}^*_{i-1}}(y_i)}_{f_{VAR}} \tag{13}$$

as discussed in Engle (2000). This concept has also been adapted in former decomposition models to study price movements (see, for example, Rydberg and Shephard (2003) or Liesenfeld, Nolte, and Pohlmeier (2005)). Extending eq. (8) and eq. (6), it is assumed in this study that

$$\begin{aligned} \Psi^*_i &= E\left(X_i|\mathcal{F}^*_{i-1}\right) \\ &= \sum_{j=1}^{p} \alpha_j X_{i-j} + \sum_{k=1}^{q} \beta_k \Psi^*_{i-k} + \boldsymbol{\kappa}\mathbf{Y}_{i-1} \end{aligned} \tag{14}$$

and

$$\underbrace{\begin{pmatrix} \Psi_i \\ \mathbf{Y}_i \end{pmatrix}}_{\mathbf{Y}^*_i} = \boldsymbol{\delta} \underbrace{\begin{pmatrix} \Psi_{i-1} \\ \mathbf{Y}_{i-1} \end{pmatrix}}_{\mathbf{Y}^*_{i-1}} + \mathbf{Z}_i \tag{15}$$

This composed framework should not only reveal the relationship between the shape of the limit order book and order duration, but also their joint dynamics, i. e. the evolution in time of the book. The dependence between X_i and $\mathbf{Y}_i$ is captured by both (14) and (15), each containing and influencing the information for the other process, and both adapted to $\mathcal{F}^*_{i-1}$. The history of the order book is embedded in both components.

4 Estimation

It is well-known that intraday data have a consistent diurnal pattern of trading activities over the course of a trading day, due to certain institutional characteristics of organized financial markets, such as opening and closing hours or intraday auctions. Since it is necessary to take the daily deterministic seasonality into account, smoothing techniques are required to get deseasonalized observations. Let $\tilde{X}_i$ denote the observed duration. Instead of applying cubic splines, where the positions of the nodes have to be carefully set, a kernel regression with an asymptotic optimal bandwidth $h = \hat{\sigma}_{\tilde{X}} \left(\frac{4}{3n}\right)^{\frac{1}{5}}$ was performed (where $\hat{\sigma}_{\tilde{X}}$ denotes the standard deviation of the raw order durations). Using the Gaussian density as kernel function $K(.)$, the diurnal periodic component (dependent on daytime d) can be computed by the Nadaraya-Watson estimator

$$NW(d_i) = \frac{\sum_{i=1}^{n} \tilde{X}_i \cdot K\left(\frac{d-d_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{d-d_i}{h}\right)} \qquad . \tag{16}$$

Thus, $X_i \equiv \frac{\tilde{X}_i}{NW(d_i)}$ is the deseasonalized duration and should have no diurnal pattern and a unit mean.

Now, according to the previous section, the subsequent estimation procedure of the final model consists of three steps: Firstly, all Gamma processes of both market curves of each observed time point are estimated separately.

Having their computed market parameters characterizing the shape of the curve, the $VAR(1)$ model is estimated in the second step in order to get the starting values for the final joint ACD-VAR model. All models are estimated with the maximum likelihood method.

Following Basawa and Brockwell (1978), the parameters α and β are obtained by maximizing

$$L(\alpha,\beta) = \exp\left(-t\int_{\varepsilon}^{\infty} v(du)\right)\left(\prod_{k=1}^{N}\frac{\alpha}{u_k}\exp(-\beta u_k)\right) \quad ,$$

yielding

$$\int_0^{\infty}\frac{\hat{\beta}\exp\left(-\hat{\beta}u\right)}{u+\varepsilon}du = \frac{N}{\sum_{k=1}^{N}U_k}$$

with

$$\hat{\beta} \to \left(-\log(\varepsilon)\sum_{k=1}^{N}\frac{U_k}{N}\right)^{-1}$$

as $\varepsilon \to 0$ and

$$\hat{\alpha} = \frac{\hat{\beta}}{\tau}\exp\left(\hat{\beta}\varepsilon\right)\sum_{k=1}^{N}U_k \quad .$$

