TIME SERIES MODELLING OF HIGH FREQUENCY STOCK TRANSACTION DATA

A.M.M. Shahiduzzaman Quoreshi

Abstract

This thesis comprises four papers concerning modelling of financial count data. Paper [1], [2] and [3] advance the integer-valued moving average model (INMA), a special case of integer-valued autoregressive moving average (INARMA) model class, and apply the models to the number of stock transactions in intra-day data. Paper [4] focuses on modelling the long memory property of time series of count data and on applying the model in a financial setting.

Paper [1] advances the INMA model to model the number of transactions in stocks in intra-day data. The conditional mean and variance properties are discussed and model extensions to include, e.g., explanatory variables are offered. Least squares and generalized method of moment estimators are presented. In a small Monte Carlo study a feasible least squares estimator comes out as the best choice. Empirically we find support for the use of long-lag moving average models in a Swedish stock series. There is evidence of asymmetric effects of news about prices on the number of transactions.

Paper [2] introduces a bivariate integer-valued moving average (BINMA) model 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 study shows that the correlation between series in the BINMA model is always smaller than one in an absolute sense. The conditional mean, variance and covariance 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.

Paper [3] introduces a vector integer-valued moving average (VINMA) model. The VINMA model allows for both positive and negative correlations between the counts. The conditional and unconditional first and second order moments are obtained. The CLS and FGLS estimators are discussed. The model is capable of capturing the covariance between and within intra-day time series of transaction frequency data due to macroeconomic news and news related to a specific stock. Empirically, it is found that the spillover effect from Ericsson B to AstraZeneca is larger than that from AstraZeneca to Ericsson B.

Paper [4] develops models to account for the long memory property in a count data framework and applies the models to high frequency stock transactions data. The unconditional and conditional first and second order moments are given. The CLS and FGLS estimators are discussed. In its empirical application to two stock series for AstraZeneca and Ericsson B, we find that both series have a fractional integration property.

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

To uncle Parvez

Acknowledgement

The following four papers and a summary are included in this thesis:

[I] Brännäs, K. and Quoreshi, A.M.M.S. (2004). Integer-Valued Moving Average Modelling of the Number of Transactions in Stocks. *Umeå Economic Studies* **637** (revised).

[II] Quoreshi, A.M.M.S. (2006). Bivariate Time Series Modelling of Financial Count Data. To appear in *Communications in Statistics: Theory and Methods,* **35, Issue 7**.

[III] Quoreshi, A.M.M.S. (2006). A Vector Integer-Valued Moving Average Model for High Frequency Financial Count Data. *Umeå Economic Studies* **674**.

[IV] Quoreshi, A.M.M.S. (2006). Long Memory, Count Data, Time Series Modelling for Financial Applications. *Umeå Economic Studies* **673**.

1   Introduction

What determines the price of a good is one of the most important questions in economics. If a household or an individual wishes to buy a good at a price and another individual agrees to sell the good at the same price we can say that the price is determined through the mutual agreement between the buyer and the seller. Since many people are interested in buying the same type of products at different prices and several suppliers or sellers are interested in selling the products at different prices the question of how market clearing price is determined arises. The market clearing price refers to the price at which the quantity demanded for a good is the same as the quantity supplied. According to classical economic theory, the market clearing price or equilibrium price is determined through the intersection of demand and supply curves.

The studies of market microstructure depart from the classical economic theory of price determination or Walrasian auctioneer approach, i.e. the auctioneer aggregates demands and supplies of a good to find a market-clearing price. Some early studies on price formation, e.g., Working (1953), not only concern the matching of demand and supply curves in equilibrium but also focus on the underlying trading mechanism. Demsetz (1968) focuses on transactions costs for the determination of prices in the securities market and analyzes the importance of the time dimension of demand and supply in the formation of market prices. The availability of high frequency data specially in stock and currency markets has spurred interest in studying market mechanisms or market microstructures. For stock and currency markets, the market microstructure studies concern, for example, the impact of transactions, bid-ask spreads, volume and time between transactions (duration) on price formation. The studies also concern how news, rumors, etc., are interpreted and used by the actors in trading.

