Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 14 (2007) 564 – 583
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Multivariate autoregressive modeling of time series count data using copulas ☆

Andréas Heinen [a], Erick Rengifo [b,*]

[a] Department of Statistics, Universidad Carlos III de Madrid, 126 Calle de Madrid, 28903 Getafe, Madrid, Spain
[b] Economics Department, Fordham University, 441 East Fordham Road, Bronx, NY 10458-9993, USA

Accepted 4 July 2006
Available online 25 April 2007

Abstract

We introduce the Multivariate Autoregressive Conditional Double Poisson model to deal with discreteness, overdispersion and both auto and cross-correlation, arising with multivariate counts. We model counts with a double Poisson and assume that conditionally on past observations the means follow a Vector Autoregression. We resort to copulas to introduce contemporaneous correlation. We apply it to the study of sector and stock-specific news related to the comovements in the number of trades per unit of time of the most important US department stores traded on the NYSE. We show that the market leaders inside a specific sector are related to their size measured by their market capitalization.
© 2007 Elsevier B.V. All rights reserved.

JEL classification: C32; C35; G10
Keywords: Copula continued extension; Counts; Factor model; Market microstructure

1. Introduction

In empirical studies of market microstructure, the Autoregressive Conditional Duration (ACD) model, introduced by Engle and Russell (1998), has been used widely to test theories with tick-

☆ The authors would like to thank Luc Bauwens, Richard Carson, Robert Engle, Clive Granger, Philippe Lambert, Bruce Lehmann, David Veredas, Ruth Williams, for helpful discussions and suggestions. We thank the editor and three anonymous referees, who helped improve this paper. This research was carried out while both authors were at the Center for Operations Research and Econometrics (CORE), at the Université Catholique de Louvain, and we wish to extend particular thanks to Luc Bauwens and members of CORE for providing a great research environment. The usual disclaimers apply.
* Corresponding author. Tel.: +1 718 817 4061; fax: +1 718 817 3518.
E-mail addresses: aheinen@est-econ.uc3m.es (A. Heinen), rengifomina@fordham.edu (E. Rengifo).

0927-5398/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2006.07.004

by-tick data in a univariate framework. This model is designed specifically to deal with the irregularly-spaced nature of financial time series of durations. However, extensions to more than one series have proven to be very difficult. The difficulty comes from the very nature of the data, that are by definition not aligned in time, i.e. the times at which an event of any type happens are random. Engle and Lunde (2003) suggest a model for the bivariate case, but the specification is not symmetric in the two processes. They analyze jointly the duration between successive trades and the duration between a trade and the next quote arrival. This is done in the framework of competing risks. Spierdijk et al. (2004) model bivariate durations using a univariate model for the duration between the arrival of all events, regardless of their type, and a probit specification which determines the type of event that occurred. These models become intractable when the number of series is greater than two.

In this paper we suggest working with counts instead of durations, especially when there are more than two series. Any duration series can easily be made into a series of counts by choosing an appropriate interval and counting the number of events that occur every period. Moreover, most applications involve relatively rare events, which makes the use of the normal distribution questionable. Thus, modeling this type of series requires one to deal explicitly with the discreteness of the data as well as its time series properties and correlation. Neglecting either of these characteristics would lead to potentially serious misspecification.

Transforming durations into counts involves making an arbitrary decision about the size of the time window. Choosing a small time window will lead to series with potentially many zeros, whereas if the time interval is too large, the loss of information due to time aggregation will be important. It is not clear what would constitute a criterion for the definition of an "optimal" observation window for counts. One possibility could be to seek a statistical answer to this question, but we think that the answer must lie in the type of application at hand and in what constitutes a reasonable time window from the point of view of the application and the question that is being addressed. This choice of time window affects a lot of empirical work in finance, like for instance studies of realized volatility (Aït-Sahalia et al. (2005) suggest using a 5-minute interval) or of commonalities in stock returns and liquidity (see Hasbrouck and Seppi (2001) who use 15-minute returns and order flow). When the number of events that have to be analyzed jointly is large, duration-based approaches become less tractable and count models are more useful. Count models are also preferable when the cross-sectional aspect is of great importance to the modeler. In that case the loss of information from considering counts will be compensated for by the possibility of flexibly modeling interactions between several series. Another case in which counts are preferable is when interest lies in forecasting, as this is difficult with multivariate durations, but straightforward with counts.

We introduce a new multivariate model for time series count data. The Multivariate Autoregressive Conditional Double Poisson model (MDACP) makes it possible to deal with issues of discreteness, over and underdispersion (variance greater or smaller than the mean) and both cross and serial correlation. This paper constitutes a multivariate extension to the univariate time series of counts model developed in Heinen (2003). We take a fully parametric approach where the counts have the double Poisson distribution proposed by Efron (1986) and their mean, conditional on past observations, is autoregressive. In order to introduce contemporaneous correlation we use a multivariate normal copula. This copula is very flexible, since it makes it possible to accommodate both positive and negative correlation, something that is impossible in most existing multivariate count distributions. The models are estimated using maximum likelihood, which makes the usual tests available. In this framework autocorrelation can be tested with a straightforward likelihood ratio test, whose simplicity is in sharp contrast with test

procedures in the latent variable time series count model of Zeger (1988). The model differs in the way the conditional mean is modeled. This model is similar to the model of Davis et al. (2001), but it differs in the use that is made of the exponential link function. In our model the time series part is linear with an identity link function, which makes statements about stationarity very simple, and the explanatory variables are multiplicative with an exponential link function. Moreover, the model proposed by Davis et al. (2001) does not consider any contemporaneous correlation that we capture here using a copula. We apply a two-stage estimation procedure developed in Patton (2006), which consists in estimating first the marginal models and then the copula, taking the parameters of the marginal models as given. This considerably eases estimation of the model. In order to capture the dynamic interactions between the series we model the conditional mean as a VARMA-type structure, focusing our attention to the (1,1) case, motivated largely by considerations of parsimony.

It is well documented in market microstructure that the trading process conveys information. According to Admati and Pfleiderer (1988) and Easley and O'Hara (1992) frequent trading implies that news is arriving to the market. Thus a higher number of trades in a given time interval is a signal for the arrival of news. Information in trading activity can be either stock-specific or sector-wide. How much sector-specific information a stock's trading activity contains has important implications from the point of view of identifying sectorial leaders. Spierdijk et al. (2004) study this question using a duration-based approach for pairs of assets. As a feasible alternative to multivariate duration models, we apply the MDACP to the study of sector and stock-specific news of the most important US department stores traded on the New York Stock Exchange during the year 1999. We model the dynamics of the number of transactions of all stocks simultaneously using an intuitive and parsimonious factor structure, whereby the conditional mean of every series depends on one lag of itself, one lag of the count and one factor of the cross-section of lagged counts. We show that the assets that contain more sector information correspond to assets with larger market capitalizations. This is in contradiction with the findings of Spierdijk et al. (2004), who find that it is the most frequently traded stocks that contain most sector-specific information.

The paper is organized as follows. Section 2 introduces the Multivariate Autoregressive Conditional Double Poisson model, shows how we use copulas in the present context, and describes the conditional mean and the marginal distribution of the model. Section 3 presents the empirical application. Section 4 concludes.

2. The Multivariate Autoregressive Conditional Double Poisson model

In this section we discuss the way in which we use copulas and the continued extension argument to generate a multivariate discrete distribution. Then we present the conditional distribution and the conditional mean of the Multivariate Autoregressive Conditional Double Poisson. Next, we summarize the features of our model and establish its properties.

2.1. A general multivariate model using copulas

In order to generate richer patterns of contemporaneous cross-correlation, we resort to copulas. Copulas provide a very general way of introducing dependence among several series with known marginals. Copula theory goes back to the work of Sklar (1959), who showed that a joint distribution can be decomposed into its K marginal distributions and a copula, that describes the dependence between the variables. This theorem provides an easy way to form valid multivariate

distributions from known marginals that need not be necessarily of the same distribution, i.e. it is possible to use normal, student or any other marginals, combine them with a copula and get a suitable joint distribution, which reflects the kind of dependence present in the series. A more detailed account of copulas can be found in Joe (1997) and in Nelsen (1999).

