

A modelling framework for the prices and times of trades made on the New York stock exchange

Tina Hviid Rydberg and Neil Shephard

Nuffield College, Oxford OX1 1NF, UK

Abstract

In this chapter we propose using compound Poisson processes to model trade-by-trade financial data. Our main focus will be on developing specific types of Cox processes in order to accurately depict the trading process. We study the problem of signal extracting the intensity of the trading process. We finish by studying the implication for price changes over pre-specified intervals of times, such as 30 seconds, 20 minutes or a day and assessing the empirical plausibility of OU based models for the intensity of the trading process..

Some keywords: Cox process; Durations; Kalman filter; Intensity; Ornstein-Uhlenbeck processes; Particle filter; Trade-by-trade dynamics.

1 Introduction

1.1 The data and model

Most modern theoretical and empirical finance is based on continuous time models with continuous sample paths or, in other words, diffusion processes which are driven by Wiener processes. Prominent recent references include Duffie (1992) and Ait-Sahalia (1996), while the most well known example is geometric Brownian motion used in the Black and Scholes (1973) option pricing model.

In practice almost all the prices at which financial assets transact live on a lattice structure. Figure 1 displays the transaction prices (in US Dollars) for the IBM stock traded on the New York Stock Exchange (NYSE) on four randomly selected days in 1995. This shows prices are integer multiples of $1/8$ and that transactions are irregularly spaced in time. Although this multiple or “tick” size varies with the financial market, the lattice structure is always present and is often important. In the case of trades on the NYSE, the tick size of $1/8$ of a dollar was determined by the NYSE, and was set in order to avoid unnecessary negotiations between seller and buyer (see, for example

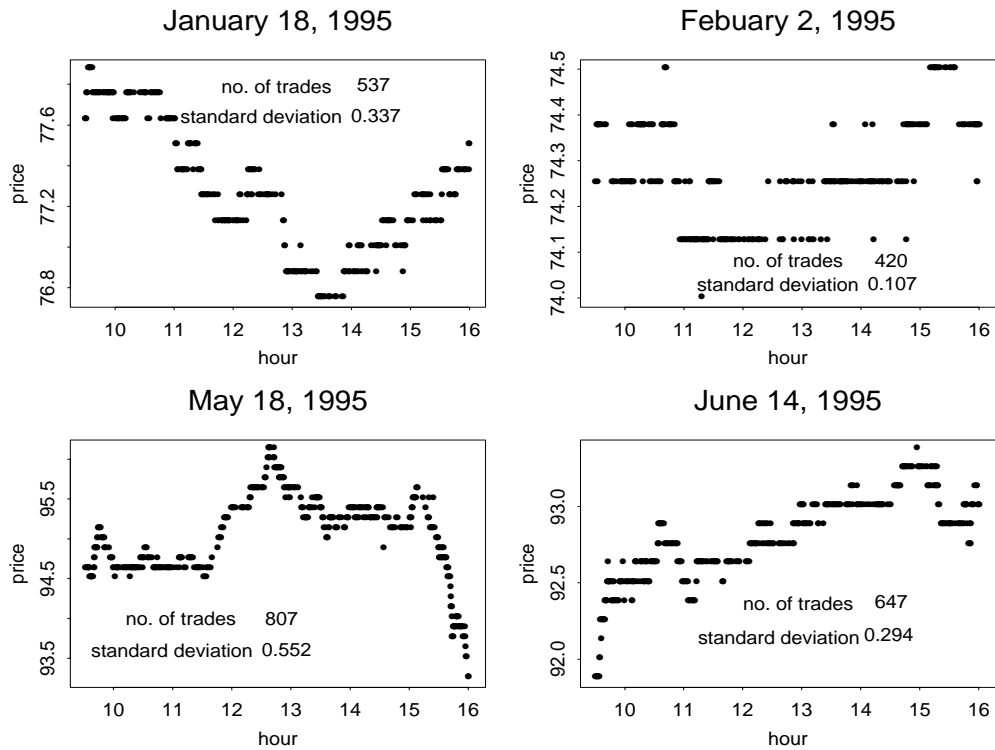


Figure 1: *Plot of all traded IBM prices at the New York stock exchange on four different days in 1995. A trade is represented as a dot (\bullet). I.e. what appears as lines in the graphs are trades at the same price.*

Hasbrouck (1996) and the references therein). In 1997 it was reduced to $1/16$ of a dollar.

There are at least three different ways of dealing with this kind of data:

1. Ignore the fact that the state space of the prices is a lattice and use a continuous sample path model. The literature assessing the effect of this type of misspecification is surveyed by Campbell, Lo, and MacKinlay (1997, pp. 109–128). Their broad conclusion is that methods which use low frequency data (e.g. daily or monthly returns) are not overly influenced by discreteness, but higher frequency analysis can be sensitive.
2. Build a model assuming an underlying unobserved continuous sample path process and a latent continuous exposure cost and then round it to match the discreteness. This method is used by Hasbrouck (1999a) in his model of discrete bid and ask quotes.
3. Directly model what is observed. Such models will have a state space which is a lattice. This approach involves a distinct stochastic model for

the times of trades as well as a model for the price changes at the times of trades.

In this paper we will discuss the third of these approaches. To simplify our exposition we have normed prices so that the tick size is one.

Our basic model structure will be a compound Poisson process in the wide sense of, for example, Grandell (1997). Let $p(u)$ denote the price of the asset at time u , then we allow the non-stationary and non-linear price process to follow

$$p(u) = p(0) + \sum_{t=1}^{N(u)} Z_t, \quad u \geq 0, \quad \text{and} \quad \sum_{t=1}^0 Z_t \stackrel{\text{def}}{=} 0,$$

where $\{N(u)\}_{u \geq 0}$ is a counting process¹ which counts the number of trades recorded up until time u , and Z_t is the price movement associated with the t -th trade. It is important to note that Z_t can be exactly zero, for many trades take place without moving the price. There is no loss of generality in writing down this representation since both $\{N(u)\}_{u \geq 0}$ and $\{Z_t\}_{t \in \mathbf{N}_+}$ can be either continuous or discrete (for ease of notation we will use the slightly abbreviated expressions $\{N(u)\}$ and $\{Z_t\}$). However, we will tend to be more specific than this. We suggest modelling $\{N(u)\}$ as Cox process², that is a Poisson process with a random intensity³. In general, the dynamics of the Cox and price movements processes can be adapted to a wide class of filtrations involving just their own past or more extensive information sets. This is purely an issue of combining both the empirical evidence and a priori economic theory, reflecting both the purpose of the modelling exercise and the data generating mechanism. A simple example of this is that the models we specify should prevent the price process from going negative. This is easy to do by careful modelling of the $\{Z_t\}$ process, which must depend upon the level of the price⁴.

¹There are many equivalent definitions of a counting process. The one which is most helpful in our context states that if $\{N(u)\}_{u \geq 0}$ is a process with state space $\mathbf{Z} \cup \{+\infty\}$ and non-decreasing right continuous paths, then $\{N(u)\}_{u \geq 0}$ is a counting process. Since the paths are non-decreasing and right continuous we automatically get that $\{N(u)\}_{u \geq 0}$ is càdlàg (continu à droite – limite à gauche).

²An convenient example of a Cox process is the influential autoregressive conditional duration (ACD) model advocated by Engle and Russell (1998), which allows straightforward likelihood based econometric inference.

³This is based on the assumption that the counting process is *simple* which means that $N(u)$ increases one unit at its epochs of increase, see Grandell (1991, p. 34). Such an assumption is fulfilled if the intensity measure is continuous.

⁴In practice trading would probably be halted if information released caused very dramatic falls in the share price. At, e.g. the NYSE there are specific rules about when trading should be paused.

1.2 Related work

Independently of our initial draft of this paper Rogers and Zane (1998) have suggested a similar type of compound Poisson process for asset prices. This paper will focus on the study of $\{N(u)\}$ and only allow the $\{Z_t\}$ to be modelled by simple descriptive Markov chains. Papers which have previously looked at the $\{N(u)\}$ process include Engle and Russell (1998) and, subsequently, Meddahi, Renault, and Werker (1998), Rydberg and Shephard (1999), Ghysels, Jasiak, and Gouriéroux (1998) and Hasbrouck (1999b). We will compare our suggestion to this earlier work in Section 5.

Following an initial draft of this paper, Rydberg and Shephard (1998) have studied the dynamics of the $\{Z_t\}$ process within the context of our compound Poisson process framework, while not discussing the specification of the counting process $\{N(u)\}$. Another approach to modelling the dynamics of the $\{Z_t\}$ has also been previously proposed by Russell and Engle (1998). These two models have rather different features since the models proposed in Rydberg and Shephard (1998) could potentially include discrete as well as continuous returns, whereas the model in Russell and Engle (1998) only allows $\{Z_t\}$ to live on a fixed number of points.

Our models are also related to stochastic volatility (SV) or time deformation models, see e.g. Clark (1973), Hull and White (1988), Stein and Stein (1991), Ghysels, Harvey, and Renault (1996) and Barndorff-Nielsen and Shephard (1998). In SV models Brownian motion is deformed, while in the compound Poisson process the Cox process is a deformed Poisson process, the intensity process playing the role of that of the volatility for the Brownian motion. Barndorff-Nielsen and Shephard (1998) have studied the connection between the modelling framework we propose and a SV model in a thickly traded market. Those results have been elaborated by Frey and Runggaldier (1998).

Compound Poisson processes with discrete innovations resemble binomial models of stock prices (see, for example, Dothan (1990)), because they also live on a grid. Our framework is more complicated since binomial models typically live in discrete time, with independent and identically distributed price movements only occurring at deterministic points in time.

Jump diffusions models, which allow discontinuous sample paths for prices, have also been used in finance. The discontinuities are usually introduced as a standard Poisson process (see e.g. Merton (1976)) and are used to model big events such as interventions by government or monetary authorities. They are not appropriate for the empirical phenomenon we are modelling in this paper.

