Modelling trade-by-trade price movements of multiple assets using multivariate compound Poisson processes

Tina Hviid Rydberg
Nuffield College, Oxford OX1 1NF, UK.
`tina.rydberg@nuf.ox.ac.uk`

Neil Shephard
Nuffield College, Oxford OX1 1NF, UK.
`neil.shephard@nuf.ox.ac.uk`

Preliminary version, June 1999

Abstract

In this paper we extend Rydberg-Shephard's activity, direction and size decomposition of trade-by-trade price movements to the multivariate case. We illustrate our ideas using a bivariate modelling problem — modelling the evolution of the prices of Ford and GM shares. Throughout we use the continuous record of trades made in the first five months of 1997 on the New York Stock Exchange (NYSE).

Keywords: Activity, AREX, autologistic, decomposition, directions, durations, forecasting, GLARMA, multivariate mdels, negative binomial distribution, prediction decomposition, size, transactions data.

1 INTRODUCTION

Usually one of the intentions of holding a portfolio of financial assets is to diversify risk, see e.g. Markowitz (1959), Sharpe (1963) and Lintner (1965) or, for a textbook exposition see Campbell, Lo, and MacKinlay (1997, Ch. 5 and 6). This is carried out by exploiting the co–dependence between the returns on the assets in the portfolio. To give appropriate measures of risk in this problem entails the development of adequate statistical models to describe the behaviour of multivariate time series of financial assets.

The standard models used in the context of portfolio allocation are based on low frequency data and do not encompass the structure of trade–by–trade dynamics. Hence, if we would like to evaluate the intra day risk of holding a portfolio of assets we need to build a multivariate model which does take the structure of trade–by–trade data into account. The use of such a model would also enable us to understand what impact trades in similar types of shares have on each other, or how transactions in a derivative and the underlying influence the price of one another.

This paper proposes a multivariate extension of the framework and decomposition developed in Rydberg and Shephard (1998b) and Rydberg and Shephard (1998a). Such an extension is essential if we are to study the dynamics of portfolios trade–by–trade, however the simple structure can be used at all aggregate levels such as days and months. The advantage of our approach is that it is very flexible and it is relatively easy to carry out the estimation of the models.

In section 2 we introduce the basic framework of a multivariate compound process, before extending the activity–direction–size (ADS) decomposition of price movements to the multivariate case. Section 4 will discuss specific modelling issues and apply the methods to share prices of Ford and GM traded on the NYSE the first five months of 1997. As a by product of our empirical work we suggest a model which involves an autoregression in the explanatory variables (AREX). Section 5 outlines various possible extensions, to our modelling framework, while Section 6 concludes.

2 MULTIVARIATE COMPOUND POISSON PROCESSES

2.1 Compound Poisson processes

In this paper we generalise the univariate continuous time model proposed by Rydberg and Shephard (1998b) for the evolution of the prices and times at which transactions are carried out on a common asset traded on a stock exchange trading only with fixed tick sizes. In almost all financial markets prices are restricted to such lattice points. For example this is the case for stocks traded at the New York Stock Exchange (NYSE) which can be seen from Figure 1 which displays the price of all NYSE based transaction on the Ford and the GM stock on February 5, 1997. In 1997 the tick size was 1/8th of a dollar until June 24th, 1997 where it changed to 1/16th of a dollar as an interim step towards trading in decimals, see New York Stock Exchange (1997)[1]. See O'Hara (1995, Ch. 1) for an introduction to markets and market making on the NYSE.

Our model will be based on $p(u)$ a $K \times 1$ vector of prices, which we will norm so that the

[1]Initially we studied a period which covered the first two quaters of 1997. This data set contained about a week of transactions where the minimum tick size was 1/16. We decided to discard this period and only study the first five months since we found that the dynamics of the price changes changed dramatically when the tick size changed. This is supported by economic theory which predicts an increase in liquidity by lowering the tick size since transaction costs decrease.

tick sizes are one. Then we propose the structure

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

where Z_t is a $K \times 1$ vector and $N(u)$ is the number of trades in any of the K assets made up to time u. In practice most of the elements of each of the Z_t will be exactly zero — indeed a trade which did not move any of the prices would result in a vector of zeros. More commonly a single element of Z_t will be non-zero, indicating a movement in a single stock, perhaps up or down one tick. When $K = 1$ we get precisely the model suggested in Rydberg and Shephard (1998b).

Rydberg and Shephard (1998b) modelled $N(u)$ as a counting process with new arrivals being generated by a Cox process[2], that is a Poisson process with a random intensity. 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. Some of the econometric issues which arise with unequally spaced financial data were discussed at length in Engle (1996). An alternative way of modelling the counting process would be via the intensity process, this is convenient if we would like to study returns over fixed intervals of time. A very flexible class of linear models for the intensity are the BIN models suggested in Rydberg and Shephard (1999).

In this framework we again think of modelling $N(u)$ as a counting process with new arrivals being generated by a Cox process whose intensity is adapted to a filtration involving, at least, the whole history of the trading times of the different assets and the price movements at those trades. More generally we could also think of including the times of quote updates. Potentially this is a complicated empirical modelling problem, however it raises no new econometric issues as it is simply a matter of a modelling a time series of intensities or equivalently the time series of durations.

Figure 1 displays the evolution of the price of two similar shares Ford and GM traded on the NYSE on May 30, 1997. We can see that Ford is more heavily traded than GM. but that the two stocks tend to move in common ways.

In Table 1 we have shown the transition probabilities between different pairs of (Z_{t-1}, Z_t). We can see that there is an increased probability of seeing Ford trades followed by another trade in Ford and the same holds for GM. It should be noted that the upper left hand corner which depicts transitions from one Ford trade to another and the lower right hand corner which are transitions between GM trades are not equal to the univariate model for each of these shares. In between two Ford shares we could see a random number of trades in GM, hence a univariate model for Ford deduced from the multivariate Markov model based on Table 1 would not be Markovian. The actual transition probabilities can be calculated as follows

$$\Pr\left(Z_t \text{ is a Ford trade} \middle| \text{ the last Ford trade} = i\right) = A_i + B\left(\sum_{j=0}^{\infty} D^j\right) C_i,$$

where A, B, C and D denotes the four corners of the matrix given in Table 1. A is the upper left hand corner, B is the upper right hand corner and C the lower left hand corner.

2.2 Univariate *ADS* decomposition

Rydberg and Shephard (1998a) introduced a new way of analysing price movements in a single asset in order to model the Z_t which is denoting the change in ticks. Then let $Z_t \in I$ be an

[2]An 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.