Let $H(y_1,\ldots,y_K)$ be a continuous K-variate cumulative distribution function with univariate margins $F_i(y_i)$, $i=1,\ldots,K$, where $F_i(y_i)=H(\infty,\ldots,y_i,\ldots,\infty)$. According to Sklar (1959), there exists a function C, called copula, mapping $[0, 1]^K$ into $[0, 1]$, such that:

$$H(y_1,\ldots,y_K) = C(F_1(y_1),\ldots,F_K(y_K)). \tag{1}$$

The joint density function is given by the product of the marginals and the copula density:

$$\frac{\partial H(y_1,\ldots,y_K)}{\partial y_1\ldots\partial y_K} = \prod_{i=1}^{K} f_i(y_i)\frac{\partial C(F_1(y_1),\ldots,F_K(y_K))}{\partial F_1(y_1)\ldots\partial F_K(y_K)}. \tag{2}$$

With this we can define the copula of a multivariate distribution with uniform [0, 1] margins as:

$$C(z_1,\ldots,z_K) = H(F_1^{-1}(z_1),\ldots,F_K^{-1}(z_K)), \tag{3}$$

where $z_i=F_i(y_i)$, for $i=1,\ldots,K$.

As we can see with the use of the copulas we are able to map the univariate marginal distributions of K random variables, each supported in the [0, 1] interval, to their K-variate distribution, supported on $[0, 1]^K$, something that holds, no matter what the dependence among the variables is (including if there is none).

Most of the literature on copulas is concerned with the bivariate case. However, we are trying to specify a general type of multivariate count model, not limited to the bivariate case. Whereas there are many alternative formulations in the bivariate case, the number of possibilities for multi-parameter multivariate copulas is rather limited. We choose to work with the most intuitive one, which is arguably the Gaussian copula, obtained by the inversion method (based on Sklar (1959)). This is a K-dimensional copula such that:

$$C(z_1,\ldots,z_K;\Sigma) = \Phi^K(\Phi^{-1}(z_1),\ldots,\Phi^{-1}(z_K);\Sigma), \tag{4}$$

and its density is given by,

$$c(z_1,\ldots,z_K;\Sigma) = |\Sigma|^{-1/2}\exp\left(\frac{1}{2}(q'(I_K-\Sigma^{-1})q)\right), \tag{5}$$

where Φ^K is the K-dimensional standard normal multivariate distribution function, Φ^{-1} is the inverse of the standard univariate normal distribution function and $q=(q_1,\ldots,q_K)'$ with normal scores $q_i=\Phi^{-1}(z_i)$, $i=1,\ldots,K$. Furthermore, it can be seen that if $Y_1,\ldots,Y_K$ are mutually independent, the matrix Σ is equal to the identity matrix I_K and the copula density is then equal to 1.

In the present paper we use a discrete marginal, the double Poisson, whose support is the set of integers, instead of continuous ones, which are defined for real values. If the marginal distribution functions are all continuous then C is unique. However, when the marginal distributions are discrete, this is no longer the case and the copula is only uniquely identified on $\bigotimes_{i=1}^{K}\mathrm{Range}(F_i)$, a K-dimensional set, which is the Cartesian product of the range of all marginals. Moreover, the problem with discrete distributions is that the Probability Integral Transformation Theorem (PITT) of Fisher (1932) does not apply, and the uniformity assumption does not hold, regardless

of the quality of the specification of the marginal model. The PITT states that if Y is a continuous variable, with cumulative distribution F, then

$$Z = F(Y)$$

is uniformly distributed on [0, 1].

We use the continued extension argument proposed by Denuit and Lambert (2005) to overcome these difficulties and apply copulas with discrete marginals. The main idea of the continued extension is to create a new random variable Y^* by adding to a discrete variable Y a continuous variable U valued in [0, 1], independent of Y, with a strictly increasing cdf, sharing no parameter with the distribution of Y, such as the uniform [0, 1] for instance:

$$Y^* = Y + (U - 1).$$

The continued extension does not alter the concordance between pairs of random variables; intuitively, two random variables Y_1 and Y_2 are concordant, if large values of Y_1 are associated with large values of Y_2. Concordance is an important concept, since it underlies many measures of association between random variables, such as Kendall's tau for instance. It is easy to see that the continued extension does not affect concordance, since $Y_1^* > Y_2^* \Leftrightarrow Y_1 > Y_2$.

Using the continued extension, we state a discrete analog of the PITT. If Y is a discrete random variable with domain χ, in N, such that $f_y = P\,(Y = y)$, $y \in \chi$, continued by U, then

$$Z^* = F^*(Y^*) = F^*(Y + (U - 1)) = F([Y^*]) + f_{[Y^*]+1}U = F(Y - 1) + f_y U$$

is uniformly distributed on [0, 1], and $[Y]$ denotes the integer part of Y. In this paper, we use the continued extension of the probability integral transformation in order to test the correct specification of the marginal models. If the marginal models are well specified, then Z^*, the PIT of the series under the estimated distribution and after the continued extension, is uniformly distributed. We also use Z^* as an argument in the copula, since, provided that the marginal model is well specified, this ensures that the conditions for use of a copula are met.

2.2. The conditional distribution and the conditional mean

In order to extend the Autoregressive Conditional Double Poisson model to a $(K \times 1)$ vector of counts N_t, we build a VARMA-type system for the conditional mean. Even though the Poisson distribution with autoregressive means is the natural starting point for counts, one of its characteristics is that the mean is equal to the variance, property referred to as equidispersion. However, by modeling the mean as an autoregressive process, we generate overdispersion in even the simple Poisson case.

In some cases one might want to break the link between overdispersion and serial correlation. It is quite probable that the overdispersion in the data is not attributable solely to the autocorrelation, but also to other factors, for instance unobserved heterogeneity. It is also imaginable that the amount of overdispersion in the data is less than the overdispersion resulting from the autocorrelation, in which case an underdispersed marginal distribution might be appropriate. In order to account for these possibilities we consider the double Poisson distribution introduced by Efron (1986) in the regression context, which is a natural extension of the Poisson model and allows one to break the equality between conditional mean and variance. The

advantages of using this distribution are that it can be both under and overdispersed, depending on whether ϕ is larger or smaller than 1. We write the model as:

$$N_{i,t}|\mathcal{F}_{t-1} \sim \mathrm{DP}(\mu_{i,t}, \phi_i), \qquad \forall i = 1, \ldots, K. \tag{6}$$

where $\mathcal{F}_{t-1}$ designates the past of all series in the system up to time $t-1$.[1] With the double Poisson, the conditional variance is equal to:

$$V[N_{i,t}|\mathcal{F}_{t-1}] = \sigma_{i,t}^2 = \frac{\mu_{i,t}}{\phi_i} \tag{7}$$

The coefficient ϕ_i of the conditional distribution will be a parameter of interest, as values different from 1 will represent departures from the Poisson distribution. The double Poisson generalizes the Poisson in the sense of allowing more flexible dispersion patterns.

The conditional means μ_t are assumed to follow a VARMA-type process:

$$E[N_t|\mathcal{F}_{t-1}] = \mu_t = \omega + AN_{t-1} + B\mu_{t-1} \tag{8}$$

This can be easily generalized to the higher order case in the obvious manner. In most empirical applications, most especially when the number of series analyzed jointly is large, some additional restrictions might have to be imposed on A and B.