Compound Poisson processes are extensively used in insurance mathematics as a model for the capital of an insurance company. The capital is effected

by the known income from premiums and the randomly arriving stream of insurance claims from policy holders. The earliest reference we know of to the use of Cox processes in this context is Ammeter (1948), while a textbook exposition of this literature is given in Grandell (1991).

The paper is organized as follows. In Section 2 we look at some descriptive statistics for the trade-by-trade data we are analysing — studying both the basic features of the counting process $\{N(u)\}$ and the price innovations process $\{Z_t\}$. Section 3 looks at the general properties of counting processes in our context. In Section 4 we discuss using two signal extraction methods for estimating the current level of intensity of trading in the market. Two alternatives are studied in Section 5, while in Section 6 we look at the implication of our model structure for the dynamics of changes in the price level over intervals of length Δ . In Section 7 we look at connecting our theoretical model of $N(u)$ and the price changes with the empirical evidence. We draw our conclusions in Section 8.

2 Basic features of trade-by-trade data

2.1 The data

The trade data used in this paper is for the IBM share recorded electronically at the New York Stock Exchange in 1995 (NYSE TAQ–data base). The market itself is not electronic but is open outcry and there is one market maker for each stock. Each market maker can make the market in several stocks, for an excellent exposition of market micro structure, see O’Hara (1995). The precision of the time stamp is one second and it is the duty of the seller to report the trade. We first construct a time series for each day on which the exchange was open, computing the price changes at each trade (rescaling the data to have a tick size of one). We cut out all trades registered after 16.00 as this is the official closing of the exchange and our initial data analysis suggested the data was significantly different when it had a time stamp which was after 16.00. Also all trades occurring with an error mark are discarded and in order to reduce the size of the data set we only consider trades which took place at the NYSE.

2.2 Price movements

The dynamics of the price level in calendar time is determined by the properties of $\{N(u)\}$ and $\{Z_t\}$. In order to formulate models for these processes, we

first look at their basic empirical features. We start with the price movements $\{Z_t\}$.

The \log_{10} of the counts of $\{Z_t, Z_{t-1}\}$ is given in Figure 2. This shows the dramatic concentration of the data on 0, 0 which accounts for 60% of the data and substantial mass along the lines

$$\{Z_t = k, Z_{t-1} = 0\}_k \quad \text{and} \quad \{Z_t = 0, Z_{t-1} = k\}_k.$$

Most importantly there is mass along the diagonal

$$\{Z_t = k, Z_{t-1} = -k\}_k,$$

which represents a move in the price which is reversed at the next trade. A high proportion of this represents a single tick up (down) which are followed by an immediate reversal of one tick down (up). This is caused by the discreteness of the sample space and the action of bid/ask bounce. This very significant diagonal is not matched by one along

$$\{Z_t = k, Z_{t-1} = k\}_k$$

which has almost no mass.

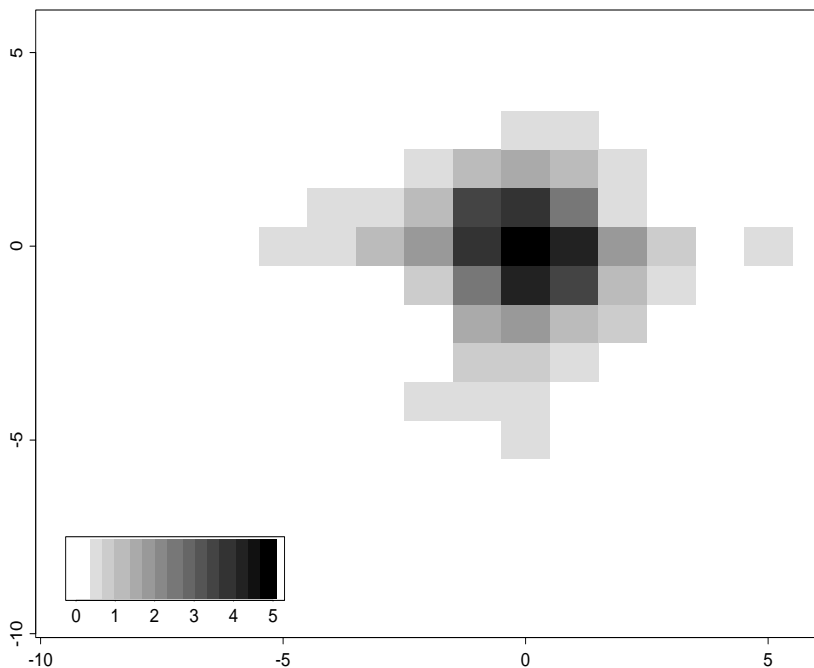


Figure 2: *This figure shows in a \log_{10} scale the number of observations in each coordinate. To all cells 1 was added before taking log in order to avoid problems with log's of zero.*

Figure 2 implies the $\{Z_t\}$ process has extremely significant negative first order autocorrelation. However, Figure 3 shows that by lag two this correlation is close to zero (although lags up to 6 are statistically different than zero). This suggests, up to a very rough Wold representation, the $\{Z_t\}$ are a Markov process.

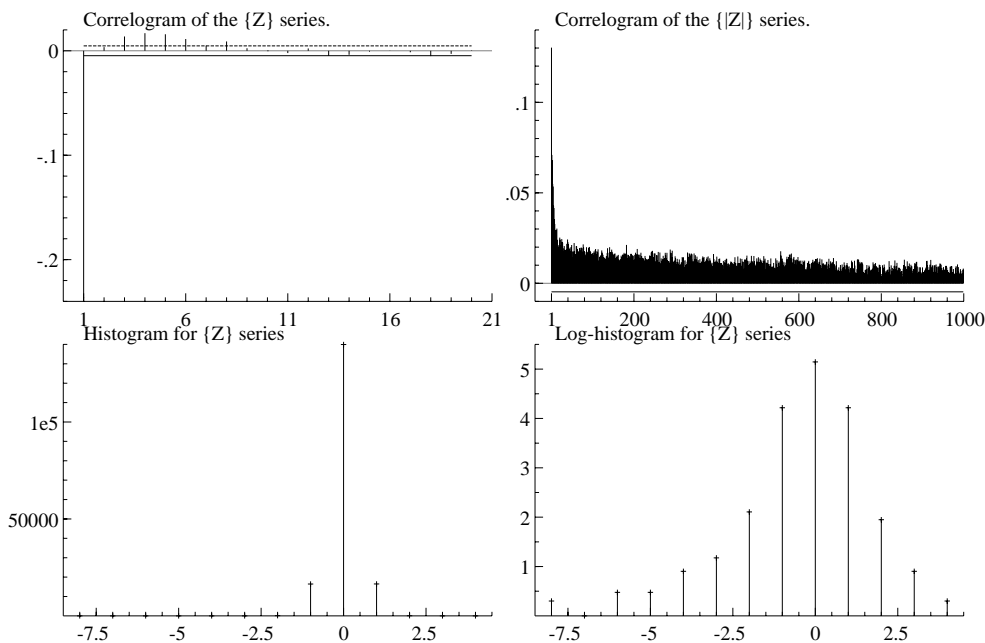


Figure 3: (a) shows the correlogram of Z_t , the price movements. (b) shows correlogram for $|Z_t|$. (c) is histogram of Z_t , while (d) is the log of the histogram plus one.

Figure 3(c) counts the numbers of price movements of the $\{Z_t\}$ equal to particular integers. It demonstrates that most of the trades on the NYSE do not move the price and only rarely does the price move by many ticks. The vast majority of the data is one of $-2, -1, 0, 1, 2$, while the density is seemingly slightly skewed to the left. This is more easily seen by looking at the log-histogram which is again given in Figure 3(d).

Figure 3(b) also gives the correlogram of the $|Z_t|$. This has a large number of lags which are significantly different from zero and take a great deal of time to die down. Indeed there are indications of long-memory type behaviour in this plot. This is not surprising as this is close to the usual volatility clustering that is often observed in financial returns over 5 minute or daily intervals.

See, for example, the econometric literature on autoregressive conditional heteroskedastic (ARCH) and SV models (Bollerslev, Engle, and Nelson (1994), Ghysels, Harvey, and Renault (1996) and Shephard (1996)). This feature of the data will not be the focus of this paper. Instead we look at a broadbrush analysis of the data, for which we maintain the Markov assumption on $\{Z_t\}$. For a thorough analysis of the $\{Z_t\}$ process, see Rydberg and Shephard (1998).

2.3 Trading times

2.3.1 Stylized intra-day, week and monthly effects

The intensity of trading on the NYSE varies considerably through time. In this subsection we study the basic features of the observed sequence of $\{N(u)\}$ for the IBM stock. We do this via a difference operator

$$N_n = N[\{(n+1)\Delta\} -] - N(n\Delta), \quad \Delta > 0, \quad (2.1)$$

which creates a discrete time series from the time continuous counting process by recording the number of trades which occurred in time intervals, or bins, of length Δ . Typically in this paper we take Δ to equal one second. Note that N_n is well defined, since $\{N(u)\}$ is càdlàg, and counts all arrivals with time stamps τ_t such that $n\Delta \leq \tau_t < (n+1)\Delta$.

The top graph of Figure 4 shows an estimate of the average number of trades which occur at each second for each day of the week. The estimate is generated using a natural cubic spline with a different bandwidth selected by generalised cross-validation for each day of the week (see, for example, Green and Silverman (1994)). We can see that for each day trading is brisk in the morning hours, slows down around lunch time and picks up again in the afternoon. In addition there are changes in these patterns between the days of the week. In particular Monday mornings and Friday afternoons are comparatively inactive, while the first 30 minutes of Friday mornings are the most active trading period of the week during 1995 for the IBM stock. Finally, we can see that the first ten minutes of each day are unlike most of the rest of the day — for the activity rate changes very dramatically during this time. Patterns of intra-day trading has been studied by several researchers, e.g. Andersen and Bollerslev (1997), Andersen and Bollerslev (1998) and Guillaume, Dacorogna, Dave, Muller, Olsen, and Pictet (1997).