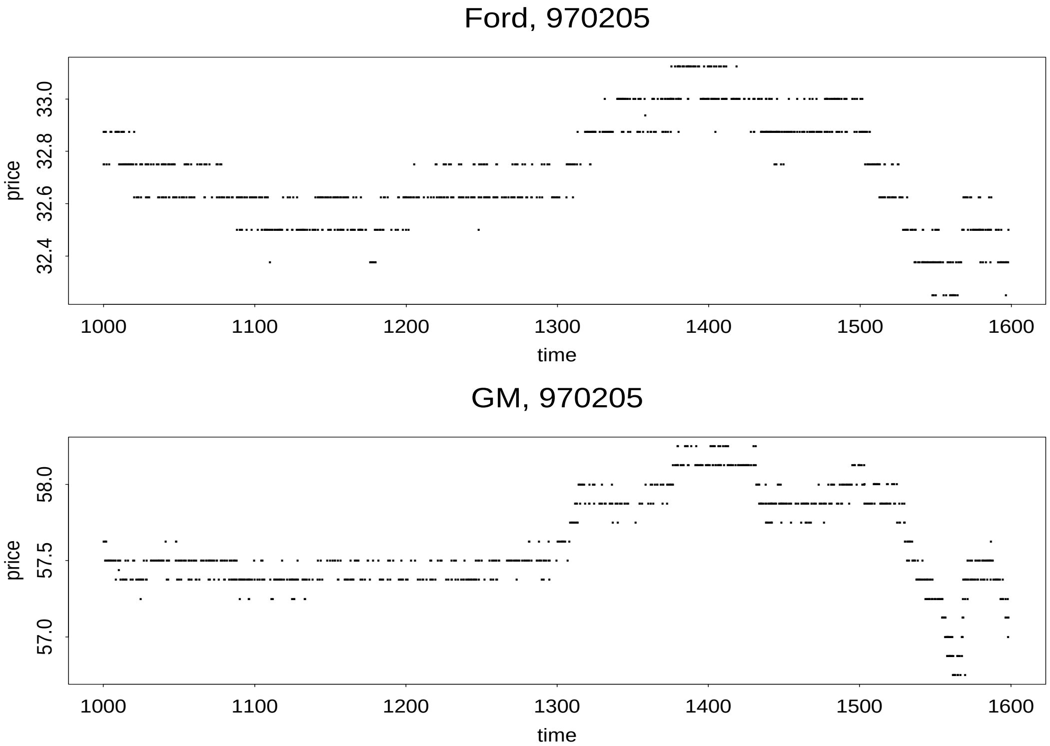

Figure 1: *Plot of all traded Ford and GM prices at the New York stock exchange on February 5, 1997. A trade is represented as a dot.*

integer process and $\mathcal{F}_t = \sigma(Z_s : s \leq t)$ be its natural sigma algebra or filtration (sequential information set). Hence $\mathcal{F}_{t-1}$ contains all the past integers from time 1 to $t-1$. Then we are interested in the joint distribution of the tick movements which are given by

$$\Pr(Z_1, ..., Z_n|\mathcal{F}_0) = \prod_{t=1}^{n} \Pr(Z_t|\mathcal{F}_{t-1}),$$

using a prediction decomposition. The problem will be specifying $\Pr(Z_t|\mathcal{F}_{t-1})$.

Prior to Rydberg and Shephard (1998a), Russell and Engle (1998) have proposed using a conditional multinomial style model for specifying $\Pr(Z_t|\mathcal{F}_{t-1})$, in a sense generalising previous work of Hausman, Lo, and MacKinlay (1992) on multivariate probit models for transaction data. Also Hasbrouck (1996) has proposed a class of dynamic latent variable models for efficient prices and traders cost which uses economically motivated truncation to enforce prices to live on lattices [3].

Potentially the distribution of $\Pr(Z_t|\mathcal{F}_{t-1})$ can be quite complicated. Rydberg and Shephard (1998a) break down the pieces of Z_t into bits and then model these sequentially. Note that there is no loss of information in this "*ADS*" decomposition.

To carry out the *ADS* decomposition define the $t-th$ price move as

$$Z_t = A_t D_t S_t. \tag{2}$$

[3] Other work which relates to this topic includes a paper by Meddahi, Renault, and Werker (1998) who combined GARCH models for price returns with autoregressive duration style models for the times between trades. The paper by Darolles, Gourieroux, and Le Fol (1998) is much closer to the framework looked at in Rydberg and Shephard (1998b) and Rogers and Zane (1998). The Darolles, Gourieroux, and Le Fol (1998) paper uses the structure (1) but assumes the Z_t process is Markov living on the points $-1, 0, 1$. This modelling assumption is combined with a reduction in the dataset by using quote data to allow them only to model buys from the market maker — which in turn makes the assumption that the price movements are at most one tick in absolute value more realistic.

			Z_{t-1}								
			Ford				GM				
			-1	0	1	$\sum$	-1	0	1	$\sum$	$\sum$
Z_t	Ford	-1	0	0.014	0.008	0.022	0.002	0.019	0.002	0.023	0.045
		0	0.018	0.23	0.019	0.268	0.013	0.137	0.013	0.166	0.431
		1	0.008	0.014	2×10^{-5}	0.022	0.002	0.018	0.002	0.020	0.045
		$\sum$	0.027	0.258	0.027	0.312	0.018	0.173	0.018	0.209	
	GM	-1	0.002	0.019	0.002	0.023	0.000	0.014	0.007	0.021	0.044
		0	0.015	0.137	0.014	0.163	0.017	0.19	0.019	0.226	0.392
		1	0.002	0.017	0.002	0.022	0.010	0.013	0.000	0.023	0.043
		$\sum$	0.017	0.175	0.017	0.210	0.027	0.217	0.026	0.271	
		$\sum$	0.045	0.431	0.045	0.521	0.044	0.392	0.043	0.479	1

Table 1: *Transition probabilities of trades in Ford and GM.*

We will let A_t take on only two values: $0, 1$. When $A_t = 0$, we define for notational convenience (there is no loss in doing this), $D_t = S_t = 0$. Otherwise, when $A_t = 1$ we let D_t and S_t live on the structure:

$$D_t = -1, 1 \text{ and } S_t = 1, 2, \dots$$

Thus we have that if A_t is zero then Z_t must be zero. This means the price does not move or, in other words, is *In-Active*. If $A_t = 1$ then there are *Active* price movements. The non–zero price movement must be $Z_t = D_t S_t$. Likewise, if we assume $A_t = 1$, then D_t controls the *Direction* of the price move. If $D_t = 1$ the price moves upwards, else it moves downwards. Finally, S_t controls the *Size* of price movements. This suggests the decomposition of price movements into

$$\Pr(Z_t = 0 | \mathcal{F}_{t-1}) = \Pr(A_t = 0 | \mathcal{F}_{t-1}),$$

while for $z_t \neq 0$

$$\begin{aligned} \Pr(Z_t &= z_t | \mathcal{F}_{t-1}) = \Pr(A_t = 1 | \mathcal{F}_{t-1}) \\ &\times \left\{ \begin{array}{c} \Pr(S_t = z_t | \mathcal{F}_{t-1}, A_t = 1, D_t = 1) \Pr(D_t = 1 | \mathcal{F}_{t-1}, A_t = 1) + \\ \Pr(S_t = -z_t | \mathcal{F}_{t-1}, A_t = 1, D_t = -1) \Pr(D_t = -1 | \mathcal{F}_{t-1}, A_t = 1) \end{array} \right\}. \end{aligned}$$

The implication of this decomposition is that there are exactly three pieces of modelling to carry out

- $\Pr(A_t | \mathcal{F}_{t-1})$ — a binary process on $\{0, 1\}$ modelling *activity* (the price moves or not).
- $\Pr(D_t | \mathcal{F}_{t-1}, A_t = 1)$ — another binary process on $\{-1, 1\}$ modelling the *direction* of the price moves.
- $\Pr(S_t | \mathcal{F}_{t-1}, A_t = 1, D_t)$ — a process on the strictly positive integers modelling the *size* of price moves.