In systems with large K, which could be found, for instance, when analyzing a large group of stocks like the constituents of an index, the full approach would not be feasible, as the number of parameters would get too large. If we assume that A and B are of full rank, the number of parameters that has to be estimated in this model would be $2K^2+K$. In situations where this is not an option, we propose to impose some additional structure on the process of the conditional mean. The most interesting structure is the "reduced rank" and "own effect" model. In this formulation it is assumed that for every series the conditional mean depends on one lag of itself, one lag of the count and r factors of the cross-section of lagged counts. $A=\mathrm{diag}(\alpha_i)+\gamma\delta'$ where γ and δ are (K, r) matrices. This is suited for large systems, where imposing a reduced rank is necessary for practical reasons, but there is reason to believe that every series' own past has explanatory power beyond the factor structure. In particular, the conditional mean can be specified as:

$$\mu_t = \omega + (\mathrm{diag}(\alpha_i) + \gamma\delta')N_{t-1} + \mathrm{diag}(\beta_i)\mu_{t-1}. \tag{9}$$

Moreover, in some cases one might want to assume that the dynamics of all the series under consideration is common, and that one factor explains the dynamics of the whole system. This can be obtained as a special case of our specification under the following set of assumptions: $A=\alpha\gamma\delta'$, $B=\beta I$, $\omega=c\gamma$, where α, β are scalars, $\gamma=(\gamma_1, \gamma_2, \ldots, \gamma_K)'$ and $\delta=(1-\Sigma_{i=2}^{K}\delta_i, \ldots, \delta_K)'$, where we have imposed the normalization that the elements of δ sum to one. This means that if we denote $\Lambda_t=\delta' N_t$, we have an autoregressive process for the factor:

$$\mu_t^0 = c + \alpha\Lambda_{t-1} + \beta\mu_{t-1}^0,$$

and $\mu_t=\gamma\mu_t^0$.

It is easy to show that the MDACP is stationary as long as the roots of the sum of the autoregressive coefficient matrices are within the unit circle, or equivalently, the eigenvalues of

[1] It is shown in Efron (1986) (Fact 2) that the mean of the Double Poisson is μ and that the variance is approximately equal to $\frac{\mu}{\phi}$. Efron (1986) shows that this approximation is highly accurate, and we will use it in our more general specifications.

$(I-A-B)$ lie within the unit circle. In that case, the unconditional mean of the MDACP(p,q) is identical to the one of a VARMA process:

$$E[N_t] = \mu = (I - A - B)^{-1}\omega \tag{10}$$

2.3. The model

Having dealt with the problems due to the discreteness and the time-varying nature of the marginal density in earlier sections, we proceed with the estimation of the model. The joint density of the counts in the double Poisson case with the Gaussian copula is:

$$h(N_{1,t}, \ldots, N_{K,t}, \theta, \Sigma) = \prod_{i=1}^{K} f_{\mathrm{DP}}(N_{i,t}, \mu_{i,t}, \phi_i) \cdot c(q_t; \Sigma),$$

$f_{\mathrm{DP}}(N_{i,t}, \mu_{i,t}, \phi_i)$ denotes the double Poisson density as a function of the observation $N_{i,t}$, the conditional mean $\mu_{i,t}$ and the dispersion parameter ϕ_i. c denotes the copula density of a multivariate normal and $\theta=(\omega, \mathrm{vec}(A), \mathrm{vec}(B))$.

The $q_{i,t}$, gathered in the vector q_t are the normal quantiles of the $z_{i,t}$:

$$q_t = (\Phi^{-1}(z_{1,t}), \ldots, \Phi^{-1}(z_{K,t}))',$$

where the $z_{i,t}$ are the PIT of the continued extension of the count data, under the marginal densities:

$$z_{i,t} = F^*_{i,t}(N^*_{i,t}) = F_{i,t}(N_{i,t} - 1) + f_{i,t}(N_{i,t}) * U_{i,t},$$

where $F^*_{i,t}$, $F_{i,t}$ and $f_{i,t}$ are the conditional cumulative distribution function (cdf) of the continued extension of the data, the conditional cdf and the conditional probability distribution function, respectively.

The $N^*_{i,t}$ are the continued extension of the original count data $N_{i,t}$:

$$N^*_{i,t} = N_{i,t} + (U_{i,t} - 1).$$

Finally the $U_{i,t}$ are uniform random variable, on [0, 1].

Taking logs, one gets:

$$\log(h_t) = \sum_{i=1}^{K} \log(f_{\mathrm{DP}}(N_{i,t}, \mu_{i,t}, \phi_i)) + \log(c(q_t; \Sigma))$$

We consider a two-stage estimator as in Patton (2006). In a first step (developed in Section 2.2), we assume that conditionally on the past, the different series are uncorrelated. This means that there is no contemporaneous correlation and that all the dependence between the series is assumed to be captured by the conditional mean. In a second step we take the parameter estimates of the marginal models as given in order to estimate the parameters of the copula. Such a procedure has been used in Patton (2006), who shows that the procedure delivers consistent estimators, even though their efficiency is reduced relative to a joint estimation of marginal and copula parameters. In our case, given the number of parameters we are estimating and the non-linearities in the copula, a joint estimation is not feasible. Moreover, given that we use the multivariate normal copula, the second step of the two-stage procedure does not require any optimization, as the MLE

of the variance–covariance matrix of a multivariate normal with a zero mean, is simply the sample counterpart:

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} q_t q_t^{'}$$

We evaluate models on the basis of their log-likelihood, but also on the basis of their Pearson residuals, which are defined as: $\epsilon_t = \frac{N_t - \mu_t}{\sigma_t}$. If a model is well specified, the Pearson residuals have mean zero, variance one and no significant autocorrelation left. The $z_{i,t}$'s are another tool for checking the specification. If the model is well specified, the $z_{i,t}$'s should be uniformly distributed and serially uncorrelated. We check this for all the models we estimate.

We now establish two properties about the unconditional variance and the autocorrelation of the MDACP with an ARMA(1,1) structure, which we denote the MDACP(1,1).

Proposition 2.1. (Unconditional variance of the MDACP(1,1) model.)

The unconditional variance of the MDACP(1,1) model, when the conditional mean is given by Eq. (8)*, is equal to*:

$$\mathrm{vec}(V[N_t]) = \left(I_{K^2} + \left((I_{K^2} - (A+B)\otimes(A+B)')^{-1}\cdot(A\otimes A')\right)\right)\cdot \mathrm{vec}(\Omega), \tag{11}$$

where $\Omega = E(\mathrm{Var}[N_t|\mathcal{F}_{t-1}])$. *Under small dispersion asymptotics of* Jorgensen (1987), $\Omega \simeq V^{\frac{1}{2}}\Sigma V^{\frac{1}{2}}$, *where* Σ *is the copula covariance and* V *is the variance of the marginal models:* $V = \mathrm{diag}\left(\frac{\mu_i}{\phi_i}\right)$.

This is a multivariate extension of Proposition 3.2 of Heinen (2003). The proof is shown in the Appendix. The variance is equal to the ratio of the mean to the dispersion parameter, the covariances are zero and therefore the variance–covariance matrix is diagonal. It can be seen that the variance–covariance of the counts is the product of a term reflecting the autoregressive part of the model, a term capturing the variance of the marginal models and a copula term responsible for the part of the contemporaneous cross-correlation which does not go through the time-varying mean.

Proposition 2.2. (Autocovariance of the MDACP(1,1) model)

The autocovariance of the MDACP(1,1) model, when the conditional mean is given by Eq. (8)*, is equal to:*

$$\mathrm{vec}(\mathrm{Cov}[N_t, N_{t-s}]) = [I\otimes A^{-1}((A+B)^s - B(A+B))]\cdot \mathrm{vec}(V[N_t] - \Omega) \tag{12}$$

where Ω *and* $V[N_t]$ *are as defined in* Proposition 2.1.

The proof is shown in the Appendix.