For $\tau \to \infty$, the estimators have limiting Normal distribution

$$\sqrt{t}\begin{bmatrix}\hat{\alpha}-\alpha \\ \hat{\beta}-\beta\end{bmatrix} \to N\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\sigma_{\alpha}^2 & \sigma_{\alpha\beta}\\ \sigma_{\alpha\beta} & \sigma_{\beta}^2\end{bmatrix}\right)$$

with

$$\begin{aligned}
\sigma_{\alpha}^2 &= \frac{\alpha(1+\beta\varepsilon)}{\frac{(1+\beta\varepsilon)}{\alpha}\left(\int_{\varepsilon}^{\infty}\frac{\alpha}{u}\exp(-\beta u)\,du\right)-\exp(-\beta\varepsilon)}\\
\sigma_{\beta}^2 &= \frac{\int_{\varepsilon}^{\infty}\frac{\alpha}{u}\exp(-\beta u)\,du}{\alpha^2\frac{\exp(-\beta\varepsilon)}{\beta^2}\left(\frac{(1+\beta\varepsilon)}{\alpha}\left(\int_{\varepsilon}^{\infty}\frac{\alpha}{u}\exp(-\beta u)\,du\right)-\exp(-\beta\varepsilon)\right)}\\
\sigma_{\alpha\beta} &= \frac{-1}{\frac{1}{\beta}\left(\frac{(1+\beta\varepsilon)}{\alpha}\left(\int_{\varepsilon}^{\infty}\frac{\alpha}{u}\exp(-\beta u)\,du\right)-\exp(-\beta\varepsilon)\right)} \quad .
\end{aligned}$$

In order to save a large amount of computation time, good starting values for the optimization procedure are important. According to *Anderson and*

Ray (1975), the slope parameter α of a Gamma distribution should be set to

$$\hat{\alpha}^* = \frac{n-3}{n}\hat{\alpha} + \frac{2}{3n}$$

to account for small sample bias, whereby

$$\hat{\alpha} = \frac{1}{4R_n}\left(1 + \sqrt{1 + \frac{4}{3}R_n}\right)$$

is the estimator suggested by *Thom* (1968) and

$$R_n = \ln\left(\frac{\frac{1}{n}\sum_{i=1}^n u_i}{\sqrt[n]{\prod_{i=1}^n u_i}}\right)$$

(see also Johnson, Kotz, and Balakrishnan (1994)). The starting value for the scale parameter β can be obtained by the common moment estimator

$$\hat{\beta} = \frac{\hat{\alpha}^*}{\frac{1}{n}\sum_{i=1}^n u_i} \qquad .$$

Considering (14) and (15) and satisfying the restrictions (9) and (10) , the final model is estimated by jointly maximizing the two partial log-likelihood functions of the general model

$$\mathcal{L} = \sum_{i=1}^n \underbrace{\ln\left(f_{ACD}\left(x_i|\mathcal{F}_{i-1}^*\right)\right)}_{\mathcal{L}_{ACD}} + \underbrace{\ln\left(f_{VAR}\left(\mathbf{y}_i|x_i, \mathcal{F}_{i-1}^*\right)\right)}_{\mathcal{L}_{VAR}}$$

(see equation (13)). The log-likelihood function of the VAR part is given by

$$\mathcal{L}_{VAR} = -\frac{6n}{2}\ln(2\pi) - \frac{n}{2}\ln|\mathbf{\Sigma}| - \sum_{i=1}^n \mathbf{z}_i'\mathbf{\Sigma}^{-1}\mathbf{z}_i$$

with

$$\mathbf{z}_i = \mathbf{y}_i^* - \boldsymbol{\delta}\mathbf{y}_{i-1}^*$$

(see Lütkepohl (2005)). The likelihood function of the ACD part with a generalized Gamma distribution for the innovations ε_i is

$$L_{ACD} = \exp\left(-\sum_{i=1}^n\left(\frac{x_i}{\Psi_i^*}\cdot\frac{\Gamma\left(\lambda+\frac{1}{\gamma}\right)}{\Gamma(\lambda)}\right)^{\gamma}\right)\prod_{i=1}^n\frac{\gamma}{x_i\cdot\Gamma(\lambda)}\left(\frac{x_i}{\Psi_i^*}\cdot\frac{\Gamma\left(\lambda+\frac{1}{\gamma}\right)}{\Gamma(\lambda)}\right)^{\gamma\lambda}$$

(see also Lunde (2000)). Taking the logarithm, one gets

$$\begin{aligned}\mathcal{L}_{ACD} &= N\ln\left(\frac{\gamma}{\Gamma(\lambda)}\right) - \left[\sum_{i=1}^{N}\ln(x_i)\right] + \gamma\lambda\left[\sum_{i=1}^{N}\ln\left(\frac{x_i}{\Psi_i^*}\right)\right] \\ &\quad + N\gamma\lambda\ln\left(\frac{\Gamma\left(\lambda+\frac{1}{\gamma}\right)}{\Gamma(\lambda)}\right) - \left(\frac{\Gamma\left(\lambda+\frac{1}{\gamma}\right)}{\Gamma(\lambda)}\right)^{\gamma}\cdot\left[\sum_{i=1}^{N}\left(\frac{x_i}{\Psi_i^*}\right)^{\gamma}\right] \quad .\end{aligned}$$