The bottom graph of Figure 4 shows the number of trades on each day that the NYSE was open. We can see very significant changes in the activity level during the year, with low levels at the beginning of the year and high levels in September and October. Some of the variation of this series arises due to

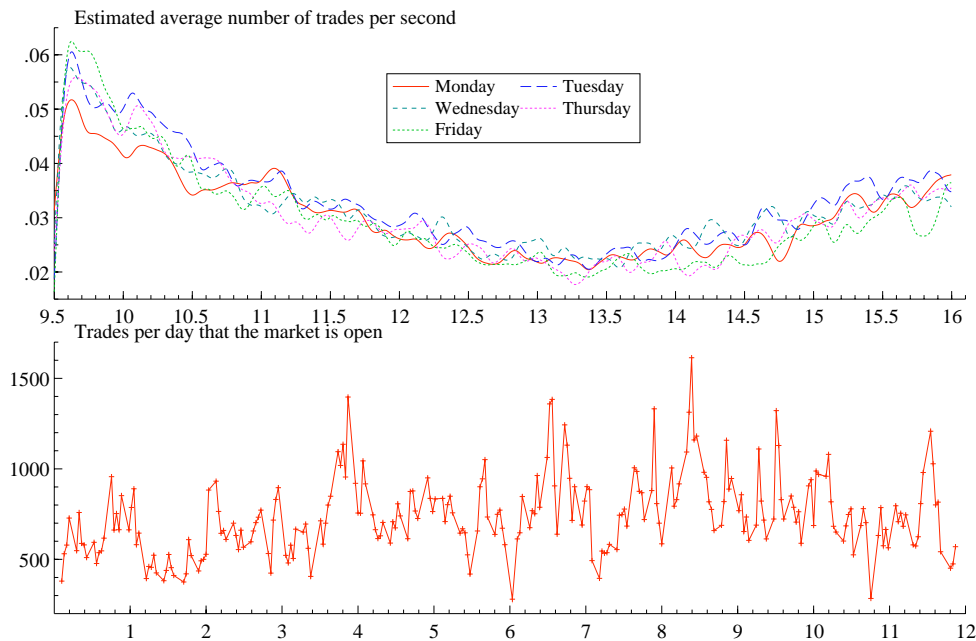


Figure 4: *Top graph is the intensity of the trading per second. Estimated daily curves using a spline with the smoothness penalty selected using generalised cross validation. Bottom graph is the number of trades per day for each day the market is open during the year. The x-axis indicates the month.*

a seasonal component. However, there is also important serial dependence in the series.

2.4 Dynamics of $\{N_n\}$

Throughout we will study the dynamics of the $\{N_n\}$ process with Δ set to one second, focusing on the stochastic properties rather than the deterministic seasonal features. Our first analysis is to look at the daily time series generated by looking at the difference between the $\{N_n\}$ sequence in each day and the corresponding daily seasonal pattern given in Figure 4. For each day we computed the correlogram and plotted the average of these 251 correlograms in Figure 5. This picture shows a negative correlation at lag one, followed by very significant (although quite small) correlations at longer lags. These die down quite slowly, but are mostly irrelevant after 5000 seconds.

The negative correlation at lag one indicates that there are less runs of trades in the series than one would expect if the trades were independently spread

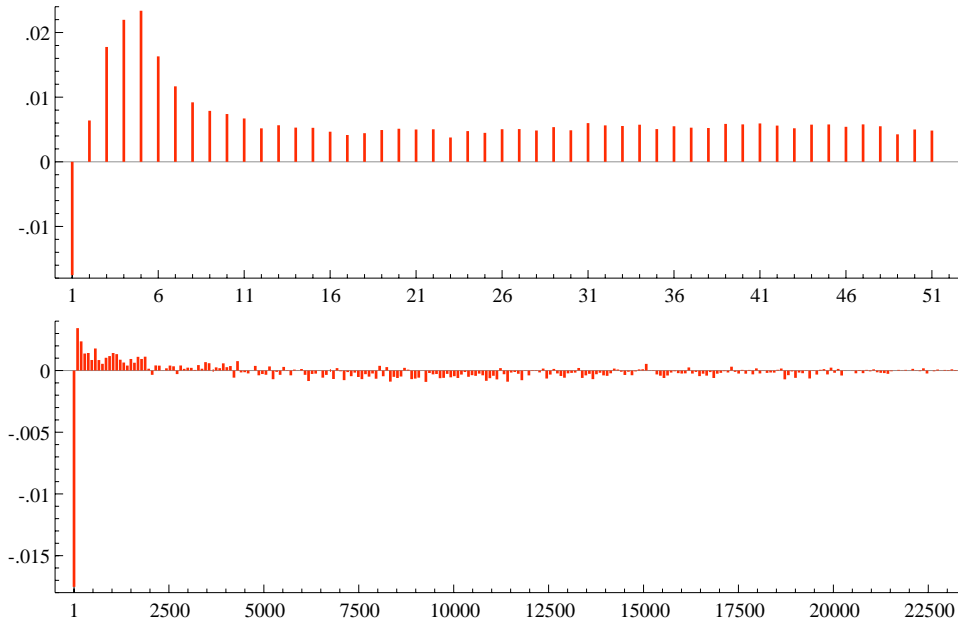


Figure 5: *Averaged correlogram for 251 active days. For each day we computed the correlogram for the day using the 23,400 second by second data. Top correlogram shows first 50 lags, bottom a thinned version of 23,400 lags.*

throughout the day. This is almost certainly due to the inability of the market maker to record trades quickly enough at active times of the market. The positive correlations at other lags are more important to the overall dynamics of the counting process as they are sustained over a large number of lags.

A difficulty with the above analysis is that it may not be picking up very long pieces of memory in the series which wash over from one day onto the next. We now try to measure this. To carry this out we will construct 4 time series each of length 60 working days by simply sticking together 60 days of the difference between $\{N_n\}$ sequence in each day and the corresponding daily seasonal pattern. Each of these series has 1,404,000 observations. Figure 6 gives the correlograms for each of these massive series. On the left we give a thinned version of the correlogram up to 50,000 lags, plotting every 200th correlation to make it easier for the eye to pick up patterns. On the right we plot the first 100 correlations in order to easily see the very short run correlations in the series. These pictures are remarkably similar over the four periods of 60 days and show very long levels of persistence in the correlations

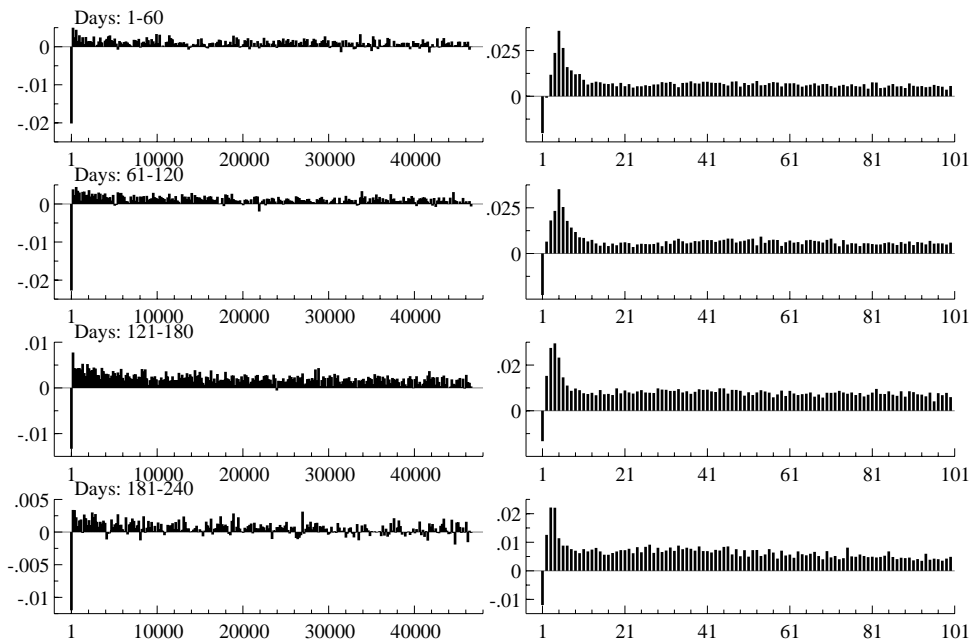


Figure 6: *Correlogram for 60 days of collated data minus seasonal pattern. Length of series is 1,404,000 observations.*

which are qualitatively different from the averaged correlogram approach that we showed earlier.

We can reconcile the long time series analysis with the average correlogram figure by working with the bottom of Figure 4 which shows the number of trades on each day that the NYSE was open. This series reveals a great deal of memory (or a neglected seasonal pattern) between days which will thus impact on the correlograms of the massive time series but have no impact on the averaged correlograms. In order to adjust for this feature of the $\{N_n\}$ we have constructed an adjusted massive series which multiplies the seasonal term by the ratio of the number of trades in the previous day to the average number of trades in the year.

The resulting correlograms for the new massive series are given in Figure 7. They show moderate correlations after about 15,000 lags which seems much more in line with the averaged correlogram analysis. Further, the analysis looks remarkably stable over time as each of these four series looks basically the same. As a result we will focus our analysis on this type of series. However, before we carry out some empirically based modelling we need to improve our

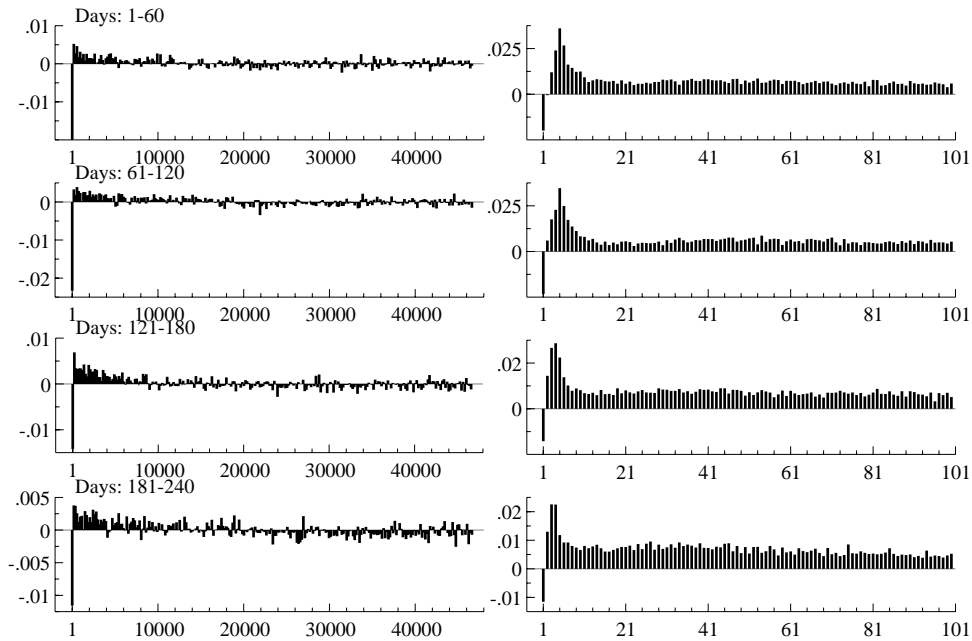


Figure 7: *Correlogram for 60 days of collated data minus seasonal pattern adjusted by dynamics of day by day data.*

understanding of the basic theory of Cox processes in order to model accurately $\{N_n\}$. This is carried out in the next section.

