Bivariate Time Series Modelling of Financial Count Data

Shahiduzzaman Quoreshi
Department of Economics & USBE, Umeå University
SE-901 87 Umeå, Sweden
Tel: +46-90-786 77 21, Fax: +46-90-77 23 02
E-mail: shahid.quoreshi@econ.umu.se

Abstract
A bivariate integer-valued moving average (BINMA) model is proposed. The BINMA model allows for both positive and negative correlation between the counts. This model can be seen as an inverse of the conditional duration model in the sense that short durations in a time interval correspond to a large count and vice versa. The conditional mean, variance and covariance of the BINMA model are given. Model extensions to include explanatory variables are suggested. Using the BINMA model for AstraZeneca and Ericsson B it is found that there is positive correlation between the stock transactions series. Empirically, we find support for the use of long-lag bivariate moving average models for the two series.

Key Words: Count data, Intra-day, High frequency, Time series, Estimation, Long memory, Finance.

JEL Classification: C13, C22, C25, C51, G12, G14.

Umeå Economic Studies 655, 2005

*The financial support from the Umeå School of Business and Economics and the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. Special thanks to Kurt Brännäs and Jörgen Hellström for their comments on the earlier versions of this paper.
1. INTRODUCTION

This paper focuses on the modelling of bivariate time series of count data that are generated from stock transactions. Each transaction refers to a trade between a buyer and a seller in a volume of stocks for a given price. Besides volume and price, a transaction is impounded with other information like, e.g., spread, i.e. the difference between bid and ask price. The used data are aggregates over five minutes intervals and computed from real time, time series data. The presented count data model can be seen as an inverse of the conditional duration model of Engle and Russell (1998) in the sense that short durations in a time interval correspond to a large count and vice versa. One obvious advantage of the current model over the conditional duration model is that there is no synchronization problem between the time series.\footnote{For a bivariate duration model the durations for transactions typically start at different times and as a consequence measuring the covariance between the series becomes intricate.} Hence, the spread of shocks and news is more easily studied in the present framework. Moreover, the bivariate count data models can easily be extended to multivariate models without much complication.

The introduced bivariate time series count data model allows for negative correlation between the counts and the integer-value property of counts is taken into account. The model that we introduce emerges from the integer-valued autoregressive moving average (INARMA) model, which is related to the conventional ARMA class of Box and Jenkins (1970). An important difference, however, between these two model classes is that INARMA comprises parameters that are interpreted as probabilities so that the values of the parameters are restricted to unit intervals. McKenzie (1986) and Al-Osh and Alzaid (1987) independently introduced the INARMA model for pure time series, while Brännäs (1995) extended an INAR model to account for explanatory variables. The application of INARMA models in economics is rather new. Some empirical applications of INARMA are due to Blundell, Griffith and Windmeijer (2002), who studied the number of patents in firms, Rudholm (2001), who studied competition in the generic pharmaceuticals market, and Brännäs, Hellström and Nordström (2002), who estimated a nonlinear INMA(1) model for tourism demand. Some introductory treatises of count data are available in Winkelmann and Zimmermann (1995) and in specialized monographs, such as Cameron and Trivedi (1998).

A large number of studies have considered the modelling of bivariate or multivariate count data assuming an underlying Poisson distribution (e.g., Gourieroux et al., 1984). Heinen and Rengifo (2003) introduce multivariate time series count data models based on the Poisson and the double Poisson distribution. Their models allow for negative correlation among the variables but depart from conventional count data regression models. Earlier models for intra-day transactions data or related financial variables have departed from traditional count data regression models or from extended versions (e.g., Brännäs and Brännäs, 2004; Rydberg and Shephard, 1999). Until now, the only related study based on
the INARMA class appears to be Brännäs and Quoreshi (2004). In a univariate setting they found that the estimated INARMA and INAR models did not satisfy the restrictions on parameters while INMA did so. Here, we develop a bivariate integer-valued moving average (BINMA) model that does satisfy the natural conditions of a count data model and that accounts for the long memory aspects of the data. The model is specified in terms of first and second order moments conditioned on historical observations.

The paper is organized as follows. The BINMA model is introduced in Section 2. The conditional and unconditional properties of the BINMA models are obtained. The extension of the BINMA model is also contained in this section. The estimation procedures, CLS, FGLS and GMM, for unknown parameters are discussed in Section 3. A detailed description of data is given in Section 4. The empirical results for the stock series are presented in Section 5 and the concluding comments are included in the final section.