Potentially each of these models has to be constructed separately — basing each on the complete history of the Z_t process. This is carried out in Rydberg and Shephard (1998a).

Our multivariate extension has two parts: we have extended the compound representation of prices on a lattice to a vector of prices; we now extend the ADS decomposition (2) to an $MADS$ decomposition.

2.3 $MADS$: multivariate ADS decomposition

The major innovation we present in this paper is in modelling the joint density of the vector process $\{Z_t\}$. We follow the univariate development and begin by being first only interested in the joint distribution of the movements which are given by

$$\Pr(Z_1, ..., Z_n|\mathcal{F}_0) = \prod_{t=1}^{n} \Pr(Z_t|\mathcal{F}_{t-1}),$$

using a prediction decomposition. The problem will then be to specify $\Pr(Z_t|\mathcal{F}_{t-1})$.

We now introduce a multivariate ADS ($MADS$) decomposition. For the j–th price process we write the price at the t–th trade as

$$Z_{j,t} = M_{j,t}A_{j,t}D_{j,t}S_{j,t}.$$

where $M_{j,t}$ is zero if the there is a trade of one of the K assets, but that trade did not involve the j–th asset. If the j–th asset was traded then $M_{j,t} = 1$. If $M_{j,t} = 0$ then we define $A_{j,t} = D_{j,t} = S_{j,t} = 0$, otherwise we use the definition of these variables given in Section 2.2. Such a decomposition loses no information and seems in keeping with the univariate analysis, while introducing only one layer of complexity to the modelling problem. That there is no loss of generality, which can be seen from the following

$$\Pr(Z_{j,t} = 0|\mathcal{F}_{t-1}) = \Pr(M_{j,t} = 0|\mathcal{F}_{t-1}) + \Pr(A_{j,t} = 0|\mathcal{F}_{t-1}, M_{j,t} = 1),$$

while for $z_{j,t} \neq 0$

$$\begin{aligned}\Pr(Z_{j,t} &= z_t|\mathcal{F}_{t-1}) = \Pr(A_{j,t} = 1|\mathcal{F}_{t-1}, M_{j,t} = 1) \\ &\times \left\{ \begin{array}{c} \Pr(S_{j,t} = z_{j,t}|\mathcal{F}_{t-1}, M_{jt} = 1, A_{j,t} = 1, D_{j,t} = 1) \\ \times \Pr(D_{j,t} = 1|\mathcal{F}_{t-1}, M_{j,t} = 1, A_{j,t} = 1) + \\ \Pr(S_{j,t} = -z_{j,t}|\mathcal{F}_{t-1}, M_{j,t} = 1, A_{j,t} = 1, D_{j,t} = -1) \\ \times \Pr(D_{j,t} = -1|\mathcal{F}_{t-1}, M_{j,t} = 1, A_{j,t} = 1) \end{array} \right\}.\end{aligned}$$

The implication of this decomposition is that there are exactly four pieces of modelling to carry out, where three of the pieces are already discussed in Rydberg and Shephard (1998a)

- $\Pr(M_t|\mathcal{F}_{t-1})$ — a binary vector process on $\{0, 1\}^K$ modelling *trade* (stock is traded or not). By construction at least one of the assets has to trade. There might be more than one asset trading at the same time.
- $\Pr(A_{j,t}|\mathcal{F}_{t-1}, M_{j,t} = 1)$ — a binary process on $\{0, 1\}$ modelling *activity* of the j–th asset (the price moves or not).
- $\Pr(D_{j,t}|\mathcal{F}_{t-1}, M_{j,t} = 1, A_{j,t} = 1)$ — another binary process on $\{-1, 1\}$ modelling the *direction* of the price moves.
- $\Pr(S_{j,t}|\mathcal{F}_{t-1}, M_{j,t} = 1, A_{j,t} = 1, D_{j,t})$ — a process on the strictly positive integers modelling the *size* of price moves.

2.4 An alternative approach

There are two obvious, conceptually different, ways of thinking about the structure of the price given by (1) in a multivariate context. Both of these, however, leads to the same mathematical model. The way we have chosen assigns the problem of deciding which asset is traded to the modelling of the price movements. Alternatively this choice could also be part of the arrival process $N(t)$, which would switch between assets controlled by a marked point process where the marks indicate which asset is traded.

3 GENERAL STRUCTURE OF THE MULTIVARIATE MODEL

3.1 A multinomial $MADS$ model

One way of thinking about the multinomial structure of our model for $\{M_t\}$ is in terms of the multinomial framework for the different jump sizes discussed in Russell and Engle (1998). However, here the multinomial distribution has ADS marginals, and the marginals models the individual assets. The state vector is given by M_t. This will lead to a multinomial logit structure (see e.g. McFadden (1984, Section 3.4)) of the following type:

$$\Pr(M_{j,t} = 1|\mathcal{F}_{t-1}) = p\left(\theta_{j,t}^{M}\right), \quad j = 1, ..., K,$$

where $\theta_t = (\theta_{1,t}, \theta_{2,t}, ..., \theta_{K,t})'$ and

$$p\left(\theta_{j,t}^{M}\right) = \frac{\exp\left(\theta_{j,t}^{M}\right)}{1 + \sum_{i=1}^{K}\exp\left(\theta_{i,t}^{M}\right)}, \quad j = 1, ..., K.$$

In practice this structure is not identified and so constraints are placed on θ_t. A typical situation would be to define $\theta_{1,t} = 0$ for all t, a solution followed by Russell and Engle (1998).

As in Russell and Engle (1998) we could now define a vector generalised linear autoregressive moving average (VGLARMA) type structure on $\theta_t^* = (\theta_{2,t}, \theta_{3,t}, ..., \theta_{K,t})'$, feeding in lagged values of $\{M_t\}$ using the variable given by (4). In particular if they define $v_t = (v_{2,t}, ..., v_{K,t})'$ with

$$v_{j,t} = \frac{M_{j,t} - p_{j,t}}{\sqrt{p_{j,t}\left(1 - p_{j,t}\right)}},$$

and

$$\theta_t^* = \alpha^* + g_t,$$

then model this system as

$$g_t = \sum_{i=1}^{p}\gamma_i g_{t-i} + \sigma v_t + \sigma\sum_{k=1}^{q}\delta_k v_{t-k},$$

where α^* is a vector, while $\{\gamma_i\}, \sigma$ and $\{\delta_k\}$ are $(K-1)\times(K-1)$ matrices. The only a priori constraint we might place on this structure is that σ should be lower triangular for identification.

If we were to adopt the framework suggested by Russell and Engle (1998) for the marginals, however, we would get a structure where each component of the M vector would again be a vector containing the possible states for the price changes.

3.2 A sequential $MADS$ model

An alternative to the multinomial approach would be to determine which stock was trading by sequentially conditioning on the previous shares not having been traded. This would give us the following expression for the distribution of M_t:

$$\Pr\left(M_t \middle| \mathcal{F}_{t-1}, \sum_{j=1}^{K} M_{j,t} = 1\right) =$$

$$\Pr\left(M_{1,t} \middle| \mathcal{F}_{t-1}, \sum_{j=1}^{K} M_{j,t} = 1\right) \times ... \times \Pr\left(M_{K,t} \middle| \mathcal{F}_{t-1}, \sum_{j=1}^{K} M_{j,t} = 1, M_{1,t}, ..., M_{K-1,t}\right)$$

notice that $M_{K,t} = 1 - \sum_{j=1}^{K-1} M_{j,t}$. This decomposition leaves us with an estimation problem which involves K binary series. In the case where we are only dealing with two shares $(K = 2)$ the multinomial and the sequential model are the same.