3 Specifying a framework for Cox processes

3.1 Background

The focus in this section will be on constructing simple models which generate the counting process $\{N(u)\}$, which counts the number of trades up to time u . These will be based upon the framework of Cox or doubly stochastic Poisson processes.

We first recall that a stochastic process $\{\Lambda(u)\}$ with

$$\Lambda(0) = 0, \quad \Lambda(u) < \infty \quad \text{for all } u < \infty$$

and non-decreasing realizations is called a random measure. Then let $\{\tilde{N}(u)\}$ be a standard Poisson process and further let Λ and \tilde{N} be independent of each other. Then the point process $\{N(u)\} = \{\tilde{N}\{\Lambda(u)\}\}$ is called a Cox process

(see, for example, Grandell (1997)). A elegant discussion of Cox processes from the viewpoint of subordination is given in Cox and Miller (1965, p. 154).

The random measure Λ is linked to the intensity λ via the integral equation

$$\Lambda(u) = \int_0^u \lambda(s) ds,$$

where $\{\lambda(u)\}$ is a càdlàg positive stochastic process, which implies that $\{\Lambda(u)\}$ is continuous.

The previous subsection analysed the number of trades in the n -th bin denoted by N_n , which occurred in intervals or bins of length one second. We call this a binned time series. Since

$$\{N(u) - N(v) | \Lambda(u) - \Lambda(v)\} \sim \text{Po} \{\Lambda(u) - \Lambda(v)\}$$

where Po denotes the Poisson distribution, we know that

$$N_n | \lambda_n \sim \text{Po}(\lambda_n), \quad \text{where} \quad \lambda_n = \Lambda \{(n+1)\Delta\} - \Lambda(n\Delta). \quad (3.1)$$

Note that we do not need $\Lambda[\{(n+1)\Delta\} -]$ since $\{\Lambda(u)\}$ is continuous by construction. Furthermore, since the Cox process conditionally on $\{\Lambda(u)\}$ has independent increments we have that the binned counts are conditionally independent. An interesting special case of this is where Δ is very small, in which case

$$\Pr(N_n = 0 | \lambda_n) = 1 - \lambda_n + o(\Delta).$$

3.2 Generic properties of the bins

It is possible to work out the autocorrelation pattern of $\{N_n\}$ simply under the condition that $\{\lambda(u)\}$ is covariance stationary. This work follows closely some related ideas on stochastic volatility due to Barndorff-Nielsen and Shephard (1998). These general results will be helpful in allowing us to derive empirically realistic and simple models for the counting process.

Let ξ , ω^2 and r denote, respectively, the mean, the variance and the autocorrelation function of the process $\{\lambda(u)\}$. It is useful to define the notation r^* for the cumulative autocorrelation function, i.e.

$$r^*(t) = \int_0^t r(u) du \quad \text{and} \quad R^*(t) = \int_0^t r^*(u) du.$$

For use below we note that

$$\int_0^t \int_0^t r(u-v) du dv = 2R^*(t)$$

and consequently, assuming that $\lambda(t)$ is square integrable,

$$\text{Var}\{\Lambda(u)\} = 2\omega^2 R^*(u)$$

and

$$\begin{aligned} \text{Cov}\{\lambda_n, \lambda_{n+s}\} &= \omega^2 [R^*\{(s+1)\Delta\} - 2R^*(s\Delta) + R^*\{(s-1)\Delta\}] \\ &= \omega^2 \diamond R^*(\Delta s), \end{aligned} \quad (3.2)$$

where $\diamond R^*(s)$ is defined as

$$\diamond R^*(s) = R^*(s + \Delta) - 2R^*(s) + R^*(s - \Delta).$$

The moments of N_n follow immediately. Notice

$$\text{E}(N_n) = \text{E}(\lambda_n) = \Delta\xi,$$

$$\text{Var}(N_n) = \text{Var}(\lambda_n) + \text{E}(\lambda_n) = 2\omega^2 R^*(\Delta) + \Delta\xi,$$

Further

$$\text{Cov}(N_n, N_{n+s}) = \text{E}(N_n N_{n+s}) - \text{E}(N_n)^2 \quad (3.3)$$

$$= \text{Cov}(\lambda_n, \lambda_{n+s}) = \omega^2 \diamond R^*(\Delta s) \quad (3.4)$$

The implication is that

$$\text{Cor}(N_n, N_{n+s}) = q \diamond R^*(\Delta s), \quad \text{where} \quad q = \frac{\omega^2}{2\omega^2 R^*(\Delta) + \Delta\xi}.$$

Example Suppose that

$$r(s) = e^{-\beta|s|}$$

for some $\beta > 0$. Such autocorrelation functions occur when we use the Ornstein-Uhlenbeck process

$$d\lambda(u) = -\beta\lambda(u)dt + dz(\beta u), \quad (3.5)$$

with $0 < \beta < \infty$. The process z is a homogeneous Lévy process⁵ with positive increments (also termed a subordinator). They are studied at length in Barndorff-Nielsen and Shephard (1998). Exactly the same autocorrelation function results from the ‘constant elasticity of variance’ process

$$d\lambda(u) = -\beta \left\{ \lambda(u) - \bar{\lambda} \right\} dt + \gamma \{ \lambda(u) \}^d dW(u), \quad d \geq 1/2, \quad (3.6)$$

⁵Note that $\text{Var}\{\lambda(u)\} = \omega^2$ does not depend on β . This is due to the reparametrisation used on the definition of Ornstein-Uhlenbeck processes, where the β enters the time of z .

where $W(u)$ is standard Brownian motion. This general structure, which is always covariance (and strictly) stationary if $0 < \beta < \infty$, has been recently highlighted by Meddahi and Renault (1996) who strongly argue that it provides a great deal of tractability in terms of studying temporal aggregation of stochastic volatility using different information sets. Then,

$$R^*(\Delta) = \frac{1}{\beta^2} (\beta\Delta + e^{-\beta\Delta} - 1)$$

and for $s > 0$,

$$\diamond R^*(\Delta s) = \beta^{-2} (1 - e^{-\beta\Delta})^2 e^{-\beta\Delta(s-1)}$$

which falls exponentially with s . Hence

$$\text{Var}(N_n) = \frac{2\omega^2}{\beta^2} (e^{(-\beta\Delta)} - 1) + \Delta \left(\xi + \frac{2\omega^2}{\beta} \right)$$

which is linear for large Δ . Furthermore, we get that

$$\text{Cor}\{N_n, N_{n+s}\} = \frac{\omega^2}{2\omega^2 R^*(\Delta) + \Delta\xi} \diamond R^*(\Delta s) = ce^{-\beta\Delta(s-1)}. \quad (3.7)$$

where

$$c = \frac{\omega^2 (1 - e^{-\beta\Delta})^2}{2\omega^2 (\beta\Delta + e^{-\beta\Delta} - 1) + \beta^2 \Delta \xi}.$$

Note that $0 < c < 1$ and that N_n has a Wold representation which is a ARMA(1,1) process with weak white noise errors.

4 Signal extraction

4.1 Estimation of intensity

We could think of $N_n | \lambda_n \sim \text{Po}(\lambda_n)$ as a state space model and then perform signal extraction on the random integrated intensity, λ_n . In this section we will study two ways of performing signal extraction in this context: a best linear method generated by the Kalman filter and an efficient method computed using a particle filter. Both are adapted to our case from the treatment of stochastic volatility developed by Barndorff-Nielsen and Shephard (1998).

The number of trades in the n -th interval of length Δ is

$$N_n = \lambda_n + (N_n - \lambda_n) = \lambda_n + u_n,$$

where $\{u_n\}$ is a Martingale difference sequence. Further, so long as $\{\lambda_n\}$ is covariance stationary, $\{u_n\}$ is a zero mean, white noise process. We will write

the variance of $\{u_n\}$ as σ_u^2 . If we were to adopt the Barndorff-Nielsen and Shephard (1998) model for $\lambda(t)$, given in (3.5), then we have two available approaches: a discrete time approximation or the exact continuous time version. For ease of exposition we have set $\Delta = 1$ throughout and we will only discuss the discrete time case. The extension to the continuous time version follows using results in Barndorff-Nielsen and Shephard (1998).