2. MODEL

This section introduces the BINMA model for the number of transactions in equidistant time intervals. The unconditional and conditional first and second order moments of the BINMA model are obtained. Later an extension to time dependent parameters and possible inclusions of explanatory variables are discussed. Finally, multivariate extensions are briefly indicated.

2.1 The BINMA Model

Assume that there are two stock series, $y_{1t}$ and $y_{2t}$, for the number of transactions in intraday data. Assume further that the dependence between $y_{1t}$ and $y_{2t}$ emerges from factor(s) like for example macro-economic news, rumors, etc. The macro-economic news may have impact on both stocks. The correlation between these stock variables can be modelled by extending the INMA($q$) model into a bivariate one that we call BINMA($q_1,q_2$). The model in its simplest form can be defined as follows

$$y_{1t} = u_{1t} + \alpha_{11} \circ u_{1t-1} + \ldots + \alpha_{1p} \circ u_{1t-q_1} \tag{1a}$$

$$y_{2t} = u_{2t} + \alpha_{21} \circ u_{2t-1} + \ldots + \alpha_{2q} \circ u_{2t-q_2}. \tag{1b}$$

The binomial thinning operator is used to account for the integer-valued property of count data. This operator can be written

$$\alpha \circ u = \sum_{i=1}^{u} z_i \tag{2}$$

with $\{z_i\}_{i=1}^{u}$ an iid sequence of 0-1 random variables, such that $\Pr(z_i = 1) = \alpha = 1 - \Pr(z_i = 0)$. Conditionally on the integer-valued $u$, $\alpha \circ u$ is binomially distributed with $E(\alpha \circ u \mid u) = \alpha u$ and $V(\alpha \circ u \mid u) = \alpha(1-\alpha)u$. Unconditionally it holds that $E(\alpha \circ u) = \alpha \lambda$ and $V(\alpha \circ u) = \alpha^2 \sigma^2 + \alpha(1-\alpha)\lambda$, where $E(u) = \lambda$ and $V(u) = \sigma^2$. Obviously, $\alpha \circ u \in [0, u]$. 
Assuming independence between and within the thinning operations and \( \{u_{jt}\}, j = 1, 2 \), an iid sequence with mean \( \lambda_j \) and variance \( \sigma^2_j = v_j \lambda_j \), the unconditional first and second order moments can be given as follows:

\[
E(y_{jt}) = \lambda_j (1 + \sum_{i=1}^{q_j} \alpha_{ji}) \\
V(y_{jt}) = \lambda_j [(v_j + \sum_{i=1}^{q_j} \alpha_{ji}) + (v_j - 1) \sum_{i=1}^{q_j} \alpha^2_{ji}] \\
\gamma_{jk} = \sigma^2_j (\alpha_{jk} + \sum_{i=1}^{q_j} \alpha_{ji} \alpha_{jk+i}), \quad k \geq 1
\]

where \( \gamma_{jk} \) denotes the autocovariance function at lag \( k \) and \( v_j > 0 \). It is clear from (3) that the mean, variance and autocovariance take only positive values since \( \lambda_j, \sigma^2_j \) and \( \alpha_{ji} \) are all positive. Note also that the variance may be larger than the mean (overdispersion), smaller than the mean (underdispersion), or equal to the mean (equidispersion) depending on whether \( v_j > 1, v_j \in (0, 1) \) or \( v_j = 1 \) respectively.

Macro-economic news, rumors, etc. can enhance the intensity of trading in both stocks or lead the intensities in opposite directions. This implies that investors in different stocks may react after the news in similar or different ways. For example, investors may increase their investments in one stock leading to a possible increase in price, while reducing their investments in another stock creating a possible price decrease. Thus, even though the prices of the two stocks move in different directions, the intensities of trading in both stocks may increase. For a fixed time interval, e.g., \([t - 1, t)\), the macro-economic news are assumed to be captured by \( u_{jt} \) for stock \( j \).

Retaining the previous assumptions and allowing for dependence between \( u_{1t} \) and \( u_{2t} \), the unconditional covariance function for \( \text{BINMA}(q_1, q_2) \) can be given as follows:

\[
\gamma_k = \begin{cases} 
\Lambda \left( \alpha_{1k} + \sum_{i=1}^{q_k} \alpha_{1k+i} \alpha_{2i} \right), & 0 \leq k \leq \min(q_1, q_2) \\
0, & k > \min(q_1, q_2) > 0
\end{cases}
\]

\[
\gamma_{-k} = \begin{cases} 
\Lambda \left( \alpha_{2k} + \sum_{i=1}^{q_k} \alpha_{2k+i} \alpha_{1i} \right), & 0 \leq k \leq \min(q_1, q_2) \\
0, & k > \min(q_1, q_2) > 0
\end{cases}
\]

where \( \gamma_k = \text{Cov}(y_{1t}, y_{2t-k}) \), \( \gamma_{-k} = \text{Cov}(y_{1t-k}, y_{2t}) \), and \( \Lambda = \varphi - \lambda_1 \lambda_2 \) where \( \varphi = E(u_{1t} u_{2t}) \).