3. Sector and stock-specific news

Much of the microstructure literature is based on the existence of asymmetric information and consequently of two types of traders: the uninformed who trade for liquidity reasons and informed traders who possess superior information. This superior information can be macroeconomic, sector or stock-specific information. Through the trading process this information is disseminated

to the public, therefore trading conveys information. According to Admati and Pfleiderer (1988) and Easley and O'Hara (1992) frequent trading implies that news is arriving to the market. Thus a higher number of trades in a given time interval is a signal for the arrival of news. In recent years, the focus of empirical microstructure has shifted from the study of an individual asset to the analysis of the cross-sectional interactions amongst stocks. Hasbrouck and Seppi (2001) document the existence of commonalities in order flow that are responsible for about two thirds of the commonalities in returns, using principal components analysis and canonical correlations on the stocks of the Dow Jones Industrial Average.

The trading activity of one asset does not only convey information about that specific asset, but can also contain information about the whole sector that this asset belongs to. In order to model comovement in trading activity within a sector, Spierdijk et al. (2004) propose a duration model for the trading intensities of pairs of stocks of department stores. Their model consists of a univariate duration model for the pooled trades of two stocks and a probit specification which determines in which stock a transaction took place. They classify stocks according to how much sector-wide information they contain, based on a series of ratios of the sample variance of the conditional intensity of the pooled and univariate ACD models for each pair of stocks.

We analyze the same data as Spierdijk et al. (2004), but the MDACP allows us to take into account the interaction amongst all stocks simultaneously as in Hasbrouck and Seppi (2001), which is helpful for the purpose of identifying leaders from the point of view of sectorial information, while at the same time modeling the dynamics in a very general framework.

3.1. Data description

We work with the five most important US department stores traded on the New York Stock Exchange during the year 1999: May Department Stores (MAY), Federated Department Stores (FD), J.C. Penney Company, Inc (JCP), Dillar]s INC (DDS) and Saks Inc (SKS). We work with the number of trades in 5-minute intervals. The data we use was taken from the Trades and Quotes (TAQ) data set, produced by the New York Stock Exchange (NYSE). This data set contains every trade and quote posted on the NYSE, the American Stock Exchange and the NASDAQ National Market System for all securities listed on NYSE. We first remove any trades that occurred with non-standard correction or G127 codes (both of these are fields in the trades data base on the TAQ CD's), such as trades that were canceled, trades that were recorded out of time sequence, and

Table 1
Descriptive statistics

	DDS	FD	JCP	MAY	SKS
Marketcap	1647	8945	3538	11,226	1612
No. trades	55,399	100,928	108,392	90,881	59,725
Mean	2.93	5.34	5.73	4.81	3.16
Median	2.00	5.00	5.00	4.00	3.00
Standard deviation	2.57	3.56	3.89	3.04	2.84
Dispersion	2.25	2.38	2.64	1.92	2.55
Maximum	37	35	38	22	32
M inimum	0	0	0	0	0
$Q(20)$	11,560	15,504	34,482	8531.7	33,679

The market capitalization is given in millions of dollars. Descriptive statistics for the number of trades. The number of observations is 18,900. $Q(20)$ is the Ljung–Box Q-statistic of order 20 on the series. The dispersion refers to the ratio of the variance to the mean.

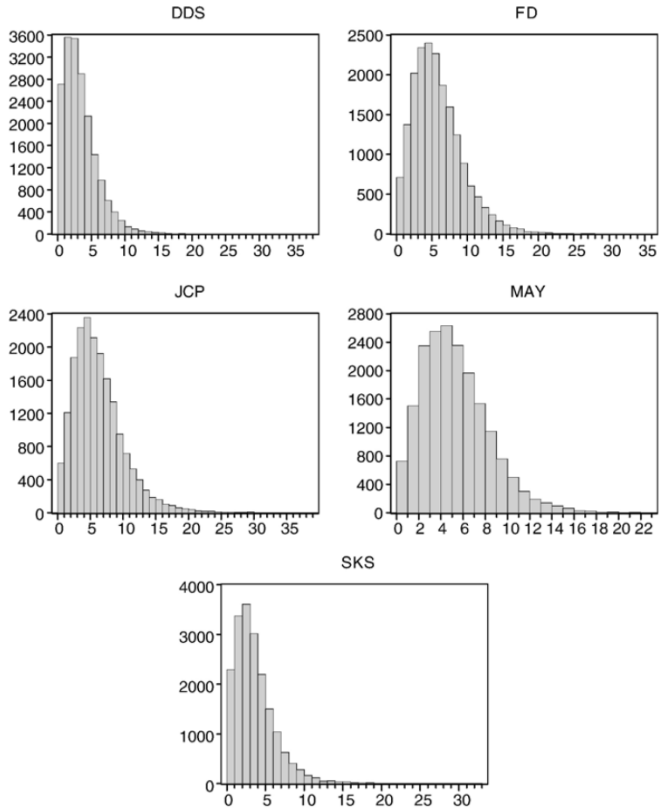

Fig. 1. Histogram of the data.

trades that were called for delivery of the stock at some later date. Any trades that were recorded to have occurred before 9:45 AM or after 4 PM (the official close of trading) were removed. The reason for starting at 9:45 instead of 9:30 AM, the official opening time, is that we wanted to make sure that none of the opening transactions were accidentally included in the sample, or that there would not be artificially low numbers of events at the start of the day, due to the fact that part of the first interval was taking place before the opening transaction. This could have biased our estimates of intradaily seasonality.

The data used was from January 2nd 1999 to December 30th 1999. This means that the sample covers 252 trading days, that represent 18, 900 observations, as there are 75 5-minute intervals every day between 9:45 AM and 4 PM. The descriptive statistics are given in Table 1. We are dealing with

Table 2
Correlation matrix of the trades data

	DDS	FD	JCP	MAY	SKS
DDS	1.00				
FD	0.27	1.00			
JCP	0.24	0.29	1.00		
MAY	0.25	0.30	0.31	1.00	
SKS	0.12	0.10	0.15	0.12	1.00

count data and the means of the series are relatively small, which makes the use of a continuous distribution like the normal problematic. If the means of the series were much larger, it would no longer be necessary to consider the discreteness explicitly. As can be seen, the data exhibits significant overdispersion (the variance is greater than the mean), which could be due alternatively to autocorrelation or to overdispersion in the marginal distribution. The presence of overdispersion is confirmed by looking at the histogram of the data in Fig. 1, which shows that, whereas the probability