4.2 Discrete time model

First we can take a Euler style approximation to λ_n so that

$$\lambda_{n+1} = e^{-\beta} \lambda_n + (1 - e^{-\beta}) \eta_n, \quad \text{and} \quad \mathbb{E}(\eta_n) = \mathbb{E}(\lambda_n) = \xi. \quad (4.1)$$

Here $\{\eta_n\}$ are independent and identically distributed strictly positive random variables. Then this model is in a linear state space form and so the Kalman filter provides the best linear estimator, written $a_{n+1|n}$, of the unobserved λ_{n+1} given N_1, \dots, N_n . In particular the Kalman filter is given by the

$$a_{n+1|n} = e^{-\beta} \left\{ a_{n|n-1} + \frac{p_{n|n-1}}{p_{n|n-1} + 1} (N_n - a_{n|n-1}) \right\} + (1 - e^{-\beta}) \xi,$$

and its associated mean square error $\sigma_u^2 p_{n+1|n}$, where $v_n = (1 - e^{-\beta}) (\eta_n - \xi)$,

$$p_{n+1|n} = \frac{e^{-2\beta} p_{n|n-1}}{p_{n|n-1} + 1} + \frac{\sigma_v^2}{\sigma_u^2}, \quad \text{where} \quad \sigma_v^2 = \text{Var}(v_n).$$

In many senses the Kalman filter solution is unsatisfactory for it does not give $f(\lambda_{n+1}|\mathcal{F}_n)$, nor in particular the fully efficient estimator $\mathbb{E}(\lambda_{n+1}|\mathcal{F}_n)$. We employ the auxiliary sampling importance resampling particle filtering of Pitt and Shephard (1999) to carry out this task.

We use the notation $f(\lambda_{n+1}|\lambda_n)$ to denote the Markov evolution of the unobserved intensity over time of the discrete time model. The particle filter has the following basic structure. The density of $\lambda_n|\mathcal{F}_n$ is approximated by a sample $\lambda_{1,n}, \dots, \lambda_{M,n}$. The particle filter regenerates these points into an approximate sample from $\lambda_{n+1}|\mathcal{F}_{n+1}$ by sampling from

$$\hat{f}(\lambda_{n+1}|\mathcal{F}_{n+1}) \propto f(N_{n+1}|\lambda_{n+1}) \sum_{k=1}^M f(\lambda_{n+1}|\lambda_{k,n}). \quad (4.2)$$

This is carried out by sampling k^j with probability proportional to

$$f(N_{n+1}|\mu_{n+1}^k), \quad \text{where} \quad \mu_{n+1}^k = \mathbb{E}(\lambda_{n+1}|\lambda_{k,n}),$$

and then drawing from $\lambda_{j,n+1} \sim \lambda_{n+1} | \lambda_{kj,n}$. This is carried out R times. The resulting population of particles are given weights proportional to

$$w_j = \frac{f(N_{n+1} | \lambda_{j,n+1})}{f(N_{n+1} | \lambda_{kj,n+1})}, \quad \pi_j = \frac{w_j}{\sum_{i=1}^R w_i}, \quad j = 1, \dots, R.$$

We resample this population with probabilities $\{\pi_j\}$ to produce a sample of size M , $\lambda_{1,n+1}, \dots, \lambda_{M,n+1}$. This sample is approximately from $\lambda_{n+1} | \mathcal{F}_{n+1}$. In this way we update the sample at each time step through the entire sample, $n = 1, 2, \dots, T$. We can estimate $E(\lambda_{n+1} | \mathcal{F}_n)$ by

$$e^{-\beta} \frac{1}{M} \sum_{j=1}^M \lambda_{j,n} + (1 - e^{-\beta}) \xi.$$

In practice when we have applied the ASIR particle filter in this context we have taken, in the order of, $M = 10,000$ and $R = 3M$.

5 Alternatives models of $N(u)$

5.1 Time between trades

In the preceding section we have studied the stochastic properties of the number of trades up to time u , written $N(u)$. We did this via modelling the number of trades occurring in bins of length Δ , $N_n = N[\{(n+1)\Delta\} -] - N(n\Delta)$. An alternative is to model the time between trades, also termed *durations*. Let τ_t be the time of the t -th trade. Then it is given by

$$\tau_t = \min_u \{N(u) = t\}, \quad t = 1, 2, \dots, N(S),$$

recalling that $N(S)$ is the number of trades in the period of length S we are studying. Then the length of time between trades is

$$L_t = \tau_t - \tau_{t-1}, \quad t = 1, 2, \dots$$

For our data a small number of these times are exactly zero.

From a statistical viewpoint we can think of $\{L_t\}$ as a time series of duration times. There is an enormous literature on the analysis of durations, although most of it does not have a time series interpretation. We refer to Lancaster (1990), Synder and Miller (1991) and Cox and Oakes (1984) for general discussions of this literature.

In the econometric literature an influential model of the durations is the autoregressive conditional duration (ACD) model of Engle and Russell (1998). This puts

$$L_t = \varepsilon_t \psi_t, \quad \varepsilon_t > 0, \quad E(\varepsilon_t) = 1$$

and the ε_t 's are independent identical distributed (*i.i.d.*), with

$$\psi_t = \alpha + \sum_{j=1}^p \gamma_j L_{t-j} + \sum_{j=1}^q \beta_j \psi_{t-j}.$$

Here $\psi_t = E(L_t | \mathcal{F}_{t-1})$, the conditional expected waiting time, where \mathcal{F}_t is a filtration, potentially containing all information up til time $t-1$. The mathematical structure of this model is identical to that of the square of an GARCH model associated with the work of Engle (1982) and Bollerslev (1986). The model has many similarities with earlier work by Wold (1948) and Cox (1972). In practice Engle and Russell (1998) have used an exponential or Weibull distribution on the $\{\varepsilon_t\}$. Straightforward alternative structures would be to parameterise the log ψ_t instead of the ψ_t .

A key feature of this model is that, conditional on ψ_0 , the likelihood can be computed via a prediction decomposition. Further the number of terms that need to be evaluated is only $N(S)$, rather than the number of seconds. However, if the focus is on events in calendar time this model has a serious drawback, namely that the intensity λ_n cannot be calculated analytically, this is discussed in more detail in Rydberg and Shephard (1999).

5.2 Modelling the price level

The ACD model implies the following structure for the price level of the stock. It has the evolution according to the process

$$p\left(\sum_{t=1}^n L_t\right) = p(0) + \sum_{t=1}^n Z_t,$$

which tells us the price at n irregularly space time points. In continuous time this has a less elegant expression as

$$p(u) = p(0) + \sum_{t=1}^{N(u)} Z_t, \quad \text{where} \quad N(u) = \arg \max_n \left\{ \sum_{t=1}^n L_t \leq u \right\}.$$

Note that we need $\sum_{t=1}^n L_t \leq u$ in order for $N(u)$ to be càdlàg. This implies

$$\begin{aligned} N_n &= N[\{(n+1)\Delta\} -] - N(n\Delta) \\ &= \arg \max_r \left\{ \sum_{t=1}^r L_t < (n+1)\Delta \right\} - \arg \max_r \left\{ \sum_{t=1}^r L_t \leq n\Delta \right\}. \end{aligned}$$

The complexity of the relationship between the $\{L_t\}$ and $\{N_n\}$ implies studying the behaviour of returns in calendar time implied by ACD style models is difficult.

Other difficulties with the ACD style of model include the following.

1. The intra-day seasonal pattern of changing activity has quite a complicated impact on the conditional waiting times, for the index t does not correspond to a physical time. This is quite important in this context as the intra-day pattern is very strong and quite quickly changing.
2. The conditional duration is only changed when there is a trade in the stock. However, in terms of economic theory it maybe easier to parameterise the model if we were to allow the conditional waiting time to change with any new information arriving in the market. At the most refined level this would argue that the conditional intensity should change every second.
3. At the end of each trading day, there is a period which does not result in a trade. This has an impact on the likelihood function, although this is easy to compute.

5.3 A BIN model

In a recent paper Rydberg and Shephard (1999) have suggested a simple direct model for $\{N_n\}$ in discrete time. That is they model, as a time series, the number of counts in the interval of length Δ . In their simplest model they write, with \mathcal{F}_n^N as the natural filtration of the $\{N_n\}$ sequence,

$$N_n | \mathcal{F}_{n-1}^N \sim \text{Po}(\lambda_n), \quad \text{where} \quad \lambda_n = \alpha + \gamma N_{n-1} + \delta \lambda_{n-1}.$$

They impose the constraints that $\alpha, \gamma, \delta > 0$. Rydberg and Shephard (1999) show that for this model the following results hold

1. The conditional likelihood $f(N_1, \dots, N_T | \lambda_1)$ can be computed.
2. The process is covariance stationary if and only if $\gamma + \delta < 1$.
3. If the process is stationary then

$$\begin{aligned} \text{Cor}(N_n, N_{n+1}) &= \frac{\gamma \{1 - \delta(\gamma + \delta)\}}{1 + \delta^2 - 2\delta(\gamma + \delta)}, \\ \text{Cor}(N_n, N_{n+s}) &= \text{Cor}(N_n, N_{n-1}) \gamma^{s-1}, \quad s = 2, 3, \dots \end{aligned}$$

This is the same autocorrelation as that derived from a Cox process for $\{N(u)\}$ when the intensity was an Ornstein–Uhlenbeck process, see formula (3.7).

This model allows us to predict future $\{N_n\}$ or time aggregations of that process. As a result it seems ideally placed if our goal is to model the price process.

6 The properties of returns

6.1 Model structure

In this section we will attempt to put together a model for the intensities and the price movements in order to produce some simple properties of returns over discrete periods of time, which are given by

$$p_n = p[\{(n+1)\Delta\} -] - p(n\Delta).$$

p_n is well defined since $\{p(u)\}$ is càdlàg. We will use the structure

$$p(u) = p(0) + \sum_{t=1}^{N(u)} Z_t,$$

together with three basic assumptions.

1. The $\{Z_t\}$ are a zero mean, first order moving average process, with autocorrelation $\rho_Z(1)$ and unconditional variance σ_Z^2 . Sometimes it will be helpful to write this out explicitly in terms of $Z_t = \varepsilon_t + \theta\varepsilon_{t-1}$, where $\{\varepsilon_t\}$ is an *i.i.d.* zero mean process. An important consequence of this structure is that $|\rho_Z(1)| \leq \frac{1}{2}$ and

$$\text{Cov}(Z_t, Z_{t-1}) = \theta \text{Var}(\varepsilon_t) = \frac{\theta}{1 + \theta^2} \sigma_Z^2.$$

2. Trades occurring irregularly in time according to a Cox process.
3. The price movements will be stochastically independent of the timing of trades.