Note that there is no cross-lag dependence among \( u_{jt} \), i.e., \( \text{Cov}(u_{1t-k}, u_{2t-k-i}) = 0 \) for \( i \neq 0 \), and the covariances \( \text{Cov}(u_{1t}, u_{2j}) \) are assumed constant over time. Note also that the sign of the covariance function in (4a−b) depends on the relative sizes of \( \varphi \) and \( \lambda_1 \lambda_2 \).

Brännäs and Hall (2001) give the conditional mean and variance for a univariate model.
The conditional mean and variance for the BINMA($q_1, q_2$) are in an analogous way

\[
E(y_{jt}|Y_{1t-1}) = E_{jt|t-1} = \lambda_j + \sum_{i=1}^{q_j} \alpha_{ji} u_{jt-i} \tag{5a}
\]

\[
V(y_{jt}|Y_{1t-1}) = V_{jt|t-1} = \nu_j \lambda_j + \sum_{i=1}^{q_j} \alpha_{ji} (1 - \alpha_{ji}) u_{jt-i}. \tag{5b}
\]

Note that the mean and variance are conditioned on only the previous observations, $Y_{1t-1} = y_{1t-1}, y_{1t-2}, \ldots$. Since the conditional variance varies with $u_{jt-i}, i \geq 1$, there is a conditional heteroskedasticity property of moving average type that Brännäs and Hall called MACH($q$). The effect of $u_{jt-i}$ on the mean is greater than on the variance. Note also that like the unconditional variance the conditional variance could be overdispersed, underdispersed or equidispersed depending on whether $\nu_j > 1 + \sum_{i=1}^{q_j} \alpha_{ji}^2 / \lambda_j$, $\nu_j \in (0, 1 + \sum_{i=1}^{q_j} \alpha_{ji}^2 / \lambda_j)$ or $\nu_j = 1 + \sum_{i=1}^{q_j} \alpha_{ji}^2 / \lambda_j$, respectively. The conditional covariance function for BINMA($q_1, q_2$) can be written

\[
\gamma_{kt|t-1} = E[(y_{1t} - E_{1t|t-1})(y_{2t} - E_{2t|t-1}) \mid Y_{t-1}] = \begin{cases} \Lambda, & k = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{6}
\]

Hence the conditional covariance does not vary with $u_{jt}$.

In order to get some additional insight into the correlation structure of the model consider the equidispersion case, i.e. $\nu_j = 1$, or assume that $\{u_{jt}\}$ are iid Poisson sequences with $\sigma_j^2 = \lambda_j$. The unconditional second order moments change to $V(y_{jt}) = E(y_{jt})$, while the first order unconditional moments (3a) remain unchanged. The covariance function remains unaffected, while the correlation function changes due to the changes in the variances. Hence, we will use the standard deviations $V^{1/2}(y_{jt})$ for the case of equidispersion, to build the correlation function. For $k = 0$ in (4), the correlation between $y_{1t}$ and $y_{2t}$ is

\[
\rho_\varphi = \frac{(1 + \sum_{i=1}^{\min(q_1,q_2)} \alpha_{i1} \alpha_{i2}) \Lambda}{V^{1/2}(y_{1t}) V^{1/2}(y_{2t})}. \tag{7}
\]

Note that this correlation can take a positive or a negative sign depending on the size of $\varphi$ relatively $\lambda_1 \lambda_2$. In an univariate setting Brännäs and Quoreshi (2004) showed that the autocorrelation function take only values in the interval $[0,1]$. By applying the Cauchy-Schwarz inequality we can show for $\rho_\varphi$ that

\[
| \varphi - \lambda_1 \lambda_2 | \leq \lambda_1^{1/2} \lambda_2^{1/2}. \]