4 EMPIRICAL MODELLING

4.1 Data

The data that we will be using throughout this section are taken from the TAQ (trades and quotes) database from the New York Stock Exchange (NYSE). We will be using the first 5 months of 1997 and we have chosen to look at the Ford share and the GM share. We believe that in a preliminary study of multivariate models considering two stock in similar types of business makes sense, since they will be likely to react to the same type of information as well as firm specific information. How this is incorporated into the models is discussed in section 4.2.

In order not having to deal with the call auction which sets off trading at the opening of the exchange we only study the period 10.00–16.00, however a few data points before 10.00 have been used to generate the explanatory variables, such that the explanatory/lagged variables for trades just after 10.00 are actually from the same day and not the previous. This is helpful since we have concatenated the days. We believe this is justifiable due to the just mentioned exercise but also the fact that we are only looking at price changes and not the price level.

We have delete all trades which had an error code and we are only studying the trades performed at NYSE. This reduces the data set to a total of 92,014 observations. Of these 47,906 (52.1%) are trades in Ford, which gives us an average duration between trades of 46.89 seconds. Of the Ford trades only 8,256 are price moves different from 0 again of these only 2 are bigger than 1 tick. There are 44,108 trades in the GM share, which gives us an average duration between trades of 50.93 seconds. Of the GM trades only 9,157 of all these trades move the price and again of these 16 are greater then 1 tick.

That so few price changes are greater than 1 tick (0.125$) is partly because the market maker, by the NYSE, is compensated for making smooth price transitions. That is if some new information moves the price by one dollar the market maker has to trade (through his order book) at the intervening prices, so that, what is actually observed is a sequence of 1 or 2 tick moves. This, however, suggests that in order for the model to reproduce the heavy tailed returns, that are usually observed at lower frequencies, it is crucial to model the correlation structure, since this will be the driving force behind large returns.

The very few returns larger than 1 tick also suggests that extreme value analysis of the NYSE data recorded trade–by–trade does not make any sense in this market. An interesting comparison in this aspect would be to the trading in the FTSE100 future at the London Stock Exchange. This market is based on a recently introduced electronic trading system (SETS) (e.g. see Taylor, van Dijk, Franses, and Lucas (1999) where implications for the size of the bid–ask spread is discussed).

4.2 The multivariate component

When we are only dealing with two stocks we simply have to model a sequence of 0 and 1. Preliminary data analysis showed that the series for trades has quite a lot of memory, see Figure 2 (a) and (b). By Figure (b), which depicts the cumulative correlogram, we see that the memory of the process is about $2,000$ lags, which is on average 13.6 hours of trading, i.e. a little more than two days. This figure furthermore shows that both shares roughly exhibit the same intra day pattern and the same pattern of activity over the year (see Figure 2 (c) and (d)).

Since we have already seen that the switching between the two shares has a lot of memory we will use the generalized linear autoregressive moving average (GLARMA) models suggested in Shephard (1994). We will use a model which puts

$$\Pr(M_t = 1|\mathcal{F}_{t-1}) = p(\theta_t^M), \quad \text{where } \ p\left(\theta_t^M\right) = \frac{\exp\left(\theta_t^M\right)}{1+\exp\left(\theta_t^M\right)}$$

and

$$\theta_t^M = x_t'\beta + g_t, \tag{3}$$

with x_t being potential combinations and subsets of $\mathcal{F}_{t-1}^{M,A,D,S}$, and

$$g_t = \sum_{i=1}^{p} \gamma_j g_{t-i} + \sigma v_t + \sigma \sum_{i=1}^{q} \delta_i v_{t-i}, \quad \sigma > 0,$$

where

$$v_t = \frac{\left\{M_{t-1} - p\left(\theta_{t-1}^M\right)\right\}}{\sqrt{p\left(\theta_{t-1}^M\right)\left\{1 - p\left(\theta_{t-1}^M\right)\right\}}}. \tag{4}$$

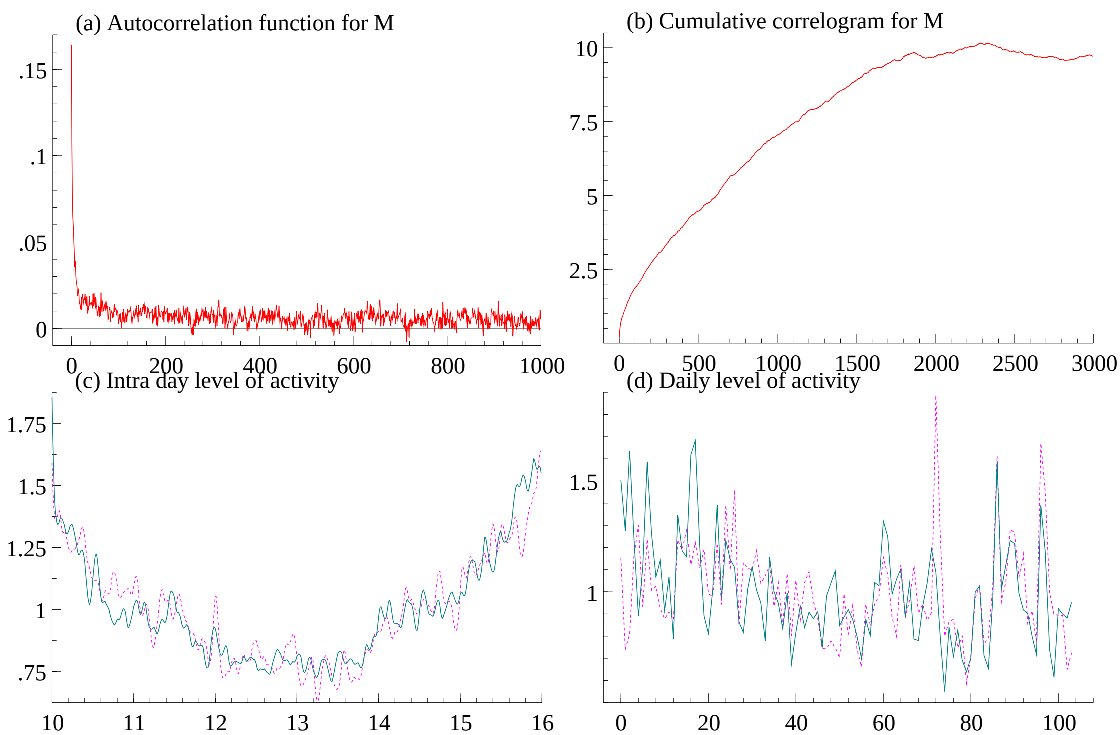

Figure 2: *(a) Shows the autocorrelation function for the M variable. (b) This is the accumulated correlation function. (c) The level of intensity over the day, the solid line is GM. (d) The daily level of activity over the five month period, the solid line is GM. The two graphs in (c) and (d) have been smoothed using a spline.*

Importantly $\{v_t\}$ is a Martingale difference sequence with a unit conditional variance. This style of model is adopted in Russell and Engle (1998) in their multinomial construction.