5 Empirical Results

The detailed order book data is extracted from the open order book of the German XETRA system, which is an order-driven market. Trading in the XETRA system is continuous during the opening hours and is based on the so-called continuous double auction mechanism. A computer keeps track of all submitted orders and order changes. The matching of supply and demand is automatically performed, generally based on the usual algorithms following a strict price-time order priority. This study only considers the continuous trading phase, where the order book is open and visible for all registered market participants. It starts after the opening auction at 9 a.m., where the opening price is determined as the price which maximizes the volume that can be traded, and ends at 6 p.m. with the closing auction.

I consider two stocks traded on the XETRA trading system, namely *Pfeiffer Vacuum* (traded in the TecDAX segment) and *Rheinmetall AG* (traded in the MDAX segment). The observation period ranges from 15^{th} October 2005 to 14^{th} December 2005. Because the data set contains the full history of all events recorded in the order book (limit orders, market orders, market-to-limit, iceberg orders, fill or kill, cancellations, changes, full/partial executions, market status) and their matching outcomes (with an accuracy as great as one hundredth of a second), it represents the close chronicle of the whole trading process during the two months. With this provided information, I reconstruct the complete status of the order book at

	Pfeiffer Vacuum	Rheinmetall AG
Sample	46212	156865
Min.	0.0100	0.0100
1^{st} Qu.	1.1550	0.8600
Median	7.8300	3.0900
Mean	54.0153	16.8053
3^{rd} Qu.	45.6800	14.5500
Max.	3201.5400	1029.3800
Std.dev.	139.0206	39.5591

Table 1: Description of Order Durations (in seconds)

each point of time, corresponding to the information available to the market participant preparing to trade at a certain time.

Descriptive statistics of the order durations are listed in Table 1. Their kernel densities and diurnalities are displayed in Figure 2.

The results of the ACD and the VAR models are reported in Table 2 and 3. The price process of the two stocks and the estimated parameters of all Gamma processes for describing the market curves are visualized in Figure 4 to 7. Generally, the estimated model parameters are quite similar and imply a Gamma density with a exponentially decreasing slope, as discernible in Figure 3. This means that most market participants only accept small price jumps, because they have little consumer or producer surplus which would tend to result in market curves with only moderate price elasticity. Conversely, increasing scale and shape parameters mean higher densities for large jumps in the price function and, thus, imply fast rising markets curves with a more arduous slope.