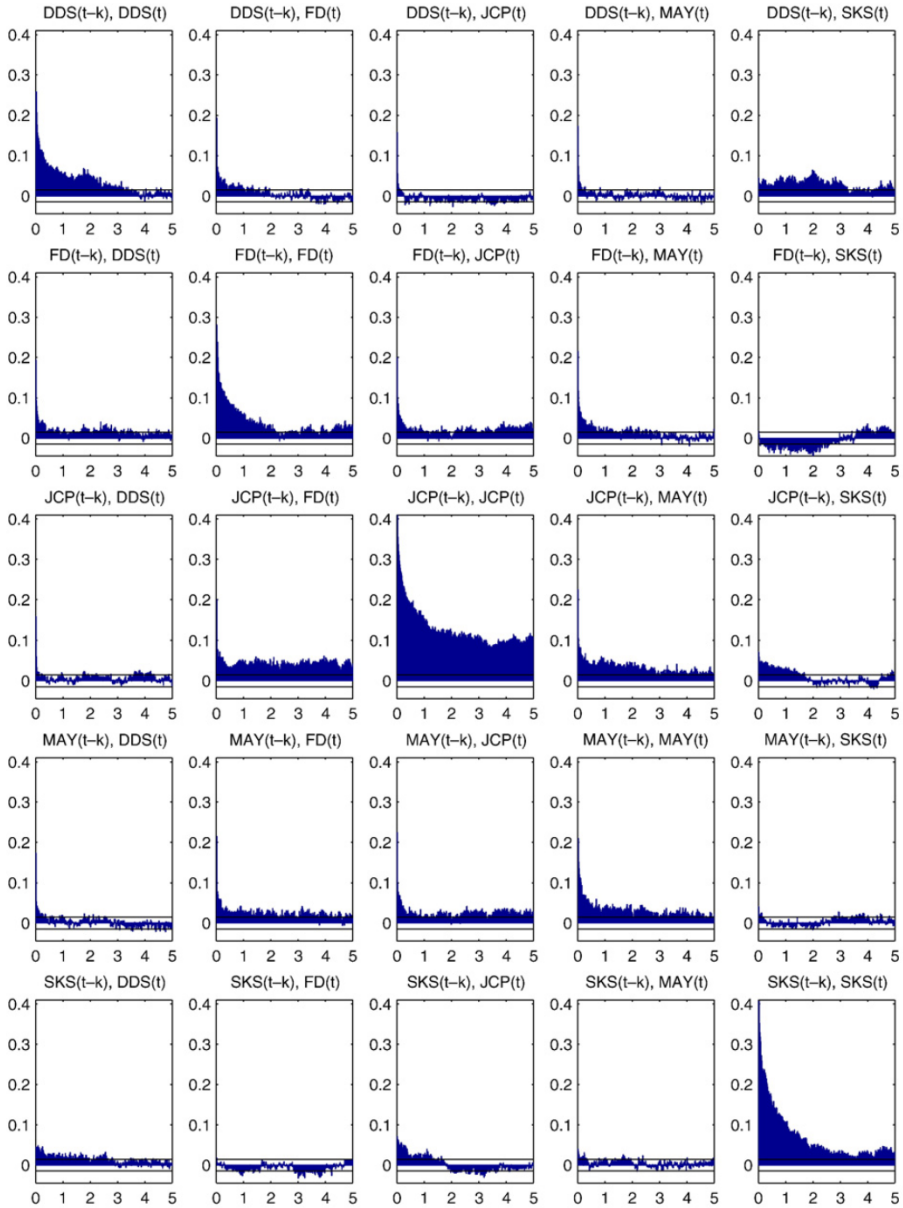

Fig. 2. Auto and cross-correlogram of the deseasonalized data.

mass is fairly concentrated around the mean, there exist large outliers. There is significant autocorrelation in each series, as can be seen from the Ljung–Box Q-statistic shown here at order 20. Table 2 presents the contemporaneous correlation matrix among the five series we analyze, obtained using the Gaussian copula on the data which marginals are assume to be double Poisson distributed. A very striking pattern of seasonality can be observed in the autocorrelogram of the raw data (not shown), which is due to the presence of diurnal seasonality of the U-shape type, which is commonly found in time series based on high frequency data. Fig. 2 shows the auto and cross-correlations of the deseasonalized number of trades, up to 75 5-minute intervals, which corresponds to 1 trading day. There are very important autocorrelations in every stock over the day and probably even a bit longer, that need to be captured by the model. However looking only at contemporaneous correlation does

Table 3
Maximum likelihood estimates of the MDACP models

θ	MDACP factor only model					MDACP with factor and own effect				
	DDS	FD	JCP	MAY	SKS	DDS	FD	JCP	MAY	SKS
ω_i	0.334	0.781	0.099	0.605	0.045	0.136	0.330	0.215	0.275	0.094
	(8.79)	(12.64)	(5.40)	(11.40)	(4.27)	(7.97)	(11.97)	(10.26)	(10.67)	(7.46)
α_{1i}						0.137				
						(25.98)				
α_{2i}							0.151			
							(17.64)			
α_{3i}								0.178		
								(36.43)		
α_{4i}									0.107	
									(11.64)	
α_{5i}										0.161
										(33.72)
γ	0.167	0.309	0.303	0.176	0.128	0.027	0.064	0.010	0.056	0.014
	(17.08)	(20.43)	(30.01)	(16.35)	(23.50)	(6.26)	(5.88)	(2.10)	(5.61)	(5.40)
δ	0.168	0.176	0.265	0.154	0.237	0.175	0.297	0.097	0.371	0.060
		(26.38)	(43.31)	(19.45)	(36.91)		(5.79)	(3.97)	(6.08)	(2.21)
β	0.694	0.663	0.789	0.759	0.838	0.811	0.777	0.814	0.819	0.825
	(34.76)	(37.73)	(105.21)	(47.86)	(110.57)	(109.80)	(97.84)	(155.57)	(103.06)	(142.95)
ϕ	0.504	0.496	0.514	0.571	0.498	0.546	0.542	0.575	0.599	0.584
	(93.78)	(96.23)	(95.83)	(95.57)	(99.46)	(92.79)	(95.58)	(101.25)	(97.26)	(95.26)
Log$L$	−219,072					−214,463				
Eigenvalue	0.99	0.68	0.71	0.83	0.77	0.94	0.97	0.99	0.99	0.93
Var(ϵ_t)	1.00	0.99	1.00	0.97	1.04	0.97	0.98	0.99	0.97	0.98
$E(\epsilon^t)$	−0.001	0.001	0.000	0.001	−0.001	0.000	0.000	0.001	0.000	0.001
$Q(20)$ of ϵ_t	4570	4234	5509	1721	13361	97	79	105	69	124

The table presents the Maximum Likelihood Estimates of the Multivariate Autoregressive Conditional Double Poisson (MDACP) models on counts based on data of the 5 most important retail department stores: DDS, FD, JCP, MAY and SKS at intervals of 5 min for the period January 1999 to the end of December 1999. These models consider the seasonality presented in the data and solved it by the use of 30 min dummies. The t-statistics are presented in parenthesis. We impose the normalization that $\delta = (1 - \sum_{i=2}^{K} \delta_i, \ldots, \delta_K)'$ in order to identify the model. $\epsilon_t = \frac{N_t - \mu_t}{\sigma_t}$ are the Pearson residuals from the model. The equation of the "factor only" model is:

$$\mu_t^0 = c + \alpha \Lambda_{t-1} + \beta \mu_{t-1}^0$$

with $\mu_t = \gamma \mu_t^0$ and the model with a factor and an own effect:

$$\mu_t = \omega + (\mathrm{diag}(\alpha_i) + \gamma\delta')N_{t-1} + \mathrm{diag}(\beta_i)\mu_{t-1}$$

not reveal the full picture, there is also a significant link across time between the stocks. We note that SKS seems to have the least amount of correlation with the other stocks.

3.2. Estimation results

In the present subsection we discuss the estimates of two different specifications of the model, one based on the idea of a common factor and the second based on a mean structure based on a common factor, a series-specific lagged term in the moving average part and a diagonal autoregressive part. In order to fit the dispersion we use the double Poisson distribution and we model seasonality using a series of half hourly dummy variables. The estimation of the factor models requires joint estimation of the dynamic means for all stocks and for this reason we use the same seasonality coefficients for all stocks in order to reduce the computational burden. The seasonality pattern for all stocks is very similar (a graph is available upon request), which means that the assumption that the diurnality is the same for all series is a reasonable one. The results for both models are shown in Table 3. Note that the estimates of the MDACP can be viewed as quasi-maximum likelihood estimates (QMLE), which deliver consistent parameter estimates, even in the case of a misspecified distribution, since their first order condition is the same as in the Poisson case (see Cameron and Trivedi (1998), pages 115–116). However, we believe that a count distribution with a dispersion parameter should be flexible enough to deliver a well specified model, and therefore we view our results as maximum likelihood estimates. We provide some evidence to this effect further down in this section. The eigenvalues of $A+B$ are smaller than 1, which means that the model is stationary. A likelihood ratio test shows that the seasonality variables (the estimates are not shown) are jointly significant. The coefficients on the seasonality shown in Fig. 3 exhibit the well-documented U-shape, which means that there is more activity at the beginning and end of the trading day and less at lunch time. The dispersion parameter ϕ of the double Poisson is also very significantly different from 1, which corresponds to the Poisson case. This means that the Poisson distribution is strongly rejected and that we now have a much better model for the conditional distribution. Furthermore, if the model is well specified, the Pearson residuals will have variance one and neither significant auto nor cross-correlation left. In Table 3