The correlation $| \rho_\varphi | = 1$ if and only if $| \varphi - \lambda_1 \lambda_2 | = \lambda_1^{1/2} \lambda_2^{1/2}$ and if in addition $\alpha_{ji} = 1$ for all $i \geq 1$. For an invertible INMA model the latter condition is not valid since then $\alpha_{ji} < 1$. Hence, it is clear that $| \rho_\varphi | < 1$. This also holds for the over- and underdispersion cases. As $\nu_j$ deviates from 1, $| \rho_\varphi |$ decreases. It can be shown that this result gets support from the coherence function for the BINMA($q_1, q_2$).\footnote{A detailed description of coherence function can be found, e.g., in Brockwell and Davis (1991, ch. 4, 10 and 11).}
2.2 Extensions of the BINMA Model

Multivariate INMA models follow directly from BINMA\((q_1, q_2)\) and can be written

\[ y_t = u_t + A_1 \circ u_{t-1} + \ldots + A_q \circ u_{t-q} \]  

(8)

where the \( A_i, i = 1, \ldots, q \), are diagonal matrices and \( u_t \sim (\lambda, \Sigma) \). Covariation between elements can also arise, even if \( \Sigma \) is a diagonal matrix but then, at least, one of the \( A_i \) must have one or several off-diagonal elements. This corresponds to letting \( y_{jt} \) depend on lags of \( u_{it} \), \( i \neq j \).

There are several other ways of extending the model. One way is to allow for time-varying \( \lambda \) as a function of explanatory variables. This can be specified as

\[ \lambda_{jt} = \exp(x_{jt} \theta_j) \geq 0 \]  

(9)

where \( M \) variables related to, e.g., previous prices for correlated stocks may be included in \( x_{jt} \). In order to obtain a more flexible conditional variance specification in (5b) we may let \( \sigma^2_{jt} \) be time dependent \( \sigma_{jt}^2 \). Allowing \( \sigma^2_{jt} \) to depend on past values of \( \sigma^2_{jt}, u_{jt} \) and explanatory variables, using an exponential form, we may specify (cf. Nelson, 1991)

\[ \sigma^2_{jt} = \exp[\phi_0 + \phi_1 \ln \sigma^2_{jt-1} + \ldots + \phi_P \ln \sigma^2_{t-P}] \]  

+ \( \omega_1(u_{jt-1} - \lambda_j)^2 + \ldots + \omega_Q(u_{jt} - \lambda_j)^2 + x_{jt} \xi_j \].

(10)

3. ESTIMATION

Here, we discuss methods for the estimation of the unknown parameters of the conditional mean and variance functions. Since we do not assume a full density function the maximum likelihood estimator is not considered. As a result we only discuss the conditional least squares (CLS), the feasible generalized least square (FGLS) and the generalized method of moments (GMM) estimator.

The three estimators, CLS, FGLS and GMM, have the following residual in common

\[ e_{jt} = y_{jt} - E_{jt|t-1}, \quad j = 1, 2. \]  

(11)

To create empirical moment conditions, instruments are to be chosen depending on the particular model specification. These moment conditions correspond to the normal equations of the CLS estimator that focuses on the unknown parameters of the conditional mean function. Alternatively and equivalently the properties \( E(e_{jt}) = 0 \) and \( E(e_{jt}e_{jt-i}) = 0 \), \( i \geq 1 \) could be used. The CLS estimator minimizes the criterion function \( S_{CLS} = \sum_{t=r}^T e^2_{jt} \), where \( r = q_j + 1 \) and \( T \) is the length of the time series, with respect to the unknown parameter vector \( \psi'_j = (\lambda_j, \alpha'_j) \) or \( \psi'_j = (\theta'_j, \alpha'_j) \) when a time-varying \( \lambda_{jt} \) is employed. To calculate the sequence \( e_{jt} \) we write the prediction error on the form

\[ e_{jt} = u_{jt} - \lambda_{jt} + \sum_{i=1}^{q_j} (\alpha_{ji} \circ u_{jt-i} - \alpha_{ji} u_{jt-i}). \]  

(12)
Instead of using, say, some randomization device to evaluate the sum we advocate using its expected value zero and so employ \( e_{jt} = u_{jt} - \lambda_{jt} \).

The parameters estimated with CLS are considered a first step of the FGLS estimator. For the next step, the conditional variance and the covariance prediction errors

\[
e_{jt} = (y_{jt} - E_{jt|t-1})^2 - V_{jt|t-1} \quad (13)
\]

\[
e_{jt} = e_{11t}e_{21t} - \gamma_{\psi|t-1} \quad (14)
\]

are used. The same prediction errors are also incorporated as moment conditions for the GMM estimator.