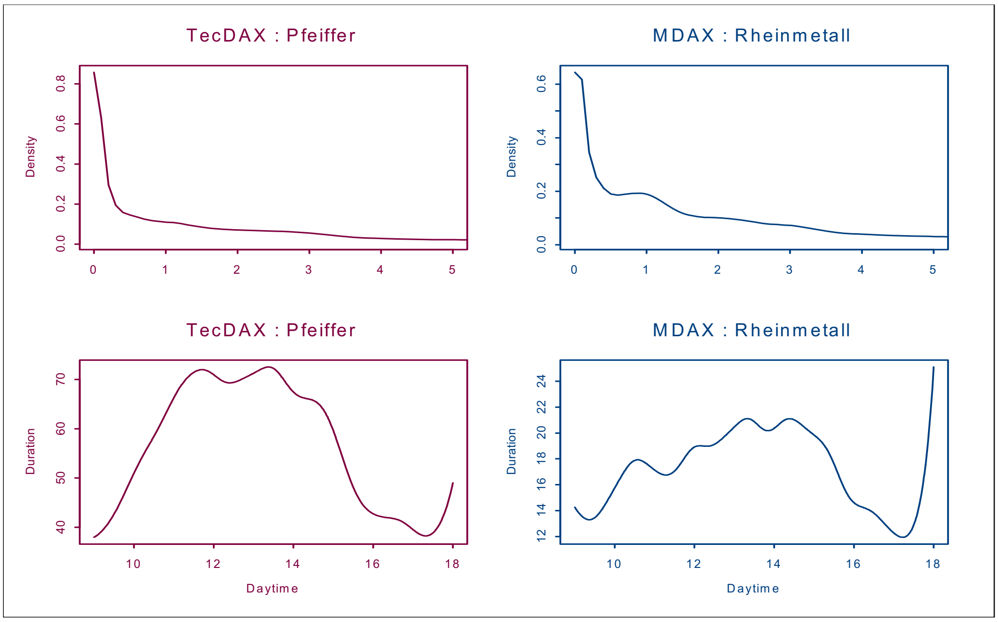

Figure 2: Kernel densities and diurnal patterns of order durations

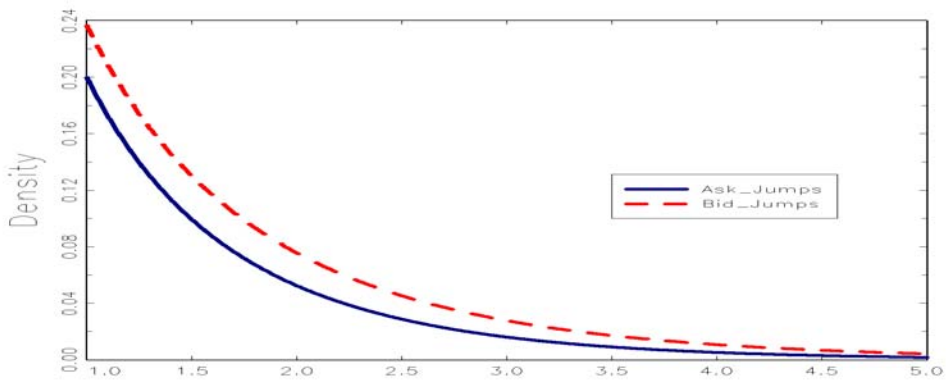

Figure 3: Example: Density of the jumps in the supply and demand curves, based on the mean values of the estimated parameters for Pfeiffer Vacuum.