GLARMA models have some of the properties of ARMA models. This can be seen by focusing in on $\{v_t\}$, which has zero mean and unit conditional and unconditional variance. The implication of this is that $\{g_t\}$ is a linear ARMA process driven by a weak white noise error term. Hence it is covariance stationary and invertible if this model obeys the usual stationarity and invertability constraints on the polynomials $1 - \sum_{i=1}^{p} \gamma_i L^i$ and $1 + \sum_{i=1}^{q} \delta_i L^i$. The implication is that the autocorrelation function of $\{g_t\}$ and the corresponding unconditional variance can be found using standard results on covariance stationary linear processes.

For our empirical work we need to define the "individual trading time" for the individual stocks, since variables based on this timing will show to be very significant. So, let

$$B_{j,t} = \sup_{s_1,\ldots,s_h} \left\{ \begin{array}{rl} s_1 < t; & M_{j,s_1} = 1 \\ s_2 < s_1; & M_{j,s_2} = 1 \\ \vdots & \\ s_h < s_{h-1}; & M_{j,s_h} = 1 \end{array} \right\}. \tag{5}$$

The $B_{j,t}$ vector is $h \times 1$ and contains the times at which the last h trades occurred in the j–th asset. E.g. $A_{B_{1,t}}$ contains the activity variable for the last h trades in share 1. Furthermore, let

$$T_{j,t} = \sup_{s_1,\ldots,s_h} \left\{ \begin{array}{rl} s_1 < t; & A_{j,s_1} = 1 \\ s_2 < s_1; & A_{j,s_2} = 1 \\ \vdots & \\ s_h < s_{h-1}; & A_{j,s_h} = 1 \end{array} \right\}. \tag{6}$$

The $T_{j,t}$ vector is $h \times 1$ and contains the times at which the last h active prices occurred in the j–th asset. We call this concept of a time scale "activity time". The vector $D_{T_{1,t}}$ will contain the direction of the last h active trades in share 1, i.e. the sign of trades that changed the price of share 1.

In our numerical work we have imposed that $\{g_t\}$ obeys the covariance stationarity and invertability constraints on the GLARMA. We impose these conditions by using the partial autocorrelations $\{\rho_i, i = 1, ..., p\}$ and the inverse partial autocorrelations $\{\bar{\rho}_i, i = 1, ..., q\}$ (see Barndorff-Nielsen and Schou (1973) and Jones (1987)). We numerically maximised the corresponding likelihood using analytic first derivatives and the BHHH algorithm[4], using the initial conditions that $g_0, ..., g_{-p+1}, v_0, ..., v_{-q+1}$ are all set to zero.

We use the normalised sequence for the fit of a binary process $\{M_t\}$ by constructing

$$u_t = \frac{M_t - p\left(\theta_t^M\right)}{\sqrt{p\left(\theta_t^M\right)\left\{1 - p\left(\theta_t^M\right)\right\}}},$$

for, if the model was correctly specified, the $\{u_t\}$ should be uncorrelated with zero mean and unit conditional (and unconditional) variance. The $\{u_t\}$ could then be used inside a Box-Pierce statistic as a measure of residual dependence.

One of the main ideas of looking at two shares is to investigate how co–dependent share prices for firms in similar types of business are. This is the reason we have chosen to look at Ford and GM which are both large automobile industries. Therefore we have the types of activity variables in the analysis. $A_{B_{1,t}}$ which indicates if a trade in Ford was active or not and A_t indicates if a trade in general was active or not. In this way we have also modelled $A_{B_{2,t}}$ since this is just $A_{B_{2,t}} = A_t - A_{B_{1,t}}$. From Table 2 it seems that knowing that the last trade was an active trade in Ford slightly increases the probability of the next trade being a trade in Ford and vice versa. The data also show some asymmetries. Since the directions in activity time are significant and positive for both shares it implies that a positive price movement increases the probability of trading in that particular stock.

[4] A plain BHHH method was used mapping the real variables being maximised into partial autocorrelations and inverse partial autocorrelations using the transform $x/(1 + |x|)$. No interventions or numerical problems were encountered.

Variable	Coef.	StR. Err.	Variable	Coef.	Str. Err.
ρ_1	0.999864		$A_{B_{1,(t,1)}}$	0.201	(0.025)
ρ_2	-0.686636		$A_{B_{1,(t,2)}}$	0.072	(0.022)
ρ_3	-0.090554		$A_{B_{1,(t,3)}}$	0.080	(0.021)
ρ_4	-0.090542		$A_{B_{1,(t,4)}}$	0.033	(0.021)
ρ_5	-0.059069		$D_{T_{1,(t,2)}}$	-0.088	(0.014)
$\bar{\rho}_1$	-0.992188		$D_{T_{1,(t,3)}}$	-0.036	(0.014)
Constant	0.089	(0.032)	A_{t-1}	-0.214	(0.021)
			A_{t-2}	-0.060	(0.020)
			A_{t-3}	-0.046	(0.019)
			A_{t-4}	-0.065	0.019
			$D_{T_{2,(t,1)}}$	-0.053	0.009
Q	$T\sum_{j=1}^{Q} r_j^2$		Log-like $=-61,598.$		
20	10.52	(31.41)			
100	99.56	(124.34)			
1500	$1,570.78$	$(1,591.21)$			

Table 2: *Estimation for the trades component using a GLARMA model. Parameter estimates of the fitted GLARMA(5,2) model, using a maximum likelihood criteria and the (4) error term. The figures in brackets are the standard errors on the regressors computed using the GLARMA model. Model order selected using AIC. r_j denotes the series correlation coefficient at lag j for the the standardised residuals u_t. The figures in brackets are corresponding 95 percentage points on the χ_Q^2 distribution to give a rough benchmark.*

4.3 The activity of prices

Our initial parametric model for $\Pr(A_{j,t}|\mathcal{F}_{t-1}, M_{j,t}=1)$ will be a GLARMA model based on $\mathcal{F}_t = \mathcal{F}_t^{M,A,D,S}$. Recall for a GLARMA we write

$$\Pr(A_{j,t}=1|\mathcal{F}_{t-1}, M_{j,t}=1) = p\left(\theta_{j,t}^A\right), \qquad \text{where} \qquad p\left(\theta_{j,t}^A\right) = \frac{\exp\left(\theta_{j,t}^A\right)}{1+\exp\left(\theta_{j,t}^A\right)},$$

and

$$\theta_{j,t}^A = x_{j,t}'\beta_j + g_{j,t}, \qquad \text{where} \qquad g_{j,t} = \sum_{i=1}^{p_j}\gamma_{j,i}g_{j,t-i} + \sigma_j v_{j,t} + \sigma_j\sum_{k=1}^{q_j}\delta_{j,k}v_{j,t-k}, \qquad \sigma_j>0, \quad (7)$$

where

$$v_{j,t} = \frac{\left\{A_{j,t-1} - p\left(\theta_{j,t-1}^A\right)\right\}}{\sqrt{p\left(\theta_{j,t-1}^A\right)\left\{1-p\left(\theta_{j,t-1}^A\right)\right\}}}. \quad (8)$$

with x_t being potential combinations and subsets of $\mathcal{F}_{t-1}^{D,S,A}$. This model structure was introduced by Cox (1958) (see Cox and Snell (1989) for an exposition) and has some significant advantages[5].