For FGLS, \( S_2 = \sum_{l=s_j} e_{jt}^2 \) and \( S_3 = \sum_{l=s_j} e_{jt}^2 \), where \( s_j = \max(q_j, P_j, Q_j) + 1 \), are minimized with respect to the respective parameters of the function \( \sigma^2_{jt} \) and with the CLS estimates \( \hat{\psi}_j \) and \( \hat{u}_{jt} \) kept fixed. Simpler and obvious moment estimators for time invariant \( \sigma_j^2 = v_j \lambda_j \) and \( \varphi \) can be written on the following forms

\[
\hat{\sigma}_j^2 = (T - r)^{-1} \sum_{l=s_j}^T \left[ e_{jit}^2 - \sum_{i=1}^{q_j} \hat{\alpha}_{ji}(1 - \hat{\alpha}_{ji})\hat{u}_{jt-i} \right]
\]

\[
\hat{\varphi} = (T - s)^{-1} \sum_{l=s+1}^T \left[ e_{11t}e_{21t} + \hat{\lambda}_1\hat{\lambda}_2 \right]
\]

where \( s = \max(s_1, s_2) - 1 \). For the third and final step of the FGLS estimator, the criterion

\[
S_{FGLS} = \sum_{l=s_j}^T (\hat{V}_{21|t-1}e_{11t}^2 + \hat{V}_{1|t-1}e_{21t}^2 - 2\hat{\gamma}_{0|t-1}e_{11t}e_{21t})/\hat{D}_t \quad (15)
\]

is minimized with respect to \( \psi_j \). In (15) \( \hat{V}_{21|t-1}, \hat{\gamma}_{0|t-1} \) and \( \hat{D}_t = \hat{V}_{1|t-1} - \hat{V}_{21|t-1} - \hat{\gamma}_{0|t-1}^2 \) are taken as given. This gives the FGLS estimates of the parameter vector \( \psi = (\psi_1', \psi_2')' \) of the bivariate conditional mean function. The covariance matrix estimator is

\[
\text{Cov}(\hat{\psi}_{FGLS}) = \left( \sum_{l=s+1}^T \frac{\partial e_t}{\partial \psi} \hat{\Omega}^{-1} \frac{\partial e_t}{\partial \psi} \right)^{-1}
\]

where \( e_t = (e_{11t}, e_{21t})' \) and \( \hat{\Omega} \) is the covariance matrix for the residual series from FGLS estimation.

One advantage of using the GMM estimator is that all parameters can be estimated jointly (Hansen, 1982). In contrast to the FGLS estimator, where weight is given with respect to individual observation, the weighting in the GMM estimator is constructed with respect to the moment conditions. We may anticipate a better performance of the FGLS estimator than that of the GMM estimator (Brünnäs, 1995). The GMM criterion function

\[
q = \mathbf{m}' \mathbf{W}^{-1} \mathbf{m}
\]

has the vector of moment conditions \( \mathbf{m} \) depending on the specification and is minimized with respect to \( \eta' = (\psi', \omega') \) where \( \omega' = (\sigma_j^2, \phi_0, \ldots, +\phi_{P_j}, \omega_1, \ldots, \omega_Q, \xi_j) \). The \( \mathbf{m} \)
comprises three vectors, i.e. \( \mathbf{m} = (\mathbf{m}_1', \mathbf{m}_2', \mathbf{m}_3') \). The moment conditions corresponding to the conditional mean, i.e. the first order condition of the CLS estimator

\[
\mathbf{m}_1 = (T - n)^{-1} \sum_{t=n+1}^{T} \mathbf{m}_{1t}
\]

where \( \mathbf{m}_{1t} = e_{j,t} \partial e_{j,t} / \partial \psi_k \) with \( n = s_1 + s_2 - 2 \). The moment conditions for the conditional variance and the covariance corresponding to (13) and (14) are collected into \( \mathbf{m}_2 \) and \( \mathbf{m}_3 \), respectively. The following is used as a consistent estimator of the weight matrix \( \hat{\Gamma} \)

\[
\hat{\Gamma} = (T - n)^{-1} \sum_{t=n+1}^{T} \mathbf{m}_t \mathbf{m}'_t.
\]

If we set \( \hat{\mathbf{W}} = \mathbf{I} \), the covariance matrix of the parameter estimator is estimated by

\[
\text{Cov}(\eta) = (T - n)^{-1} (\hat{\mathbf{G}}' \hat{\mathbf{G}})^{-1} \hat{\mathbf{F}} \hat{\mathbf{G}}' (\hat{\mathbf{G}}' \hat{\mathbf{G}})^{-1},
\]

where \( \hat{\mathbf{G}} = \partial \hat{\mathbf{m}} / \partial \eta \). The covariance matrix of the parameter estimator becomes

\[
\text{Cov}(\eta) = (T - n)^{-1} \hat{\mathbf{G}}^{-1} \hat{\mathbf{F}} (\hat{\mathbf{G}}')^{-1}
\]

when the numbers of parameters and moment conditions are equal.