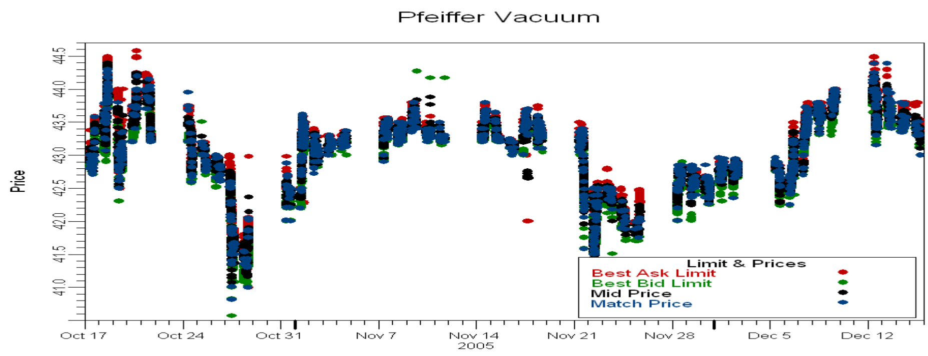

Figure 4: The Price Process of Pfeiffer Vacuum

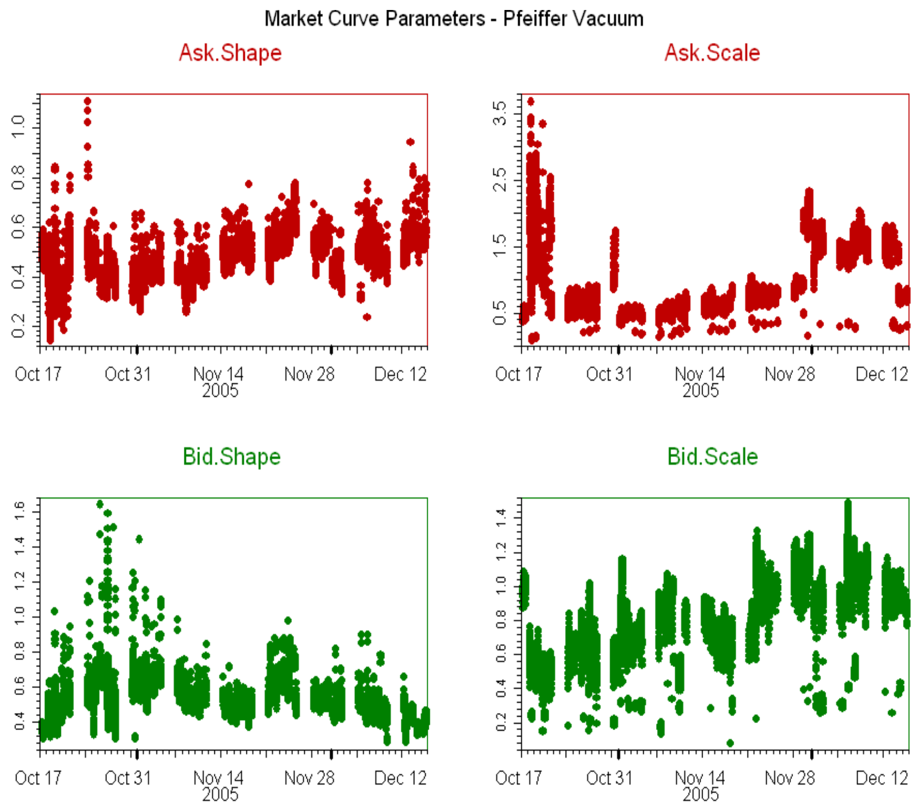

Figure 5: The (irregularly spaced) stochastic process of the Gamma Process parameters for Pfeiffer Vacuum

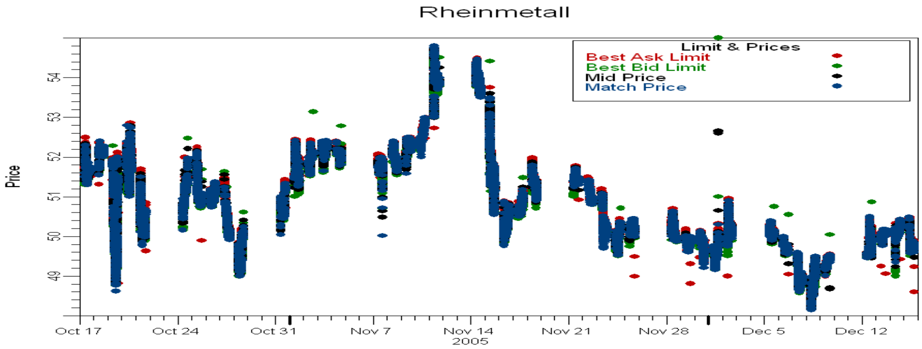

Figure 6: The Price Process of Rheinmetall AG

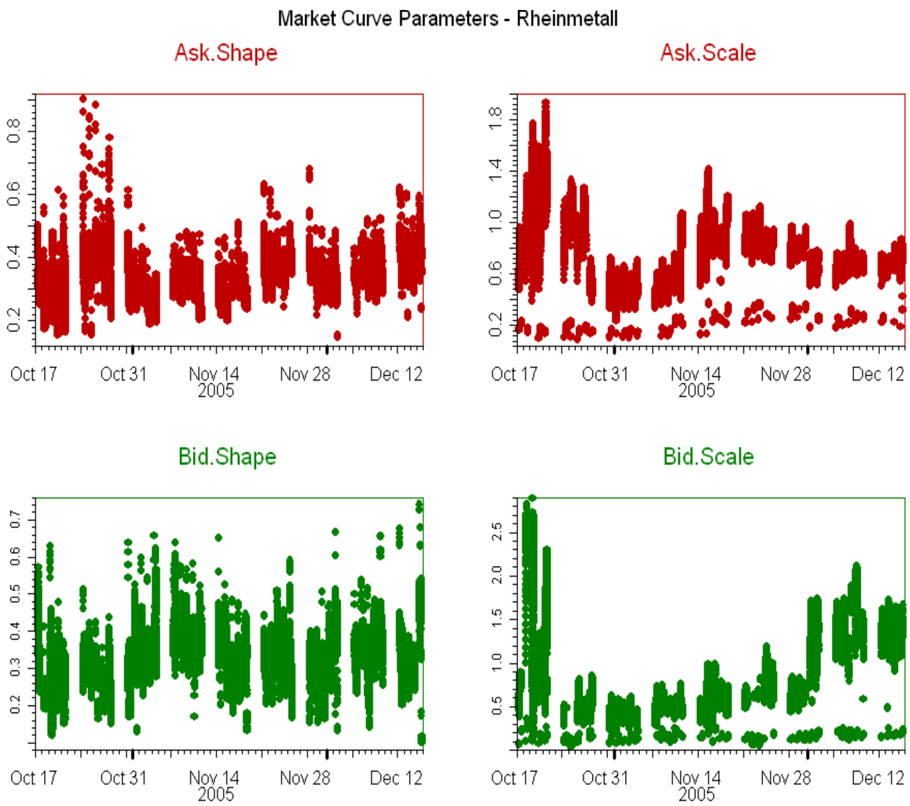

Figure 7: The (irregularly spaced) stochastic process of the Gamma Process parameters for Rheinmetall AG.