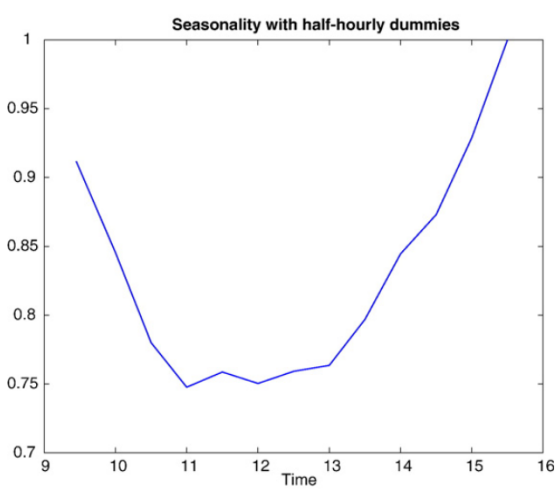

Fig. 3. Coefficients of the seasonality dummies in the “factor and own effect” model.

we can appreciate that the variance of the Pearson residuals is close to one and that the Ljung–Box of the residuals has significantly decreased relative to the original data. Comparing the "factor only" with the "factor and own effect" model shows that the latter captures the dynamics of the system correctly, whereas the former fails.

Visual inspection of the Q–Q plots of the Z statistic of the "factor and own effect" model in Fig. 4 suggests that the distribution is well specified, since the Q–Q plots nearly coincide with the 45-degree line. Besides the Q–Q plots of the sequence of probability integral transforms (the $\{z_{i,t}\}$'s) we use the specification test proposed by Diebold et al. (1998). Fig. 5 shows the histograms of the $\{z_{i,t}\}$'s, estimated using the continued extension proposed by Denuit and Lambert (2005). The histograms reveal that we have some problems in the tails, which are notoriously difficult to model, particularly for discrete data. Formal Kolmogorov–Smirnov tests reject the uniformity assumption of the probability integral transforms, which is not surprising given the sample size of 18,900 observations. This is commonly seen in high frequency models, see for instance Bauwens et al. (2004). This might also be due to the presence of a few outliers in the right tail of the distribution. These episodes of frantic trading with more than 25 trades in 5 minutes, compared to means in the range of 3 to 6 are difficult to capture and more flexible distributions are needed. We prefer to leave these extreme observations in though, as they represent real market events. Similar results hold for the "factor only" model.

The autocorrelations of the Z statistic of the "factor and own effect" model, shown in Fig. 6 are essentially not significant, which indicates that the dynamics of the series is well accounted for. For space reasons we do not present the cross-correlations of the Z statistics which are also insignificant. The auto and cross-correlations of the Pearson residuals of the series (not shown) confirm that there is no more seasonal pattern left and the correlations are well below significance. This however is not the case for the "factor only" model, for which there still remains autocorrelation in the residuals.

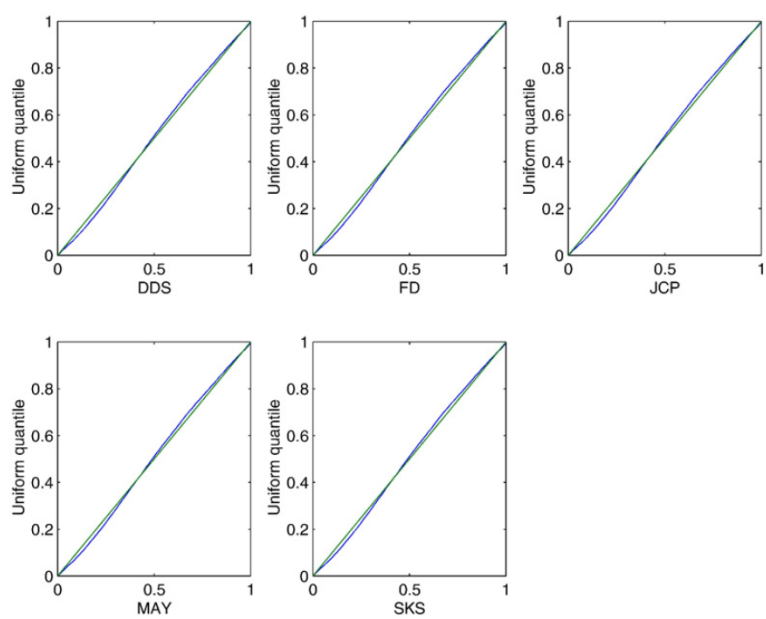

Fig. 4. Quantile plots of the Z statistics of the MDACP model with factor and own effect.

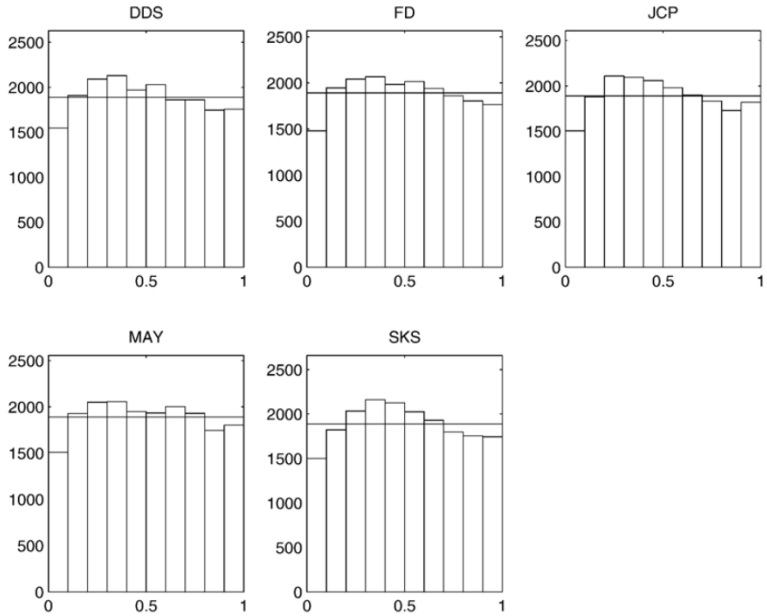

Fig. 5. Histogram of the Z statistics of the MDACP model with factor and own effect.

In order to model the contemporaneous correlations we estimate a multivariate normal copula. As this model is somewhat involved in terms of the number of parameters, we use the two-step procedure of Patton (2006). For the "factor and own effect" model, Table 4 shows the copula

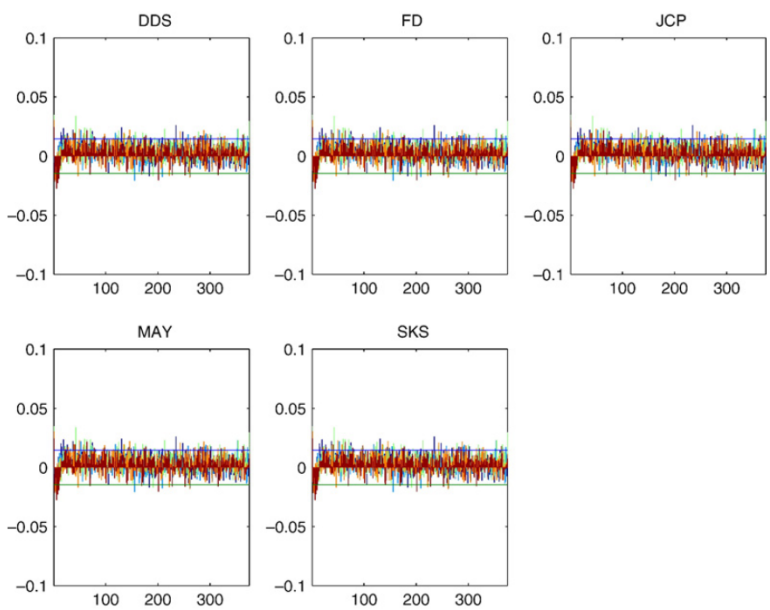

Fig. 6. Autocorrelation of the Z statistics of the MDACP model with factor and own effect.