Assumptions (1) or (3) are not entirely reasonable however, to a first approximation, they will be helpful. For a detailed discussion and model of this process, see Rydberg and Shephard (1998).

6.2 Two moments

We first discuss the linear structure of the model. We immediately get that

$$\text{E}\{p(u)\} = p(0).$$

For the variance, by using that $\mu_N^e(u) = \text{E}\{N(u)|N(u) > 0\}$, we get the following structure:

$$\text{Var}\{p(u)\} = \text{E}\left[\text{Var}\left\{\sum_{t=1}^{N(u)} Z_t \middle| N(u)\right\}\right] \quad (6.1)$$

$$= \sigma_Z^2 \text{Pr}\{N(u) > 0\} [\mu_N^e(u) + 2\{\mu_N^e(u) - 1\}\rho_Z(1)] \quad (6.2)$$

$$= \sigma_Z^2 [\text{E}\{N(u)\}\{1 + 2\rho_Z(1)\} - 2\rho_Z(1)\text{Pr}\{N(u) > 0\}] \quad (6.3)$$

If u is quite large then for an active stock $N(u)$ must also be large. Thus

$$\text{Var}\{p(u)\} \simeq \sigma_Z^2 \{1 + 2\rho_Z(1)\} E\{N(u)\}, \quad \text{for large } u.$$

In practice we have found that $2\rho_Z(1)$ to be around -0.5 , so the dependence in the price movements has a very considerable influence on this volatility measure. An important feature of this result is that $\text{Var}\{p(u)\}$ is approximately proportional to $E\{N(u)\}$ for large u , with the constant being $\sigma_Z^2 \{1 + 2\rho_Z(1)\}$. However, it should be noted that u has to be pretty large in practice for this to hold. If $\Delta = 30$ seconds then $\Pr\{N(u) > 0\}$ is around $\frac{1}{2}$. One use of this result is that

$$\text{Var}(p_n) = E \left[\text{Var} \left\{ \sum_{t=1}^{N_n} Z_t \middle| N_n \right\} \right] = \text{Var}\{p(\Delta)\},$$

and so for large Δ the

$$\text{Var}(p_n) \simeq \sigma_Z^2 \{1 + 2\rho_Z(1)\} E\{N(\Delta)\}. \quad (6.4)$$

When Δ is very small we have that $\Pr\{N(u) > 0\} \simeq E\{N(\Delta)\}$ and so

$$\text{Var}(p_n) \simeq \sigma_Z^2 E\{N(\Delta)\}, \quad \text{for small } \Delta. \quad (6.5)$$

These results for returns are not immediately obvious, but they are important. For small Δ it is very rare to get more than a single price movement and so the presence of correlation amongst the price movements is irrelevant (note the variance would be exactly $\sigma_Z^2 E\{N(\Delta)\}$ if the $\{Z_t\}$ are *i.i.d.* and have zero mean). For larger Δ it is possible to get more than a single price movement occurring in the interval and hence the correlation reduces the variation in the returns and causes the non-linearity in the relationship between $\text{Var}(p_n)$ and $E\{N(\Delta)\}$.

6.3 Conditional independence

In order to study the dependence between p_n and p_{n+s} it is helpful to work conditionally on

$$N_n, N\{(n+s)\Delta\} - N\{(n+1)\Delta\}, N_{n+s}.$$

A crucial feature of our setup is that

$$p_n \perp\!\!\!\perp p_{n+s} \middle| N_n, N_{n+s}, N\{(n+s)\Delta\} - N\{(n+1)\Delta\} > 0,$$

that is the returns will be conditionally independent if there are trades in-between these time periods. Hence when we look at any dependence structure

we will only be interested in situations where $N \{(n+s)\Delta\} = N \{(n+1)\Delta\}$. An implication of this, which we will use twice in our calculations, is that we can write

$$\begin{cases} p_n &= \varepsilon_{N(n\Delta+\Delta)} + x \\ p_{n+s} &= \theta\varepsilon_{N(n\Delta+\Delta)} + y \end{cases}, \quad (6.6)$$

where

$$\begin{cases} x &= \theta\varepsilon_{N(n\Delta+\Delta)-1} + Z_{N(n\Delta+\Delta)-1} + \dots + Z_{N(n\Delta)+1} \\ y &= Z_{N\{(n+s)\Delta+\Delta\}} + \dots + Z_{N\{(n+s)\Delta\}+2} + \varepsilon_{N(n\Delta+\Delta)+1} \end{cases}.$$

We can think of x as shocks to the price which occur before $\varepsilon_{N(n\Delta+\Delta)}$, while y are the shocks after that period. The only shared shock in this framework is $\varepsilon_{N(n\Delta+\Delta)}$. As a result y , $\varepsilon_{N(n\Delta+\Delta)}$ and x are conditionally (on $N_{n+s}, N_n, N \{(n+s)\Delta\} - N \{(n+1)\Delta\} = 0$), mutually independent.

6.4 Autocorrelation

Let $\gamma_p(s) = \text{Cov}(p_{n+s}, p_n)$. The only way there can be linear dependence is if there is a trade between times $n\Delta, n\Delta + \Delta$, then no trade in the interval $(n+1)\Delta, (n+s)\Delta$ and finally a trade in times $(n+s)\Delta, (n+s+1)\Delta$. Thus the autocorrelation $\gamma_p(s)$ equals

$$\begin{aligned} & \text{Cov} \left\{ \varepsilon_{N(n\Delta+\Delta)}, \theta\varepsilon_{N(n\Delta+\Delta)} \right\} \\ & \times \Pr [N_{n+s} > 0, N \{(n+s)\Delta\} = N \{(n+1)\Delta\}, N_n > 0], \end{aligned}$$

which in turn equals

$$\sigma_Z^2 \rho_Z(1) \Pr [N_{n+s} > 0, N \{(n+s)\Delta\} = N \{(n+1)\Delta\} | N_n > 0] \Pr (N_n > 0).$$

Hence

$$\begin{aligned} \rho_p(s) &= \text{Cor}(p_{n+s}, p_n) \\ &= \frac{\rho_Z(1) \Pr [N_{n+s} > 0, N \{(n+s)\Delta\} = N \{(n+1)\Delta\} | N_n > 0]}{\text{Var}(p_n)}. \end{aligned}$$

As $\rho_Z(1) < 0$, this correlation will be negative for every $s > 0$. If $s = 1$, then as $\Delta \rightarrow 0$ this correlation approaches zero, as there is little chance there will be two contiguous trades and so there cannot be any linear dependence. For larger Δ the dependence becomes apparent with a strong negative correlation, but for large Δ the series will be basically uncorrelated. All of these effects weaken as s increases beyond one due to the requirement that there is no trade in the interval $(n+1)\Delta, (n+s)\Delta$.

6.5 Volatility clustering

6.5.1 No dependence case

In order to study volatility clustering we will initially assume the $\{Z_t\}$ are independent over time. Then let $s > 0$,

$$\begin{aligned}
\gamma_{p^2}(s) &= \text{Cov}(p_{n+s}^2, p_n^2) \\
&= \text{E} \left\{ \text{Cov} \left(p_{n+s}^2, p_n^2 \mid N_{n+s}, N_n \right) \right\} + \text{Cov} \left\{ \text{E} \left(p_{n+s}^2 \mid N_{n+s} \right), \text{E} \left(p_n^2 \mid N_n \right) \right\} \\
&= \text{Cov} \left\{ \text{Var} \left(p_{n+s} \mid N_{n+s} \right), \text{Var} \left(p_n \mid N_n \right) \right\} \\
&= \sigma_Z^4 \text{Cov} \left(N_{n+s}, N_n \right) \\
&= \sigma_Z^4 \omega^2 \diamond R^*(\Delta s),
\end{aligned}$$

using (3.3) and $\text{Cov} \left(p_{n+s}^2, p_n^2 \mid N_{n+s}, N_n \right) = 0$. This is a completely general expression for volatility clustering under the overly strong independence assumption on the $\{Z_t\}$.

6.5.2 Dependence case

We are interested in the problem where there is dependence amongst the $\{Z_t\}$. We start by noting that we still have

$$\gamma_{p^2}(s) = \text{E} \left\{ \text{Cov} \left(p_{n+s}^2, p_n^2 \mid N_{n+s}, N_n \right) \right\} + \text{Cov} \left\{ \text{Var} \left(p_{n+s} \mid N_{n+s} \right), \text{Var} \left(p_n \mid N_n \right) \right\}$$

and that, using (6.1),

$$\begin{aligned}
&\text{Var} \left(p_n \mid N_n \right) - \text{E} \left\{ \text{Var} \left(p_n \mid N_n \right) \right\} \\
&= \sigma_Z^2 \left\langle \begin{array}{c} \{1 + 2\rho_Z(1)\} [N_n - \text{E} \{N(\Delta)\}] \\ -2\rho_Z(1) [I \{N_n > 0\} - \text{Pr} \{N(\Delta) > 0\}] \end{array} \right\rangle.
\end{aligned}$$

As a result, for $s > 0$

$$\begin{aligned}
&\text{Cov} \left\{ \text{Var} \left(p_{n+s} \mid N_{n+s} \right), \text{Var} \left(p_n \mid N_n \right) \right\} \tag{6.7} \\
&= \sigma_Z^4 \left\langle \begin{array}{c} \{1 + 2\rho_Z(1)\}^2 \text{Cov} \left(N_{n+s}, N_n \right) \\ +4\rho_Z(1)^2 \text{Cov} [I \{N_{n+s} > 0\}, I \{N_n > 0\}] \\ -4 \{1 + 2\rho_Z(1)\} \rho_Z(1) \text{Cov} [I \{N_{n+s} > 0\}, N_n] \end{array} \right\rangle.
\end{aligned}$$

For large values of Δ , in a thickly traded market $N_{n+s} > 0$ and so

$$\text{Cov} [I \{N_{n+s} > 0\}, N_n] = 0$$

and

$$\text{Cov} [I \{N_{n+s} > 0\}, I \{N_n > 0\}] = 0.$$

The implication will be that

$$\text{Cov} \{ \text{Var} (p_{n+s}|N_{n+s}), \text{Var}(p_n|N_n) \} \simeq \{1 + 2\rho_Z(1)\}^2 \text{Cov} (N_{n+s}, N_n). \quad (6.8)$$

For small Δ ,

$$\begin{aligned} \text{Cov} (N_{n+s}, N_n) &\simeq \text{Cov} [I \{N_{n+s} > 0\}, N_n] \\ &\simeq \text{Cov} [I \{N_{n+s} > 0\}, I \{N_n > 0\}] \end{aligned}$$

and so

$$\text{Cov} \{ \text{Var} (p_{n+s}|N_{n+s}), \text{Var}(p_n|N_n) \} \simeq \text{Cov} (N_{n+s}, N_n), \quad (6.9)$$

which again reflects the irrelevance of the dependence amongst the Z_t for small Δ .