(1) ACD part:

	Duration Process			Covariates	
ω	intercept	0.0745 (0.0030)	κ_1	Spread	-0.0616 (0.0023)
α	last duration	0.2862 (0.0170)	κ_2	Ask Shape	-0.0260 (0.0069)
β	last cond. dur.	0.7158 (0.0128)	κ_3	Ask Scale	-0.0001 (0.0026)
γ	G.Gamma shape1	0.3410 (0.0117)	κ_4	Bid Shape	0.0137 (0.0030)
λ	G.Gamma shape2	1.7910 (0.1062)	κ_5	Bid Scale	0.0103 (0.0057)

(2) VAR part:

$\boldsymbol{\delta}$	Ψ_i	Spr_i	A_i^{Ask}	B_i^{Ask}	A_i^{Bid}	B_i^{Bid}
Ψ_{i-1}	0.8357 (0.5480)	-0.0002 (0.0359)	-0.0003 (0.0149)	0.0001 (0.0510)	-0.0004 (0.0228)	0.0001 (0.0245)
Spr_{i-1}	-0.1847 (5.0743)	0.9433 (0.3324)	-0.0013 (0.1382)	-0.0001 (0.4723)	0.0008 (0.2115)	0.0013 (0.2272)
A_{i-1}^{Ask}	0.1747 (8.9629)	0.0108 (0.5871)	0.9699 (0.2441)	-0.0234 (0.8343)	-0.0031 (0.3737)	0.0161 (0.4013)
B_{i-1}^{Ask}	0.0054 (1.4219)	0.0021 (0.0931)	-0.0003 (0.0387)	0.9914 (0.1323)	-0.0013 (0.0592)	-0.0003 (0.0636)
A_{i-1}^{Bid}	0.1082 (6.7196)	0.0067 (0.4402)	-0.0007 (0.1830)	-0.014 (0.6255)	0.9602 (0.2801)	-0.0016 (0.3008)
B_{i-1}^{Bid}	-0.0201 (3.4322)	-0.0110 (0.2248)	0.0047 (0.0935)	0.0033 (0.3194)	-0.0018 (0.1431)	0.9882 (0.1536)

Table 2: Estimation results for Pfeiffer Vacuum

(1) ACD part:

	Duration Process			Covariates	
ω	intercept	0.1189 (0.0046)	κ_1	Spread	−0.0797 (0.0034)
α	last duration	0.3518 (0.0096)	κ_2	Ask Shape	0.0010 (0.0113)
β	last cond. dur.	0.6283 (0.0081)	κ_3	Ask Scale	−0.0341 (0.0029)
γ	G.Gamma shape1	0.3203 (0.0059)	κ_4	Bid Shape	0.0078 (0.0129)
λ	G.Gamma shape2	2.8584 (0.0973)	κ_5	Bid Scale	0.0013 (0.0022)

(2) VAR part:

$\boldsymbol{\delta}$	Ψ_i	Spr_i	A_i^{Ask}	B_i^{Ask}	A_i^{Bid}	B_i^{Bid}
Ψ_{i-1}	0.7941 (0.6097)	−0.0005 (0.0249)	0.0003 (0.0094)	−0.0004 (0.0211)	0.0002 (0.0083)	−0.0004 (0.0346)
Spr_{i-1}	−0.3426 (7.9725)	0.9456 (0.3259)	−0.0001 (0.1232)	0.0006 (0.2770)	−0.0006 (0.1088)	0.0008 (0.4532)
A_{i-1}^{Ask}	−0.0191 (12.2169)	0.0027 (0.4995)	0.9820 (0.1888)	0.0011 (0.4245)	−0.0031 (0.1668)	0.0072 (0.6944)
B_{i-1}^{Ask}	−0.1267 (3.0662)	−0.0021 (0.1253)	−0.0005 (0.0473)	0.9950 (0.1065)	−0.0009 (0.0418)	−0.0015 (0.1743)
A_{i-1}^{Bid}	0.0595 (11.2377)	0.0044 (0.4594)	0.0003 (0.1737)	0.0010 (0.3905)	0.9883 (0.1534)	0.0003 (0.6388)
B_{i-1}^{Bid}	0.0061 (1.4270)	0.0005 (0.0583)	0.0002 (0.0221)	−0.0004 (0.0495)	0.0001 (0.0194)	0.9967 (0.0811)