Table 4
Correlation matrix of the Q estimated by the MDACP model

	Copula–MDACP				
	DDS	FD	JCP	MAY	SKS
DDS	1.00				
FD	0.16	1.00			
JCP	0.17	0.18	1.00		
MAY	0.15	0.17	0.20	1.00	
SKS	0.02	0.02	0.04	0.03	1.00

The table presents the correlation matrix of Q, based on the probability integral transformation, Z, of the continued extension of the count data under the marginal densities estimated using the "factor and own effect" model estimated with the two-step procedure.

correlation matrix Σ of Proposition 2.1, which is responsible for the part of the contemporaneous and lagged cross-correlation which does not go through the time-varying mean.

We apply the goodness of fit test of Chen et al. (2004) to the Gaussian copula and we get a test statistic of -0.14, which follows a standard normal distribution under the null of normality. We cannot reject the null hypothesis of normality. In order to check the robustness of our result we have also estimated a multivariate t copula, which is the most well known of very few alternatives in a multivariate framework (recently Demarta and McNeil (2005) have introduced the grouped t and the skewed t copulas) and the estimated degrees of freedom are 31.69, which implies that our estimated t copula can hardly be distinguished from the normal one.

Our first results are based on the "factor only" model (left panel of Table 3), in which we assume that the dynamics of all the series under consideration is common, and that one factor explains the dynamics of the whole system. To see the influence on the factor of each of the assets involved we just need to take a look at the vector of factor weights (the δ's). According to this the ranking of sectorial influence is JCP, SKS, DDS, FD and MAY. These results are exactly the same as the results of Spierdijk et al. (2004). This ranking has a Spearman's rank correlation of 0.20 with trading activity of the assets and is highly negatively correlated with market capitalization with a Spearman's correlation of -0.90 (see Table 1 for trading activity and market capitalization of the stocks). However, if instead we rely on the intuitive idea that every stock's past trading activity plays a special role for that asset, in addition to an effect through a common factor, we find quite a different result.

The results in the right side of Table 3, obtained with a series-specific lagged term, a common factor in the moving average part and a diagonal autoregressive part, imply a quite different ranking: MAY, FD, DDS, JCP and SKS. This has a Spearman's rank correlation of 0.50 with trading activity, meaning that to some extent more frequently traded assets contain more sectorial information. However, the striking feature is the high correlation with market capitalization of the stocks (MAY, FD, JCP, DDS and SKS), with a Spearman's rank correlation of 0.90. Based on these results, we can conclude that within a sector there exist two kinds of information that matter for traders: stock-specific information, related to the series-specific autocorrelation coefficients (the $\alpha_{i,i}$'s) and sector specific news, captured by the common factor (the δ's and γ's).

From our analysis we can identify stocks whose weight in the common factor is most important and which therefore influence other assets in the sector. A dealer trading assets in that sector will watch these stocks with particular attention to changes in their trading activity, since they can convey information that is also relevant for other stocks in the sector. Of course, to obtain

Table 5
Factor weights for 5, 10 and 15-minute models

	MDACP factor only model					MDACP with factor and own effect				
	DDS	FD	JCP	MAY	SKS	DDS	FD	JCP	MAY	SKS
5 min	0.168	0.176	0.265	0.154	0.237	0.175	0.297	0.097	0.371	0.060
10 min	0.189	0.191	0.235	0.167	0.219	0.214	0.294	0.124	0.295	0.072
15 min	0.180	0.175	0.250	0.158	0.237	0.229	0.295	0.093	0.309	0.074

more general results, one would need to incorporate all stocks of that sector and not only the biggest ones as we do. One caveat applies in that what we have labeled "sector-specific" also contains market-wide news. In order to disentangle these two effects, one would need to work with more stocks in different sectors. This is left for future research.

The comparison of our results with the ones of duration-based models suggests that taking into consideration all the assets simultaneously does make a difference. We are able to capture cross-sectional interactions with an intuitive factor structure, commonly used in finance since the CAPM and also used more recently in the context of liquidity and order flow by Hasbrouck and Seppi (2001). As we are more interested in getting the cross-sectional aspect right and the strict timing of events seems less important to us, we believe that the advantage of our multivariate specification compensates for the potential loss of information due to the aggregation. Moreover, we have estimated our models for different time intervals (10 and 15 min) and we obtain the same results. Results for the factor weights for 5, 10 and 15 min models are shown in Table 5 (full results are available upon request). This robustness over time aggregation and the accordance of our results with economic intuition increases our confidence in the findings.

4. Conclusion

In this paper we introduce new models for the analysis of multivariate time series of count data with many possible specifications. These models have proved to be very flexible and easy to estimate. We discuss how to adapt copulas to the case of time series of counts and show that the Multivariate Autoregressive Conditional Double Poisson model (MDACP) can accommodate many features of multivariate count data, such as discreteness, over and underdispersion (variance greater and smaller than the mean) and both auto and cross-correlation. Hypothesis testing in this context is straightforward because all the usual likelihood-based tests can be applied. Another important advantage of this model is that it can accommodate both positive and negative correlation among variables, which most multivariate count models cannot do, and which could be important in some applications. Finally, the model can very easily be generalized to distributions other than the double Poisson. The model could be used for instance to model jointly counts, continuous positive variables and a real-valued continuous variable in a time-varying framework with a copula to capture the dependence amongst the variables.

As a feasible alternative to multivariate duration models, the model is applied to the study of sector and stock-specific news related to the comovements in the number of trades per unit of time of the most important US department stocks traded on the New York Stock Exchange. We show that the informational leaders inside a specific sector are related to their size measured by their market capitalization rather than to their trading activity.

We advocate the use of the Multivariate Autoregressive Conditional Double Poisson model for the study of multivariate point processes in finance, when the number of variables considered simultaneously exceeds two and looking at durations becomes too difficult. Plans for further

research include evaluating the forecasting ability of these models, both in terms of point and density forecasts and we left more empirical applications for further work with more detailed tick-by-tick data sets.

Appendix A

Proof of Proposition 2.1. Upon substitution of the mean equation in the autoregressive intensity, one obtains:

$$\mu_t - \mu = A(N_{t-1} - \mu) + B(\mu_{t-1} - \mu) \tag{13}$$

$$\mu_t - \mu = A(N_{t-1} - \mu_{t-1}) + (A + B)(\mu_{t-1} - \mu) \tag{14}$$

Squaring and taking expectations gives:

$$V[\mu_t] = AE[(N_{t-1} - \mu_{t-1})(N_{t-1} - \mu_{t-1})']A + (A + B)V[\mu_{t-1}](A + B)' \tag{15}$$

Using the law of iterated expectations and denoting $\Omega = V[N_t|\mathcal{F}_{t-1}]$, one gets:

$$V[\mu_t] = A\Omega A + (A + B)V[\mu_{t-1}](A + B)' \tag{16}$$

Vectorializing and collecting terms, one gets:

$$\mathrm{vec}(V[\mu_t]) = (I_{K^2} - (A + B) \otimes (A + B)')^{-1} \cdot (A \otimes A') \cdot \mathrm{vec}(\Omega) \tag{17}$$

Now, applying the following property on conditional variance

$$V[y] = E_x[V_{y|x}(y|x)] + V_x[E_{y|x}(y|x)] \tag{18}$$

to the counts and vectorializing, one obtains:

$$\mathrm{vec}(V[N_t]) = \mathrm{vec}(\Omega) + \mathrm{vec}(V[\mu_t]) \tag{19}$$