4. DATA AND DESCRIPTIVES

The tick-by-tick data for Ericsson B and AstraZeneca have been downloaded from the Ecovision system and later filtered by the author. These stocks are frequently traded stocks and have the highest turnover at the Stockholmsbörsen. The stock series are collected for the period November 5-December 12, 2002. Due to a technical problem in downloading data there are no data for November 12 in the time series and the first captured minute of December 5 is 1037. To analyze the data we have deleted all trading before 0935 (trading opens at 0930) and after 1714 (order book closes at 1720) since our intention is to capture ordinary transactions. The transactions in the first few minutes are subject to a different trading mechanism while there is practically no trading after 1714. The data are aggregated into five minute intervals of time. There are altogether 2392 observations for both Ericsson B and AstraZeneca. The series are exhibited in Figure 1. There are frequent zero frequencies in the AstraZeneca series and hence the application of count data modelling is called for.

The autocorrelation functions for both the Ericsson B and the AstraZeneca series and their first differences are displayed in Figures 2 and 3, respectively. For the first differences of both series, the autocorrelations are nearly zero after lag one. The partial autocorrelation functions for both series die out gradually for both the level and the difference series. The autocorrelation functions for both level series suggest that a low order AR-part together with a low order MA-part be included in the model. The differenced series suggest that a low order AR-part together with an MA(1) parameter be in the model. The cross-correlation function for Ericsson B and AstraZeneca is presented in Figure 4. The correlation coefficient at lag zero is 0.28. The AstraZeneca series leads the
Figure 1: Time series plots for Ericsson B (mean 58.64, variance 1193.70, maximum 249) and AstraZeneca (mean 6.64, variance 34.38, maximum 64) for the period November 5-December 12, 2002.
Figure 2: Autocorrelation functions from lag one for the Ericsson B series and its first difference.

Figure 3: Autocorrelation functions from lag one for the AstraZeneca series and its first difference.
Figure 4: Cross-correlation function between Ericsson B and AstraZeneca from lag 0 to ±20.

Ericsson B series at lag 1 with a correlation coefficient of 0.23, and the Ericsson B series leads the AstraZeneca series at lag −1 with a similar correlation coefficient. After a few lags the correlation function decays slowly but without any particular pattern. It is worth noting that for higher lags the correlations for the negative lags are generally higher than those for positive lags.

Applying conventional Box-Jenkins methodology, we have estimated INARMA models for both the level and differenced Ericsson B and AstraZeneca series. These results together with the results for pure but higher order INAR model support the findings in Brännäs and Quoreshi (2004). The estimated INARMA model does not satisfy the restrictions on parameters while higher order INAR models are not successful in eliminating serial correlation. The estimated INAR(1) parameters in the INARMA models for both Ericsson B and AstraZeneca are close to 1 indicating the presence of unit roots.

Alternatively, the positive autocorrelations and the slow hyperbolic decay in the autocorrelation functions may suggest a long memory model. There are several ways of studying whether these data series have long memory properties or not. Based on the variance time function

\[ R(k) = k \frac{\sigma_1^2}{\sigma_k^2}, \quad k \geq 1 \]
where \( \sigma^2_k = V(y_t - y_{t-k}) \) and \( \sigma^2_k \sim O(k^{2d-1}) \) for an I(d) process, a test for the presence of I(d) can be conducted (Diebold, 1989). The growth in \( R(k) \) is constant for \( d < 1/2 \). There is a decreasing growth rate for \( 1/2 < d < 1 \) while an increasing growth rate for \( 1 < d < 3/2 \). For \( 0 < d < 1/2 \), the process is called a long memory process in the sense of \( \lim_{n \to \infty} \sum_{-n}^{n} | \rho_k | = \infty \), where \( \rho_k \) is the autocorrelation function at lag \( k \) of \( y_t \) (McLeod and Hipel, 1978). In Figure 5, \( R(k) \) for the AstraZeneca series exhibits roughly a linear function in \( k \), while \( R(k) \) for the Ericsson B series initially has larger increments than for larger \( k \). One possible way of estimating a long memory process is based on the ARFIMA class. The estimated parameters for ARFIMA \((1, d, 1)\) models for both the Ericsson B and the AstraZeneca series contain non-negative parameters. Moreover, the estimated \( d \) for the AstraZeneca series is consistent with a long memory process while the estimated \( d \) for the Ericsson B series indicates a process that has a mean reversion property but is not covariance stationary.