Table 3: Estimation results for Rheinmetall

In this study, the ACD part of the model detects a strong cluster structure of order durations, signaling a specific behavioral pattern of traders. According to the results of other former studies, long durations tend to be followed by long ones and short durations by short ones, $\hat{a}_1, \hat{b}_1 > 0$. This finding is in line with the information-based market microstructure theory, where the trading and ordering process represents a source of information (see O'Hara (1997)). Uninformed market participants must participate in order to update their pool of information, whereas informed traders only enter the market when they have private information. Long durations suggests that uninformed traders still believe that the underlying value of the asset has not changed and only trade in order to optimize their own portfolio or due to demand for liquidity. In contrast, short durations and, hence, fast and intensive trading signalize the presence of asymmetric information, where insiders are assumed to make money by using their informational advantage. Thus, an understanding of the time-varying speed of transactions is important, in order to determine when to enter the trading platform to demand or supply liquidity.

Interestingly, similar to the spread that decreases the order duration (see O'Hara (1997)), the ask curve parameters reduce the conditional mean duration, whereas the bid curve parameters increase it. These results also confirm the hypothesis in market microstructure theory, where rising prices – in this case *expected* rising prices as expressed in the fast increasing ask curve – are concerned as "good news" attracting insiders who used to trade immediately and, thus, short durations. In contrast, *expected* falling prices as exhibited in the rapidly decreasing bid curve signalize "bad news", where only noise trader stay in market to trade (long durations). The estimates of the VAR model reveal that the market parameters follow a strong autoregressive structure without "cross-over" effects. This signalizes that both supply and demand curves depend on their own past history, but not on the dynamics of the contrary market side.

6 Conclusion

This paper develops a multivariate modeling framework for analyzing the dynamics of supply and demand curves in electronic markets generated by the arrival process of ask and bid orders in an electronic order book market. The detailed ultra-high frequency XETRA data set does not only shows the price, volume and time stamp of the transaction , but also the status of the submitted order, allowing the reconstruction of all recorded transactions. Based on this detailed data, it is possible to rebuild all order book compositions observed in the irregular time intervals. Since stock prices and trading quantity have discrete values in reality, the resulting empirical bid and ask curves represent a step functions, that can be approximated via flexible Gamma processes that allow non-integer jumps. Subsequently, a VAR(1) model was applied in order to analyze the dynamics of the market parameters, and a common linear $\text{ACD}(1,1)$ model was affixed to recover the temporal structure of the general stochastic process, reflecting the information flow and market activity.

The empirical results show that fast increasing ask curves accelerate order submission process, whereas rapidly decreasing bid curves decelerate it. Further, the market parameters follow a strong autoregressive structure without "cross-over" effects. This indicates that both supply and demand curves depend on their own past history, but not on the dynamics of the contrary market side.

References

Bagwell, L. S. (1992): "Dutch Auction Repurchases: An Analysis of Share Holder Heterogenity," *The Journal of Finance*, 47(1), 71–105.

Basawa, I. V., and P. J. Brockwell (1978): "Inference for gamma and stable processes," *Biometrika*, 67(1), 129–133.

Bauwens, L., P. Giot, J. Grammig, and D. Veredas (2004): "A Comparison of Financial Duration Models Via Density Forecasts," *International Journal of Forecasting*, 20(4), 589–609.

Bauwens, L., and N. Hautsch (2006): "Stochastic Conditional Intensity Process," *Journal of Financial Econometrics*, 119(2), 450–493.

Bowsher, C. (2004): "Modelling the Dynamics of Cross-Sectional Price Functions: Econometric Analysis of the Bid and Ask Curves of an Automated Exchange," Nuffield college economics discussion paper 2004-w21, University of Oxford.

Cao, C., O. Hansch, and X. Wang (2004): "The Informational Content of an Open Limit Order Book," Working paper, Department of Finance, Pennsylvania State University.

Ellul, A., C. W. Holden, P. Jain, and R. Jennings (2007): "Order Dynamics: Recent Evidence from the NYSE," Working paper, Kelley School of Business, Indiana University.

Engle, R. F. (2000): "The Econometrics of Ultra-High Frequency Data," *Econometrica*, 68(1), 1–22.

Engle, R. F., and J. R. Russell (1997): "Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model," *Journal of Empirical Finance*, 4(2-3), 187–212.

Escribano, A., and R. Pascual (2006): “Asymmetries in bid and ask resposes to innovations in the trading process,” *Empirical Economics*, 30(4), 913–946.

Fernandes, M., and J. Grammig (2006): “A Family of Autoregressive Conditional Durations Models,” *Journal of Econometrics*, 130(1), 1–23.