[5]A standard latent variable interpretation of these models writes

$$\Pr(Y_t=1|X_t) = \Pr(\theta_t + U > 0) = p(\theta_t),$$

where U has a logistic distribution with parameters $(0,1)$. Replacing the logistic distribution by a standard normal produces a probit models.

The log–likelihood for the autologistic is concave and so numerical optimisation is completely straightforward, allowing standard logistic regression software to be used to rapidly and reliably fit the model (e.g. McCullagh and Nelder (1989)).

We use the normalised sequence for the fit of a binary process $\{A_{j,t}\}$ by constructing

$$u_{j,t} = \frac{A_{j,t} - p\left(\theta^A_{j,t}\right)}{\sqrt{p\left(\theta^A_{j,t}\right)\left\{1 - p\left(\theta^A_{j,t}\right)\right\}}},$$

for, if the model was correctly specified, the $\{u_{j,t}\}$ should be uncorrelated with zero mean and unit conditional (and unconditional) variance. The $\{u_{j,t}\}$ could then be used inside a Box–Pierce statistic as a measure of residual dependence.

The results from the estimation are shown in Tables 3 and 4. For both shares there seem to be no dependence on the level of activity in the other share, hence none of the explanatory variables include the other share. The result of a "from general to simple" approach results in rather parsimonious models for both stocks, there seem to be slightly more memory in the activity series for GM than that for Ford, hence the chosen model for the activity series of Ford is a GLARMA(2,1) while we choose a GLARMA(4,1) for the GM. However, the way in which past directions influence the level of activity is opposite for the two shares. For the Ford share the parameters for the two past directions in activity time are positive, hence upward movements increase the probability of having a price chance different from zero, while for the GM share the sign of the previous direction in activity time has a negative effect.

Variable	Coef.	StR. Err.	Variable	Coef.	Str. Err.
ρ_1	0.981023		April	-0.253	(0.079)
ρ_2	-0.48449		$D_{T_{1,(t,1)}}$	0.063	(0.023)
$\bar{\rho}_1$	-0.712929		$D_{T_{1,(t,2)}}$	0.055	(0.023)
			Constant	-1.707	(0.035)
Q	$T\sum_{j=1}^{Q} r_j^2$		Log-like = -20,279.		
20	16.20	(31.41)			
100	99.40	(124.34)			
1500	$1,428.24$	$(1,591.21)$			

Table 3: *Estimation for the activity component for Ford using a GLARMA model. Parameter estimates of the fitted GLARMA(2,1) model, using a maximum likelihood criteria and the (4) error term. The figures in brackets are the standard errors on the regressors computed using the GLARMA model. Model order selected using AIC. r_j denotes the series correlation coefficient at lag j for the the standardised residuals u_t. The figures in brackets are corresponding 95 percentage points on the χ^2_Q distribution to give a rough benchmark.*

4.4 The direction of price changes

4.4.1 Autologistic model

An important feature of our decomposition is that we are now able to focus on a model of the directions of the price changes, given that the price has changed: $\Pr(D_{j,t}|\mathcal{F}_{t-1}, M_{j,t} = 1, A_{j,t} = 1)$. This maybe helpful for high frequency traders who are interested in the movement of prices over very small periods of time.

Variable	Coef.	StR. Err.	Variable	Coef.	Str. Err.
ρ_1	0.999223		$D_{T_{2,(t,1)}}$	-0.034	(0.013)
ρ_2	-0.623985		February	-0.332	(0.090)
ρ_3	-0.13421		Constant	-1.545	(0.037)
ρ_4	-0.106838				
$\bar{\rho}_1$	-0.981715				
Q	$T\sum_{j=1}^{Q} r_j^2$		Log-like $= -19,881$.		
20	13.19	(31.41)			
100	80.01	(124.34)			
1500	1,496.5	(1591.21)			

Table 4: *Estimation for the activity component for GM using a GLARMA model. Parameter estimates of the fitted GLARMA(4,1) model, using a maximum likelihood criteria and the (4) error term. The figures in brackets are the standard errors on the regressors computed using the GLARMA model. Model order selected using AIC. r_j denotes the series correlation coefficient at lag j for the the standardised residuals u_t. The figures in brackets are corresponding 95 percentage points on the χ^2_Q distribution to give a rough benchmark.*

We will use an autologistic model, where the outcome variable will live on the support $\{-1, 1\}$, rather than $\{0, 1\}$. To start out with we will let $\mathcal{F}_t = \mathcal{F}_t^{M,A,D,S}$.

This gives us an autologistic model for

$$\Pr\left(D_{j,t} = 1 | \mathcal{F}_{t-1}, A_{j,t} = 1, M_{j,t} = 1\right) = p\left(\theta_{j,t}^D\right), \qquad \text{where} \qquad p\left(\theta_{j,t}^D\right) = \frac{\exp\left(\theta_{j,t}^D\right)}{1 + \exp\left(\theta_{j,t}^D\right)},$$

and so

$$\Pr\left(D_{j,t} = -1 | \mathcal{F}_{t-1}, A_{j,t} = 1, M_{j,t} = 1\right) = \frac{1}{1 + \exp\left(\theta_{j,t}^D\right)}.$$

After testing out insignificant explanatory variables we end up with directions and the price changes themselves being significant.

Recall the concept of "activity time" defined in 6. In Rydberg and Shephard (1998a) we found that this concept of time was extremely significant statistically when modelling the direction of a univariate series. $D_{Tj,t}$ will be a vector of the signs of the last k price changes different from 0 in the j–th asset. We refer to the i'th element of this vector as $D_{T_{j,(t,i)}}$, which is the sign of the i'th last price move that has been observed, standing at time t. Also recall the concept of "individual trading time" $B_{j,t}$, then $D_{B_{j,t}}$ will be a vector of the signs of the last k price changes in the j–th asset, most of which will be 0. We refer to the i'th element of this vector as $D_{B_{j,(t,i)}}$, which is the direction of the i'th last trade in the j–th asset that has been observed, standing at time t. A simple example of this is $D_{B_{j,(t,1)}}$ which is the direction of the last trade in asset j. In the empirical modelling of direction we find these types of variables very significant statistically when working with multivariate time series. By going from general to specific we find that direction has more memory than we have modelled, see Table 5. From the estimated parameters we see that the process $D_{j,t}$ is strongly mean–reverting (in activity time) reflecting the observed directions. An implication of this fitted model is that the dynamics generating the directions seems symmetrical for Ford and GM since all the $Z_{B_{2,t}}$ variables are positive. So what we initially guessed, namely that Ford and GM moves together has turned out to be true.