In order to evaluate $\text{E} \{ \text{Cov} (p_{n+s}^2, p_n^2 | N_{n+s}, N_n) \}$ it will be convenient to use the structure introduced in (6.6). Then

$$\begin{aligned} &\text{Cov} (p_{n+s}^2, p_n^2 | N_{n+s} > 0, N_n > 0, N \{(n+s)\Delta\} = N \{(n+1)\Delta\}) \\ &= \text{Cov} \left\{ \left(\theta \varepsilon_{N(n\Delta+\Delta)} + y \right)^2, \left(\varepsilon_{N(n\Delta+\Delta)} + x \right)^2 \middle| N_{n+s}, N_n \right\} \\ &= \theta^2 \text{Var} \left(\varepsilon_{N(n\Delta+\Delta)}^2 \middle| N_{n+s}, N_n \right) \\ &= \sigma_Z^4 \rho_Z(1)^2 \end{aligned}$$

Of course unconditionally

$$\begin{aligned} &\text{E} \left\{ \text{Cov} (p_{n+s}^2, p_n^2 | N_{n+s}, N_n) \right\} \\ &= \sigma_Z^4 \rho_Z(1)^2 \text{E} [N_{n+s} > 0, N_n > 0, N \{(n+s)\Delta\} = N \{(n+1)\Delta\}]. \end{aligned}$$

If $s = 1$ this term can have a reasonably important contribution when Δ is small. Then

$$\gamma_{p^2}(1) \simeq \{1 + 2\rho_Z(1)^2\} \text{Cov} (N_{n+1}, N_n),$$

but for larger values of s it is likely to be swamped by (6.7) and is thus irrelevant. For large values of Δ the contribution of this term will be basically irrelevant even at lag 1.

The dominant terms in these expressions are given in (6.9) and (6.8). They suggest the dynamics features of the autocorrelations amongst the squares changes with Δ in an important way which is not apparent when there is no dependence in the price movements or in the corresponding SV type model. It is the combination of discreteness and serial dependence which is needed in order to get this feature of the model.

7 Empirical analysis

7.1 The arrivals

Let us briefly study the behaviour of the arrival process. Recall that in the example in Section 3.2 we discussed Ornstein–Uhlenbeck processes as possible models for the intensity process. For such types the autocorrelation function is of the type

$$r(s) = e^{-\beta|s|}$$

for some $\beta > 0$, and

$$R^*(\Delta) = \frac{1}{\beta^2} (\beta\Delta + e^{-\beta\Delta} - 1).$$

This implies that the overdispersion of the counts is

$$\text{Var}(N_n) - \text{E}(N_n) = \frac{2\omega^2}{\beta^2} (\beta\Delta + e^{-\beta\Delta} - 1).$$

In order to investigate if this very simple type of model, is at least in the right ball park, we randomly chose four days in 1995. The four days are rather different in intensity levels, the lowest of the four was 21 February where the average duration was 62.6 seconds while the highest was on 27 June where the average duration was 31.7 seconds, almost twice as busy. In order not to have to deal with individual day effects we for each of these four days calculated the empirical values of $\text{Var}(N_n) - \text{E}(\lambda_n)$ for $1 \leq \Delta \leq 300$ and estimated the two parameters β and ω^2 . The results can be found in Figure 8. For all the days the shape of the curve seem to capture the relationship quite well. There is some variation in both β and ω^2 showing different levels of over–dispersion and memory. The estimated values of β correspond to half–lives of between 80 and 127 seconds. This is not quite enough to capture all of the memory we have seen in the intensity process in Figures 5 and 6.

One way to go forward from here could be to assume that the intensity process follows a sum of Ornstein–Uhlenbeck processes, studied in e.g. Barndorff-Nielsen and Shephard (1999) as a way of modelling the volatility process. For such processes the autocorrelation function is given by

$$r(s) = \sum_{i=1}^k w_i e^{-\beta_i |s|}$$

for some $\beta_i > 0$ and $\sum_{i=1}^k w_i = 1$. Furthermore,

$$R^*(\Delta) = \sum_{i=1}^k w_i R_i^*(\Delta)$$

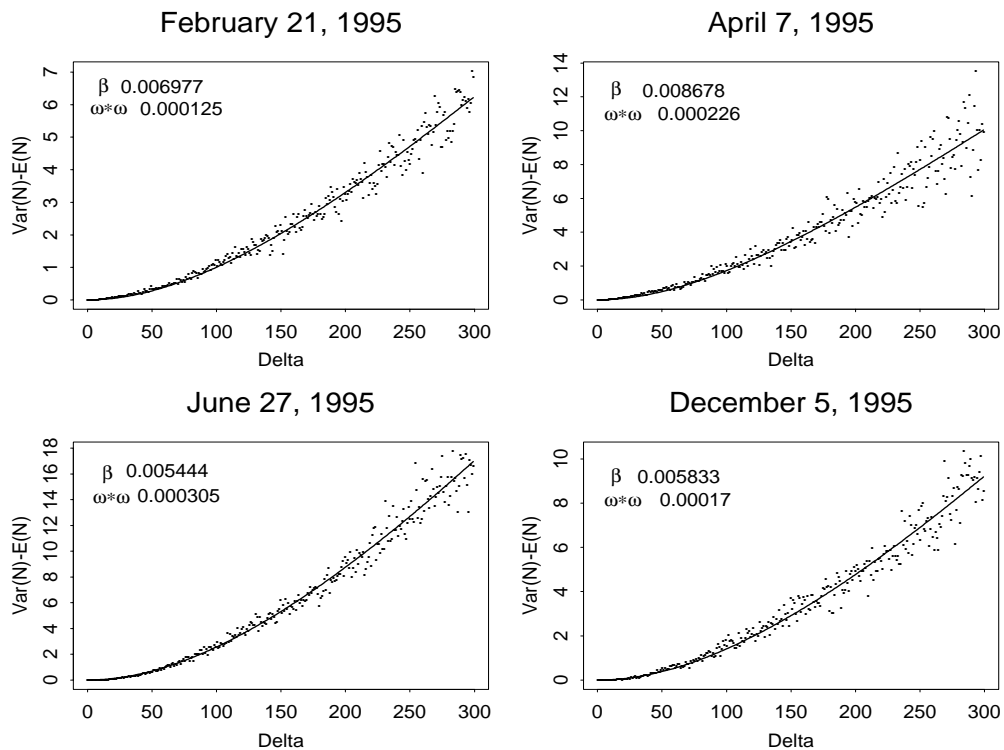


Figure 8: *The observed relationship between $\text{Var}(N_n) - \text{E}(N_n)$ and Δ (represented as a \cdot) and the estimated curve resulting from the assumption that the intensity process can be modelled as an Ornstein–Uhlenbeck process (represented as a solid line).*

where

$$R_i^*(\Delta) = \frac{1}{\beta_i^2} (\beta_i \Delta + e^{-\beta_i \Delta} - 1).$$

This implies that

$$\text{Var}(N_n) - \text{E}(N_n) = \sum_{i=1}^k \frac{\varpi_i}{\beta_i^2} (\beta_i \Delta + e^{-\beta_i \Delta} - 1),$$

where $\varpi_i = 2\omega^2 w_i$. The different decay rates could then represent the persistence of different types of shocks to the intensity level.

7.2 Price changes

In Section 6 we studied a simple model which did allow for serial dependence between price changes in that they were assumed to follow a first order moving average process. We furthermore assumed that the arrivals were

stochastically independent of the price changes. In this setting we have that the variance of p_n can be estimated by:

$$\widehat{\text{Var}}(p_n) = \hat{\sigma}_Z^2 \left[\hat{\lambda}_1 \Delta \{1 + 2\hat{\rho}_Z(1)\} - 2\hat{\rho}_Z(1) (1 - e^{-\hat{\lambda}_1 \Delta}) \right], \quad (7.1)$$

where $\hat{\sigma}_Z^2$ is the estimate of the variance of the price changes, $\hat{\rho}_Z(1)$ is the estimate of the first order correlation and $\hat{\lambda}_1$ is the estimate of the mean intensity when $\Delta = 1$.