Hall, A., and N. Hautsch (2004): “A Continuous-Time Measurement of the Buy-Sell Pressure in a Limit Order Book Market,” Working paper, University of Technology, Sydney and University of Copenhagen.

——— (2006): “Order Aggressiveness and Order Book Dynamics,” *Empirical Economics*, 30(4), 973–1005.

Hall, A., N. Hautsch, and J. MacCulloch (2003): “Estimating the Intensity of Buy and Sell Arrivals in a Limit Order Book Market,” Working paper, CoFE & UTS.

Hausman, J. A., and W. K. Newey (1995): “Nonparametric Estimation of Exact Consumer Surplus and Deadweight Loss,” *Econometrica*, 63(6), 1445–1476.

Hautsch, N. (2004): *Modelling Irregularly Spaced Financial Data*. Springer, Heidelberg.

Hedvall, K., and J. R. Niemeyer (1997): “Do buyers and Sellers behave similarly in a limit orderbook? A high-frequency data examination of the Finnish stock exchange,” *Journal of Empirical Finance*, 4(2-3), 279–293.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994): *Continuous Univariate Distributions*, vol. 1. John Wiley & Sons, New York et al., 2nd edn.

Kaul, A., V. Mehrotra, and R. Morck (2000): "Demand Curves for Stocks Do Slope Down: New Evidence from an Index Weigths Adjustment," *The Journal of Finance*, 55(2), 893–912.

Kim, J.-W., J. Lee, and R. Morck (2004): "Heterogeneous Investors and their Changing Demand and Supply Schedules for Individual Common stocks," Discussion paper no. 2031, Harvard Institute of Economic Research.

Kleiber, C., and S. Kotz (2003): *Statistical Size Distributions in Economics and Actuarial Sciences*. John Wiley & Sons, New York et al.

Levin, E. J., and R. E. Wright (2001): "Downwards Sloping Demand Curves for Stock?," Working paper, University of Stirling.

Liesenfeld, R., I. Nolte, and W. Pohlmeier (2005): "Modelling Financial Transaction Price Movements: A Dynamic Integer Count Data Model," *Empirical Economics*, 30(4), 795–825.

Loderer, C., J. W. Cooney, and L. D. Van Drunen (1991): "The Price Elasticity of Demand for Common Stock," *The Journal of Finance*, 46(2), 621–651.

Lunde, A. (2000): "A Generalized Gamma Autoregressive Conditional Duration Model," Discussion paper, Aalborg University.

Lütkepohl, H. (2005): *New Introduction to Multiple Time Series Analysis*. Springer, Berlin et al.

Maverick, L. A. (1940): "Demand and Supply Curves," *The Quaterly Journal of Economics*, 54(2), 307–313.

Obizhaeva, A., and J. Wang (2006): "Optimal Trading Strategy and Supply/Demand Dynamics," Working paper, MIT Sloan School of Management.

O'Hara, M. (1997): *Market Microstructure Theory.* Blackwell, Oxford.

Ramsey, J. B. (1972): "Limiting Functional Forms for Market Demand Curves," *Econometrica*, 40(2), 327–341.

Ranaldo, A. (2004): "Order aggressiveness in limit order book markets," *Journal of Financial Markets*, 7(1), 53–74.

Russell, J. R., and R. F. Engle (2005): "A Discrete-State Continuous-Time Model of Financial Transactions Prices and Times: The Autoregressive Conditional Multinomial-Autoregressive Conditional Duration Model," *Journal of Business & Economic Statistics*, 23(2), 166–180.

Rydberg, T. H., and N. Shephard (2003): "Dynamics of trade-by-trade price movements: decomposition and models," *Journal of Financial Econometrics*, 1(1), 2–25.

Scholes, M. S. (1972): "The Market for Securities: Substitution versus Price Pressure and the Effects of Information on Share Prices," *Journal of Business*, 42(2), 179–211.

Shepherd, G. (1936): "Vertical and Horizontal Shifts in Demand Curves," *Econometrica*, 4(4), 361–367.

Shleifer, A. (1986): "Do Demand Curves for Stocks Slope Down?," *The Journal of Finance*, 41(3), 579–590.

Tintner, G. (1938): "The Theoretical Derivation of Dynamic Demand Curves," *Econometrica*, 6(4), 375–380.

Wrugler, J., and E. Zhuravskaya (2002): "Does Arbitrage Flatten Demand Curves for Stocks," *Journal of Business*, 75(4), 583–608.