Taken together, there is empirical justification for developing bivariate INMA models with long lags even though these models are less parsimoniously parameterized.
5. EMPIRICAL RESULTS

Both FGLS and GMM methods are employed for estimation. FGLS turns out to be the best in terms of residual properties. The models are estimated under the assumption of conditional heteroskedasticity. AIC and SBIC criteria for both univariate and bivariate models are used to select the lag length of the BINMA model. With FGLS a BINMA(18,16 with additional lags 20 and 22) appears to be the best model while with GMM a BINMA(17,15) is selected. The standardized residuals, estimated with GMM, are serially correlated by the Ljung-Box test statistic. However, the squared standardized residuals from the FGLS estimator for both series do not pass the Ljung-Box test statistic. However, the squared standardized residuals are not of particular interest in this model since we are interested in estimating the mean number of transactions but not in capturing the volatility property which is of particular interest only in price processes.

The Ljung-Box statistic $Q_{n,k}$ for a bivariate model has a $p$-value close to zero at 60 degrees of freedom. This indicates that there is cross-correlation between the residual series.

The estimation results for the FGLS and GMM estimators for the final models are given in Table 1. The estimated parameters are indexed by 1 for Ericsson B, while with 2 for AstraZeneca. Employing FGLS, all the estimated BINMA coefficients are positive and highly significant. Employing GMM, all the estimated BINMA parameters are also positive and all but one are significant at the 5 percent level. The coefficient for AstraZeneca at lag 12 is significant at the 10 percent level, though. For FGLS, the $\hat{\alpha}_{1i}$s decrease, with some exceptions, all the way as lag increases, while $\hat{\alpha}_{2i}$s decrease in the same way until lag 12. After lag 12, the $\hat{\alpha}_{2i}$s fluctuate around 0.06. This downward trend in parameters can be interpreted that the impact of macro-economic and other common news on both stocks are similar in a sense that the intensity of trading increases as the news breaks out and fades away with time. For FGLS and $\hat{\alpha}_{1i}$, a linear regression gives $\hat{\alpha}_{1i} = 0.358 - 0.019i$ ($R^2 = 0.87$). We get a better result in terms of goodness-of-fit when $\hat{\alpha}_{1i}$ is dropped from the regression. In an univariate setting, Brännäs and Quoreshi (2004) find a similar result for the same stock, Ericsson B, but for a different data series.

The FGLS (GMM) estimate for the correlation coefficient $\hat{\rho}_{y_{1t-1}}$ is 0.15 (0.68). The large difference between the estimated correlation coefficient from the two estimators is due to the large difference in variances for the AstraZeneca series. The corresponding correlation between $y_{1t}$ and $y_{2t}$ in the sample is 0.28.

For FGLS (GMM), the $\hat{\alpha}_{1i}$ estimates give a mean lag of 5.18 (5.52) and a median lag of 4 (5), while the $\hat{\alpha}_{2i}$ estimates give a mean lag of 4.93 (3.78) and a median lag of 3 (3). Since the used data are aggregates over five minutes intervals, the mean lags must

\[^3\text{As measures of reaction times to macroeconomic news/rumors in the }\{u_{jt}\}\text{ sequence we use the mean lag }\sum_{i=0}^{w} \alpha_{ji}/w, \text{ where } w = \sum_{i=0}^{w} \alpha_{j0} \text{ and where } \alpha_{j0} = 1. \text{ Alternatively, we use the median lag, which is the smallest } k \text{ such that } \sum_{i=0}^{k} \alpha_{ji}/w \geq 0.5.\]
Table 1: Results for BINMA models for Ericsson B and AstraZeneca estimated by FGLS and GMM.