Variable	Coef.	StR. Err.	Variable	Coef.	Str. Err.
$Z_{B_{1,(t,1)}}$	-9.429	(0.505)	$Z_{B_{2,(t,1)}}$	0.327	(0.114)
$Z_{B_{1,(t,2)}}$	-5.931	(0.299)	$Z_{B_{2,(t,2)}}$	0.329	(0.119)
$Z_{B_{1,(t,3)}}$	-3.989	(0.229)	$Z_{B_{2,(t,3)}}$	0.525	(0.121)
$Z_{B_{1,(t,4)}}$	-2.728	(0.186)	$Z_{B_{2,(t,4)}}$	0.718	(0.120)
$Z_{B_{1,(t,5)}}$	-1.952	(0.164)	$Z_{B_{2,(t,5)}}$	0.645	(0.121)
$Z_{B_{1,(t,6)}}$	-1.528	(0.155)	$Z_{B_{2,(t,6)}}$	0.604	(0.118)
$Z_{B_{1,(t,7)}}$	-1.094	(0.146)	$Z_{B_{2,(t,7)}}$	0.562	(0.117)
$Z_{B_{1,(t,8)}}$	-0.960	(0.140)	$Z_{B_{2,(t,8)}}$	0.425	(0.117)
$Z_{B_{1,(t,9)}}$	-0.580	(0.132)	$Z_{B_{2,(t,9)}}$	0.478	(0.117)
$Z_{B_{1,(t,10)}}$	-0.426	(0.115)	$Z_{B_{2,(t,10)}}$	0.305	(0.108)
$D_{T_{1,(t,1)}}$	-0.519	(0.062)			
$D_{T_{1,(t,2)}}$	0.162	(0.051)			
$D_{T_{1,(t,9)}}$	-0.134	(0.038)			
Q	$T(T+2)\sum_{j=1}^{Q}\frac{1}{T-j}r_j^2$		Log-like $=-1784.5$		
20	20.38	31.4104			
100	97.16	124.342			
1500	1,508.72	1,591.21			

Table 5: *Estimation results for the direction of active Ford trades using an autologistic model. Variable is the explanatory variable. The figures in brackets are the standard errors on the regressors computed using the autologistic model.* r_j *denotes the series correlation coefficient at lag* j *for the the standardised residuals* u_t*. The figures in brackets are corresponding 95 percentage points on the* χ^2_Q *distribution to give a rough benchmark. AIC=3521.*

4.4.2 AREX model

In the previous section we saw that the direction of Ford had much more memory than we have modelled. Hence, we will have to introduce a more sophisticated and more parsimonious way of modelling the memory. One way of doing this is via an AREX model (AutoRegressive in the EXplanatories). This will give us

$$\theta^D_{j,t} = x'_{j,t}\beta_j + \sum_{h=1}^{K} g_{B_{h,t}},$$

where

$$g_{B_{h,t}} = \sum_{i=1}^{p_h} \gamma_{h,i} g_{B_{h,(t,i+1)}} + \sum_{k=1}^{q_h} \delta_{h,k} Z_{B_{h,(t,k)}},$$

with $x_{j,t}$ being potential combinations and subsets of $F^{D,S,A}_{t-1}$. In order to asses the fit of the model we use the residuals given by

$$u^D_{j,t} = 1_{\{D_{j,t}>0\}} - p\left(\theta^D_{j,t}\right).$$

The reason for choosing this type of residual, which is also white noise, over the standardized one we have used in the previous section is that $p\left(\theta^D_{j,t}\right)$ gets very close to one and zero, hence the standard deviation given by $\sqrt{p\left(\theta^D_{j,t}\right)\left\{1-p\left(\theta^D_{j,t}\right)\right\}}$ is very close to zero. This distorts the correlation structure and when we tried to use the standardized residuals in our Box–Pierce statistics we got a test statistic of about 805 at 1,500 lags (see Table 5).

Variable	Coef.	StR. Err.	Variable	Coef.	Str. Err.
g_1			$D_{T_{1,(t,1)}}$	-0.400	(0.064)
$\rho_{1,1}$	0.453231		$D_{T_{1,(t,2)}}$	0.238	(0.059)
$\rho_{1,2}$	0.396639		Constant	-0.026	(0.044)
$D_{B_{1,(t,1)}}$	-9.382	(0.128)			
$D_{B_{1,(t,2)}}$	-3.458	(1.121)			
$D_{B_{1,(t,3)}}$	1.256	(0.247)			
$D_{B_{1,(t,4)}}$	0.466	(0.145)			
g_2					
$\rho_{2,1}$	0.955177				
$\rho_{2,2}$	-0.391946				
$D_{B_{2,(t,1)}}$	0.191	(0.026)			
$D_{B_{2,(t,2)}}$	0.306	(0.025)			
Q	$T(T+2)\sum_{j=1}^{Q}\frac{1}{T-j}r_j^2$		Log-like $=-1774.7$		
20	31.03	(31.41)	$\mathrm{E}\left(u^D\right)$	-2×10^{-9}	
100	109.11	(124.34)	$\sqrt{\mathrm{Var}\left(u^D\right)}$	0.255	
1500	$1,548.34$	$(1,591.21)$			

Table 6: *Estimation results for the direction of active Ford trades using an autologistic model. Variable is the explanatory variable. The figures in brackets are the standard errors on the regressors computed using the autologistic model.* r_j *denotes the series correlation coefficient at lag* j *for the residuals* u_t^D*. The figures in brackets are corresponding 95 percentage points on the* χ_Q^2 *distribution to give a rough benchmark. AIC=3522.*

4.5 The size of price changes

As noted earlier, of the 8,256 Ford trades where the price moves is different from 0 only 2 are bigger than 1 tick. On the basis of this it seems unnecessary to try and model the sizes. This could of course be done using that same methods as in Rydberg and Shephard (1998a). There a negative binomial model where the parameters are a function of a GLARMA model is suggested. Under different market conditions the size distribution could potentially be very important. That there are so few price changes bigger than 1 tick is in some sense artificial and due to the before mentioned reason that the specialist at the NYSE is imposed to make the price move in a smooth fashion. If the market was purely electronic it would probably react must faster to news, and hence we would see many more large price changes.

Also if we used the model for data aggregated over for instance 5 minute intervals we would need the S part of the decomposition, since this would blur the effect of the market maker.

5 Extensions

5.1 Quotes

One type of information that would be potentially useful would be to know if particular trades where buys or sells. A way of obtaining this information would be to use the quotes data, which are also available from the TAQ database[6] to categorize the data in: buys, sells and undeter-

[6]The NYSE does not release information about the limit order book. At the Paris Bourse the five best bid and offer prices are avilable.

Variable	Coef.	StR. Err.	Variable	Coef.	Str. Err.
	g_2		$D_{T_{2,(t,1)}}$	-0.262	(0.035)
$\rho_{2,1}$	0.73405		$D_{T_{2,(t,2)}}$	0.207	(0.034)
$\rho_{2,2}$	-0.37817		Constant	-0.076	(0.030)
$D_{B_{2,(t,1)}}$	-3.943	(0.055)			
$D_{B_{2,(t,2)}}$	2.080	(0.335)			
$D_{B_{2,(t,3)}}$	-0.370	(0.114)			
	g_1				
$\rho_{1,1}$	0.951885				
$\rho_{1,2}$	-0.307395				
$D_{B_{1,(t,1)}}$	0.209	(0.018)			
$D_{B_{1,(t,2)}}$	0.105	(0.025)			
Q	$T(T+2)\sum_{j=1}^{Q}\frac{1}{T-j}r_j^2$		Log-like $=-3480.7$		
20	29.17	(31.41)	$\mathrm{E}\left(u^D\right)$	1×10^{-9}	
100	136.88	(124.34)	$\sqrt{\mathrm{Var}\left(u^D\right)}$	0.378	
1500	1,404.53	(1,591.21)			

Table 7: *Estimation results for the direction of active GM trades using an autologistic model. Variable is the explanatory variable. The figures in brackets are the standard errors on the regressors computed using the autologistic model.* r_j *denotes the series correlation coefficient at lag* j *for the residuals* u_t^D*. The figures in brackets are corresponding 95 percentage points on the* χ_Q^2 *distribution to give a rough benchmark.*

mined. One way of generating such a variable is suggested in Hausman, Lo, and MacKinlay (1992) where they, for some shares, find that three lags significant at a 5% level.