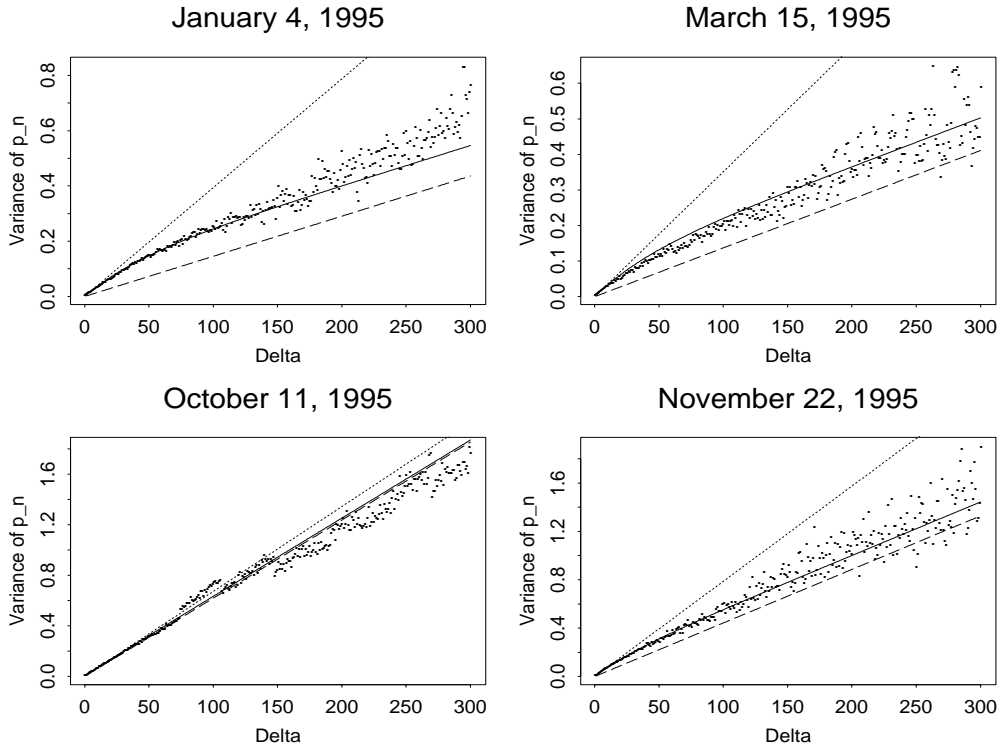


Figure 9: *The observed relationship between $\text{Var}(p_n)$ and Δ (represented as a \cdot) and the estimated curve resulting from the assumption that the variance can be described by (7.1) (represented as a solid line). The approximations given by 6.5 and 6.4 are plotted as a dotted and a dashed line.*

In Figure 9 we have plotted $\text{Var}(p_n)$ for $1 \leq \Delta \leq 300$ and the estimated curve resulting from the assumption that the variance can be described by (7.1), again for four randomly chosen days. The estimate from 7.1 does not seem too far off, but it is obvious that we have to build a more sophisticated model for the $\{Z_t\}$ process (this is the topic of Rydberg and Shephard (1998)). We have also plotted the approximations given by (6.5) and (6.4). From the graph it is seen that (6.5) works reasonably well for $\Delta < 20$ and that (6.4) seriously underestimates the variance when $-\hat{\rho}_Z(1)$ is large.

8 Conclusion

In this paper we have set out a framework for the study of the trade-by-trade price movements of speculative assets traded on major stock markets. The emphasis has been on the discrete nature of prices, irregularly spaced trading and returns (price movements) over intervals of length Δ . We have used compound processes as the basis of our modelling, with trading occurring at times determined by a Cox process. We studied the use of signal extraction methods for estimating the time varying intensity of this process. In Section 7.1 of the paper we studied the implications of this style of model for the dynamics of price changes in the market. The empirical evidence suggests that we are on the right track but that the models have to be some what more sophisticated. Further development of the structure for the arrival rate process was hinted at in Section 3.2 and is developed in Rydberg and Shephard (1999). Also the model for the price changes only included the most dominant feature, the first order moving average term, this is extended in Rydberg and Shephard (1998), where also other explanatory variables, such as duration and volume, are included.

9 Acknowledgments

Tina Rydberg thanks the Danish Natural Science Research Council for their financial support through a post-doctoral fellowship. Both authors are grateful for support from The Centre for Analytical Finance, Aarhus, Denmark. The initial draft of this paper was written in December 1997 and presented at a seminar at Oxford University that month and at the Workshop on Mathematical Finance, University of Bremen, Germany, February 1998. Subsequently it was presented at numerous other international meetings. We thank Richard Spady for various helpful comments on this line of research, Richard L Smith and Bent Nielsen for detailed comments on an earlier draft, while we have a big debt to Ole Barndorff-Nielsen for his work on SV models (with Neil Shephard) which allowed us to better understand the behaviour of $\{N_n\}$. We alone are responsible for any errors in the paper.

References

- Ait-Sahalia, Y. (1996). Nonparametric pricing of interest rate derivative securities. *Econometrica* 64, 527–560.
- Ammeter, H. (1948). A generalization of the collective theory of risk in

- regard to fluctuating basic probabilities. *Skand. AktuarTidskr.* 31, 171–198.
- Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *J. Empirical Finance* 4, 115–58.
- Andersen, T. G. and T. Bollerslev (1998). Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *J. Finance* 53, 219–265.
- Barndorff-Nielsen, O. E. and N. Shephard (1998). Aggregation and model construction for volatility models. Unpublished discussion paper: Nuffield College, Oxford.
- Barndorff-Nielsen, O. E. and N. Shephard (1999). Non-Gaussian OU based models and some of their uses in financial economics. Unpublished discussion paper: Nuffield College, Oxford.
- Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *J. Political Economy* 81, 637–654.
- Bollerslev, T. (1986). Generalised autoregressive conditional heteroskedasticity. *J. Econometrics* 51, 307–327.
- Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. In R. F. Engle and D. McFadden (Eds.), *The Handbook of Econometrics, Volume 4*, pp. 2959–3038. North-Holland.
- Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton, New Jersey: Princeton University Press.
- Clark, P. K. (1973). A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica* 41, 135–156.
- Cox, D. R. (1972). The statistical analysis of dependencies in point processes. In P. A. P. Lewis (Ed.), *Symposium on Point Processes*, pp. 55–66. New York: Wiley and Sons.
- Cox, D. R. and H. D. Miller (1965). *The Theory of Stochastic Processes*. London: Chapman & Hall.
- Cox, D. R. and D. Oakes (1984). *Analysis of Survival Data*. London: Chapman & Hall.
- Dothan, M. U. (1990). *Prices in Financial Markets*. New York: Oxford University Press.
- Duffie, D. (1992). *Dynamic Asset Pricing theory*. New Jersey: Princeton University Press.

- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica* 50, 987–1007.
- Engle, R. F. and J. R. Russell (1998). Forecasting transaction rates: the autoregressive conditional duration model. *Econometrica* 66, 1127–1162.
- Frey, R. and W. Runggaldier (1998). Nonlinear filtering techniques for estimation and risk management in partially observed stochastic volatility models. Unpublished paper: Mathematics Department, EZH.
- Ghysels, E., A. C. Harvey, and E. Renault (1996). Stochastic volatility. In C. R. Rao and G. S. Maddala (Eds.), *Statistical Methods in Finance*, pp. 119–191. Amsterdam: North-Holland.
- Ghysels, E., J. Jasiak, and C. Gouriou (1998). Stochastic volatility duration models. Unpublished paper: Pennsylvania State University. Presented at Second international conference on high frequency data in finance, Zurich, Switzerland, April.
- Grandell, J. (1991). *Aspects of Risk Theory*. Berlin, Heidelberg, New York: Springer-Verlag.
- Grandell, J. (1997). *Mixed Poisson Processes*, Volume 77 of *Monographs on Statistics and Applied Probability*. London: Chapman and Hall.
- Green, P. and B. W. Silverman (1994). *Nonparameteric Regression and Generalized Linear Models: A Roughness Penalty Approach*. London: Chapman & Hall.
- Guillaume, D. M., M. M. Dacorogna, R. R. Dave, U. A. Muller, R. B. Olsen, and O. V. Pictet (1997). From the bird’s eye view to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics* 2, 95–130.
- Hasbrouck, J. (1996). Modeling Market Microstructure Time Series. In C. R. Rao and G. S. Maddala (Eds.), *Statistical Methods in Finance*, pp. 647–692. Amsterdam: North-Holland.
- Hasbrouck, J. (1999a). The dynamics of discrete bid and ask quotes. *J. Finance* 54. Forthcoming.
- Hasbrouck, J. (1999b). Trading fast and slow: Security market events in real time. Unpublished paper: Stern Business School, New York University.
- Hull, J. and A. White (1988). An analysis of the bias in option prices caused by stochastic volatility. *Advances in Futures and Options Pricing Research* 3, 29–61.

- Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge: Cambridge University Press.
- Meddahi, N. and E. Renault (1996). Aggregation and marginalization of GARCH and stochastic volatility models. Unpublished paper: CREST-INSEE.
- Meddahi, N., E. Renault, and B. Werker (1998). Modelling high frequency data in continuous time. Unpublished paper: CIRANO, CRDE, Montreal University.
- Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3, 125–144.
- O’Hara, M. (1995). *Market Microstructure Theory*. Oxford: Blackwell Publishers.
- Pitt, M. K. and N. Shephard (1999). Filtering via simulation based on auxiliary particle filters. *J. American Statistical Association* 94. Forthcoming.
- Rogers, L. C. G. and O. Zane (1998). Designing and estimating models of high frequency data. Unpublished paper: Department of Mathematics, University of Bath. Presented at Workshop on Mathematical Finance, University of Bremen, Germany, February.
- Russell, J. R. and R. F. Engle (1998). Econometric analysis of discrete-valued, irregularly-spaced financial transactions data using a new autoregressive conditional multinomial models. Unpublished paper: Graduate School of Business, University of Chicago. Presented at Second international conference on high frequency data in finance, Zurich, Switzerland, April.
- Rydberg, T. H. and N. Shephard (1998). Dynamics of trade-by-trade price movements: decomposition and models. Working paper, Nuffield College, Oxford. Presented at Workshop on Econometrics and Finance, Isaac Newton Institute, Cambridge University, October 1998.
- Rydberg, T. H. and N. Shephard (1999). BIN models for trade-by-trade data. Modelling the number of trades in a fixed interval of time. Working paper, Nuffield College, Oxford.
- Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In D. R. Cox, D. V. Hinkley, and O. E. Barndorff-Nielsen (Eds.), *Time Series Models in Econometrics, Finance and Other Fields*, pp. 1–67. London: Chapman & Hall.
- Stein, E. M. and J. Stein (1991). Stock price distributions with stochastic volatility: an analytic approach. *Rev. Financial Studies* 4, 727–752.

- Snyder, D. L. and M. I. Miller (1991). *Random Point Processes in Time and Space* (2 ed.). New York: Springer-Verlag.
- Wold, H. (1948). On stationary point processes and Markov chains. *Skand. AktuarTidskr.* 31, 229–240.