<table>
<thead>
<tr>
<th>Lag</th>
<th>FGLS</th>
<th>GMM</th>
<th>Lag</th>
<th>FGLS</th>
<th>GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{a}_{11}$</td>
<td>s.e.</td>
<td>$\hat{a}_{11}$</td>
<td>s.e.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.453</td>
<td>0.029</td>
<td>0.576</td>
<td>0.032</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.335</td>
<td>0.025</td>
<td>0.501</td>
<td>0.033</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0.283</td>
<td>0.024</td>
<td>0.446</td>
<td>0.029</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>0.231</td>
<td>0.022</td>
<td>0.382</td>
<td>0.029</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0.234</td>
<td>0.022</td>
<td>0.436</td>
<td>0.028</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>0.266</td>
<td>0.023</td>
<td>0.455</td>
<td>0.033</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>0.192</td>
<td>0.021</td>
<td>0.370</td>
<td>0.027</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>0.100</td>
<td>0.031</td>
<td>0.356</td>
<td>0.028</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>0.289</td>
<td>0.021</td>
<td>0.269</td>
<td>0.024</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>0.188</td>
<td>0.021</td>
<td>0.306</td>
<td>0.024</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>0.118</td>
<td>0.020</td>
<td>0.196</td>
<td>0.022</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>0.123</td>
<td>0.020</td>
<td>0.179</td>
<td>0.020</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>0.103</td>
<td>0.020</td>
<td>0.119</td>
<td>0.021</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>0.092</td>
<td>0.020</td>
<td>0.132</td>
<td>0.022</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>0.084</td>
<td>0.020</td>
<td>0.118</td>
<td>0.025</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>0.098</td>
<td>0.020</td>
<td>0.150</td>
<td>0.024</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>0.070</td>
<td>0.020</td>
<td>0.081</td>
<td>0.022</td>
<td>20</td>
</tr>
<tr>
<td>18</td>
<td>0.040</td>
<td>0.019</td>
<td></td>
<td></td>
<td>22</td>
</tr>
</tbody>
</table>

$\hat{\lambda}_1$ 13.54 0.320 9.95 3.18 $\hat{\lambda}_2$ 2.31 0.032 2.867 0.91
$\hat{\sigma}_1^2$ 550.96 746.53 38.71 $\hat{\sigma}_2^2$ 22.75 10.57 2.81
$\hat{\bar{\varphi}}$ 50.17 98.79 28.56 $\hat{\rho}_{1\text{,}t-1}$ 0.15 0.68
$LB_{1,30}$ 34.67 96.39 $LB_{2,30}$ 18.41 184.71
$LB_{1,30}^2$ 143.90 151.12 $LB_{2,30}^2$ 67.70 36.69
be multiplied by 5 to express them in terms of minute. Hence, for FGLS, the $\hat{\alpha}_{1i}$ estimates give a mean reaction time of 25.91 minutes and a median reaction time of 20 minutes for Ericsson B, while the corresponding mean and median for AstraZeneca are 24.64 and 15 minutes, respectively. Hence, for the measurement of reaction time, the choice of mean or median lag matters, specially for the FGLS estimator. In an earlier study, in a univariate model it is found that the mean reaction time for Ericsson B is 15.8 minutes and the corresponding median is 14 minutes (Brännäs and Quoreshi, 2004). There are several possible explanations for these differences in mean and median reaction time. First, the time gap between the data sets used in these studies is 4 months. Second, the data used in Brännäs and Quoreshi (2004) are collected in the period July 2 – 22, 2002 that is one month after the decision of issuing new shares, while the data used in this study are collected after about two months of the realization of new issuing shares. Hence, we may expect to have more intensity in trading in the former period than the latter.

6. CONCLUDING REMARKS

This study introduces a bivariate integer-valued moving average model (BINMA) and applies the BINMA model to the number of stock transactions in intra-day data. The BINMA model allows for both positive and negative correlations between the count data series. The conditional and unconditional first and second moments are given. The study shows that the correlation between series in the BINMA model is always smaller than 1 in an absolute sense. For the number of transactions in Ericsson B and AstraZeneca, we find promising and less promising features of the model. The conditional mean, variance and covariance have successfully been estimated. The standardized residuals based on FGLS are serially uncorrelated. But the model could not eliminate the serial correlation in the squared standardized residual series that, however, is not of particular interest in this study. Further study is required to eliminate that serial correlation. One way of getting possibly better performance in eliminating serial correlation might be using the extended model, i.e. letting $\lambda_j$ or $\sigma_j^2$ be time-varying. In a univariate setting, by letting $\lambda$ vary with time, the serial correlation reduces drastically (Brännäs and Quoreshi, 2004). Alternatively, by introducing non-diagonal $\mathbf{A}$ matrices as in (8), we could allow for an asymmetric flow of news from say Ericsson B to AstraZeneca but not the other way.
REFERENCE


Econometrica 59, 347-370.