The quotes and the depth (i.e. how many shares the one who has posted the quote is willing to buy and sell at the quotes) of the quotes could also be indicative of the liquidity of the market. It would probably be more helpful also to have information about the limit and market orders (they are for instance available in the TORQ database complied by Hasbrouck, see Hasbrouck (1992)) as they are the indicators of the liquidity of the market. However such information is not usually publicly available.

5.2 Durations and volume

More obvious types of explanatory variables to include would be information about the volume of previous trades and the durations between them. In a previous study, see Rydberg and Shephard (1998a), we did find interesting dependencies between the activity variables A and lagged values of log(durations) and the direction D and lagged values of log(volume). However we also found that the numerical size of the effects from the variables contained in $\mathcal{F}^{A,D,S}$ did not change by including this information. So since our main objective was to study how the information from the trading in other stocks impacted the trading sequence we have left this open.

5.3 Continuous price changes

In this paper we have assumed that S_t lives on a lattice structure, due to the fact that at NYSE trading is restricted to fixed tick sizes. This, however, is not a crucial assumption for our model. The model extends easily to a situation where there is no underlying fixed tick size by letting

S_t follow a continuous distribution. Even though there is no longer discrete price changes our modelling framework is still helpful in situations where price changes occur at irregular spaces time points.

5.4 Zeros and non-zeros

An important part of our model is the A structure, the reason being that for the highest frequencies of data, the trade–by–trade data, it is approximately the absolute returns series. Since almost 80% of the returns are zero one could consider modelling only the changes, which would be equivalent to conditioning on A, i.e. $D_t S_t | A_t, \mathcal{F}_{t-1}$. This route is taken in Engle and Russell (1997) where they study the frequency of changes in quoted foreign exchange rates.

If we applied our model to aggregated data, e.g. 10 minute returns, or mid pries of quotes the underlying tick size would be smoothed out and the zero price changes would no longer be dominant. In such a situation we would have $A_t = 1$, for all t, i.e. $Z_t = D_t S_t$.

5.5 A market portfolio as an explanatory variable

Here we have only focused on two shares which are in the same line of business. Also we have only used explanatory variables from the filtration generated by these two processes. It would be natural to think that what is actually going on is that both shares are subject to common market fluctuations as well as branch and firm specific information. In order to test such a hypothesis we would need to include some type of market portfolio, this could for instance be the S&P100, the Dow Jones index or the automotive related Dow Jones Automotive index. However, it is not clear if this would actually reveal the desired information since the indices are composite of the shares we are modelling and move according to the activity in the included shares, hence the information should be slightly delayed in the indices.

6 Conclusion

In this paper we studied a period where the minimum tick size was constant. The major reason being that changing the tick size should by economic considerations change the liquidity since a smaller tick size should decrease the spread. We originally did work with a series that also included June 1997, which did see a tick change. We decided to exclude this month since we did see signs of changes in the dynamics. Work continues on this aspect.

In this paper we have proposed a multivariate decomposition of the price movements of trade–by–trade portfolios. The decomposition means we model sequentially which of the assets is being traded, price activity, direction of moves and size of moves. The advantage of the decomposition is that each modelling exercise is straightforward and interpretable. By applying this model to a bivariate time series of the trade in the Ford and the GM share, two large automotive producers, we show that the trading sequence in the two shares holds quite a lot of information about each other.

When combined with a good model for the times between trades in the portfolio this analysis provides a complete model for the evolution of prices in real time.

7 Acknowledgements

Tina Rydberg thanks the Danish National 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 authors are grateful for comments from the participants at the International Workshop on Statistics in Finance, Hong Kong, July, 1999. Thanks to Cilve Granger for helpful conversations on issues regarding the use of market indices.

References

Barndorff-Nielsen, O. E. and G. Schou (1973). On the reparameterization of autoregressive models by partial autocorrelations. *J. Multivariate Analysis 3*, 408–419.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton, New Jersey: Princeton University Press.

Cox, D. R. (1958). The regression analysis of binary sequences (with discussion). *J. Royal Statistical Society B 20*, 215–42.

Cox, D. R. and E. J. Snell (1989). *The Analysis of Binary Data* (2 ed.). London: Chapman & Hall.

Darolles, S., C. Gourieroux, and G. Le Fol (1998). Intra-day transaction price dynamics. Unpublished paper: CERMSEM-Paris I Univ. and CREST.

Engle, R. F. (1996). The econometrics of ultra-high frequency data. Department of Economics, UC. San Diego, forthcomming in *Econometrica*.

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 12*, 187–212.

Engle, R. F. and J. R. Russell (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica 66*, 1127–1162.

Hasbrouck, J. (1992). Using the torq database. Unpublished paper: New York Stock Exchange.

Hasbrouck, J. (1996). The dynamics of discrete bid and ask quotes. Unpublished paper: Stern Business School, New York University.

Hausman, J., A. W. Lo, and A. C. MacKinlay (1992). An ordered probit analysis of transaction stock prices. *Journal of Financial Economics 31*, 319–30.

Jones, M. C. (1987). Randomly choosing parameters for the stationary and invertibility region of autoregressive-moving average models. *Applied Statistics 36*, 134–138.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics 47*, 13–37.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley.

McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models*. London: Chapman & Hall. 2nd Edition.

McFadden, D. L. (1984). Qualitative response models. In Z. Griliches and M. Intriligator (Eds.), *The Handbook of Econometrics, Volume 2*, pp. 1395–1457. North-Holland.

Meddahi, N., E. Renault, and B. Werker (1998). Modelling high frequency data in continuous time. Unpublished paper: CIRANO, CRDE, Montreal University.

New York Stock Exchange (1997, June). *New York Stock Exchange Newsletter*. New York Stock Exchange.

O'Hara, M. (1995). *Market Microstructure Theory*. Oxford: Blackwell Publishers.

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 (1998a). Dynamics of trade–by–trade price movements: decomposition and models. Working paper, Nuffield College, Oxford. Presented at Workshop on Econometrics and Finance, Issac Newton Institute, Cambridge University, October 1998.

Rydberg, T. H. and N. Shephard (1998b). A modelling framework for the prices and times of trades made on the NYSE. To appear in *Nonlinear and nonstationary signal processing* edited by W.J. Fitzgerald, R.L. Smith, A.T. Walden and P. C. Young. Cambridge University Press, 2000.

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.

Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science 9*, 227–293.

Shephard, N. (1994). Autoregressive based generalized linear models. Unpublished paper: Nuffield College, Oxford. Presented at the Econometric Society World Congress, Tokyo, 1995.

Taylor, N., D. van Dijk, P. H. Franses, and A. Lucas (1999). Sets, arbitrage activity, and stock price dynamics. Working paper, University of Warwick.