Again using the law of iterated expectations, substituting the conditional variance σ_t for its expression, then making use of the previous result, and after finally collecting terms, one gets the announced result.

$$\mathrm{vec}(V[N_t]) = \left(I_{K^2} + \left((I_{K^2} - (A + B) \otimes (A + B)')^{-1} \cdot (A \otimes A')\right)\right) \cdot \mathrm{vec}(\Omega) \tag{20}$$

Based on Song (2000), and on tail area approximations Jorgensen (1997), we can approximate the Pearson residual as follows:

$$F(N_{i,t}, \mu_{i,t}, \phi) \simeq \Phi\left(\frac{N_{i,t} - \mu_{i,t}}{\sqrt{\frac{\mu_{i,t}}{\phi_i}}}\right), \tag{21}$$

Equivalently, we have:

$$q_{i,t} \equiv \Phi^{-1}(F(N_{i,t}, \mu_{i,t}, \phi)) \simeq \frac{N_{i,t} - \mu_{i,t}}{\sqrt{\frac{\mu_{i,t}}{\phi_i}}} \equiv \epsilon_{i,t}, \tag{22}$$

Therefore, we can approximate the variance–covariance of the Pearson residuals with the copula covariance:

$$\Sigma = \mathrm{Cov}(q_t) \simeq \mathrm{Cov}(\epsilon_{i,t}) \tag{23}$$

Now the average conditional variance–covariance matrix Ω can be obtained simply from Σ as:

$$\Omega \simeq V^{\frac{1}{2}} \Sigma V^{\frac{1}{2}} \tag{24}$$

□

Proof of Proposition 2.2. As a consequence of the martingale property, deviations between the time t value of the dependent variable and the conditional mean are independent from the information set at time t. Therefore:

$$E[(N_t - \mu_t)(\mu_{t-s} - \mu)'] = 0 \qquad \forall s \geq 0 \tag{25}$$

By distributing $N_t - \mu_t$, one gets:

$$\mathrm{Cov}[N_t, \mu_{t-s}] = \mathrm{Cov}[\mu_t, \mu_{t-s}] \qquad \forall s \geq 0 \tag{26}$$

By the same "non-anticipation" condition as used above, it must be true that:

$$E[(N_t - \mu_t)(N_{t-s} - \mu)'] = 0 \qquad \forall s \geq 0 \tag{27}$$

Again, distributing $N_t - \mu_t$, one gets:

$$\mathrm{Cov}[N_t, N_{t-s}] = \mathrm{Cov}[\mu_t, N_{t-s}] \qquad \forall s \geq 0 \tag{28}$$

Now,

$$\begin{aligned} \mathrm{Cov}[\mu_t, \mu_{t-s}] &= A\mathrm{Cov}[N_t, \mu_{t-s+1}] + B\mathrm{Cov}[\mu_t, \mu_{t-s}] \\ &= (A+B)\mathrm{Cov}[\mu_t, \mu_{t-s}] \\ &= (A+B)^s V[\mu_t] \end{aligned} \tag{29}$$

The first line was obtained by replacing μ_t by its expression, the second line by making use of Eq. (26), the last line follows from iterating line two.

$$\mathrm{Cov}[\mu_t, \mu_{t-s+1}] = A\mathrm{Cov}[\mu_t, N_{t-s}] + B\mathrm{Cov}[\mu_t, \mu_{t-s}] \tag{30}$$

Rearranging and making use of Eq. (28), one gets:

$$\begin{aligned} A\mathrm{Cov}[N_t, N_{t-s}] &= \mathrm{Cov}[\mu_t, \mu_{t-s+1}] - B\mathrm{Cov}[\mu_t, \mu_{t-s}] \\ &= ((A+B)^s - B(A+B))V[\mu_t] \end{aligned} \tag{31}$$

Under the condition that A is invertible, which is not an innocuous assumption, as it excludes the pure factor model, we get after vectorializing:

$$\mathrm{vec}(\mathrm{Cov}[N_t, N_{t-s}]) = [I \otimes (A^{-1}(A+B)^s - A^{-1}B(A+B))]\mathrm{vec}(V[\mu_t]) \tag{32}$$

After substituting in Eq. (20), we get:

$$\begin{aligned} \mathrm{vec}(\mathrm{Cov}[N_t, N_{t-s}]) = {} & [I \otimes A^{-1}((A+B)^s - B(A+B))] \\ & \times \left(I_{K^2} + \left((I_{K^2} - (A+B) \otimes (A+B)')^{-1} \cdot (A \otimes A')\right)\right) \cdot \mathrm{vec}(\Omega) \end{aligned} \tag{33}$$

□

References

Admati, Anat, Pfleiderer, Paul, 1988. A theory of intraday patterns: volume and price variability. Review of Financial Studies 1, 3–40.

Aït-Sahalia, Y., Mykland, P., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. Review of Financial Studies 18, 351–416.

Bauwens, L., Giot, P., Grammig, J., Veredas, D., 2004. A comparison of financial duration models via density forecasts. International Journal of Forecasting 20, 589–609.

Cameron, A. Colin, Trivedi, Pravin K., 1998. Regression Analysis of Count Data. Cambridge University Press, Cambridge.

Chen, X., Fan, Y., Patton, A., 2004. Simple tests for models of dependence between multiple financial time series, with applications to U.S. equity returns and exchange rates. Working Paper. London School of Economics.

Davis, R., Rydberg, T., Shephard, N., Streett, S., 2001. The cbin model for counts: testing for common features in the speed of trading, quote changes, limit and market order arrivals. Working Paper. Nuffield College.

Demarta, S., McNeil, A.J., 2005. The *t* copula and related copulas. International Statistical Review 73.

Denuit, M., Lambert, P., 2005. Constraints on concordance measures in bivariate discrete data. Journal of Multivariate Analysis 93, 40–57.

Diebold, Francis X., Gunther, Todd A., Tay, Anthony S., 1998. Evaluating density forecasts with applications to financial risk management. International Economic Review 39, 863–883.

Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. Journal of Finance 47, 577–605.

Efron, Bradley, 1986. Double exponential families and their use in generalized linear regression. Journal of the American Statistical Association 81, 709–721.

Engle, R., Lunde, A., 2003. Trades and quotes: a bivariate point process. Journal of Financial Econometrics 1, 159–188.

Engle, Robert F., Russell, Jeffrey, 1998. Autoregressive Conditional Duration: a new model for irregularly spaced transaction data. Econometrica 66, 1127–1162.

Fisher, R.A., 1932. Statistical Methods for Research Workers.

Hasbrouck, Joel, Seppi, Duane J., 2001. Common factors in prices, order flows and liquidity. Journal of Financial Economics 59, 383–411.

Heinen, A., 2003. Modeling time series count data: an autoregressive conditional Poisson model. CORE Discussion Paper 2003/62.

Joe, Harry, 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

Jorgensen, Bent, 1987. Exponential dispersion model. Journal of the Royal Statistical Society 49, 127–162.

Jorgensen, Bent, 1997. The Theory of Dispersion Models. Chapman and Hall, London.

Nelsen, Roger B., 1999. An Introduction to Copulas. Springer, New York.

Patton, A.J., 2006. Estimation of multivariate models for time series of possibly different lengths. Journal of Applied Econometrics 21, 147–173.

Sklar, A., 1959. Fonctions de répartitions à *n* dimensions et leurs marges. Public Institute of Statistics of the Univeristy of Paris 8, 229–231.

Song, Peter Xue-Kun, 2000. Multivariate dispersion models generated from Gaussian copulas. Scandinavian Journal of Statistics 27, 305–320.

Spierdijk, L., Nijman, T., Van Soest, A., 2004. Modeling comovements in trading intensities to distinguish sector and specific news. Working Paper. Econometric and Finance Group, CentER, Tilburg University.

Zeger, Scott L., 1988. A regression model for time series of counts. Biometrika 75, 621–629.