A transaction or a trade takes place when a buyer and a seller agree to exchange a volume of stocks at a given price. A transaction is impounded with information such as volume, price, spread, i.e., the difference between bid and ask prices. The time between transactions and numbers of transactions or trades are related due to the nature of these kinds of data. The more time elapses between successive transactions the fewer trades take place in a fixed time interval. Hence, the trading intensity and the durations can be seen as inversely related. The trading intensity and durations have played a central roll in understanding price processes in the market microstructure research during the last two decades. Diamond and Verrecchia (1987) show that a low trading intensity implies the presence of bad news, while Easley and O'Harra

(1992) shows that a low trading intensity implies no news. Engle (2000) finds that longer durations are associated with lower price volatilities. The stock transactions data are counts for a fixed interval of time. Until now there is no study of pure time series models for count data in this area and this thesis contributes to filling this gap.

A time series of count data is an integer-valued non-negative sequence of count observations observed at equidistant instants of time. There is a growing literature of various aspect of how to model, estimate and use such data. Jacobs and Lewis (1978ab, 1983) develop discrete ARMA (DARMA) models that introduce time dependence through a mixture process. McKenzie (1986) and Al-Osh and Alzaid (1987) introduce independently the integer-valued autoregressive moving average (INARMA) model for pure time series data, while Brännäs (1995) extends the INAR model to incorporate explanatory variables. The regression analysis of count data is relatively new, though the statistical analysis of count data has a long and rich history. The increased availability of count data in recent years has stimulated the development of models for both panel and time series count data. For reviews of these and other models, see, e.g., Cameron and Trivedi (1998, ch. 7) and McKenzie (2003). In INARMA, the parameters are interpreted as probabilities and hence restricted to unit intervals. Some empirical applications of INAR 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.

In this thesis, we focus on advancing and employing an integer-valued moving average model of order q [INMA(q)], i.e. a special case of the INARMA model class, for analyzing high frequency financial data in the form of stock transactions data aggregated over one or five minute intervals of time. Later, we propose a bivariate integer-valued moving average (BINMA) model, a vector integer-valued moving average (VINMA) model and an integer-valued autoregressive fractionally integrated moving average (INARFIMA) model. The BINMA model is developed to capture the covariance between stock transactions data due to macroeconomic news or rumors, while the VINMA Model is more general than the BINMA model and enables the study of the spillover effects of news from one stock to other. Macroeconomic news refer to the news that may have impact on the stock markets as a whole and necessarily on a particular stocks. For example, news related to interest rates, unemployment statistics for a country, etc. may influence all stocks. Rumors are the information related to, e.g., macroeconomic news or news related to a particular stock

that spread unofficially. The INARFIMA model is developed to study the long memory property of high frequency count data. The models introduced in this thesis can also be used to measure the reaction times to shocks or news. A description of high frequency data, the INMA model, the BINMA, VINMA model, long memory and the INARFIMA model is given below.

2   High Frequency Data

Financial market data are tick-by-tick data. Each tick represents a change in, e.g., a quote or corresponds to a transaction. For a liquid stock or a currency, these tick-by-tick data generate high frequency data. Such financial data are also characterized by lack of synchronization, in the sense that only rarely is there more than one transaction at a given instant of time. For reviews of high frequency data and their characteristics, see, e.g., Tsay (2002, ch. 5), Dacorogna et al. (2001) and Gourieroux and Jasiak (2001, ch. 14). The access to high frequency data is getting less and less of a problem for individual researchers and costs are low. As a consequence, many issues related to the trading process and the market microstructure are under study.

Transactions data are collected from an electronic limited order book for each stock. Incoming orders are ranked according to price and time of entry and are continuously updated. Hence, new incoming buy and sell orders and the automatic match of the buy and sell orders are recorded. The automatic match of a buy and a sell order generates a transaction. In Figure 1, we see that the transactions in the two stocks are not synchronized, i.e. the transactions appear at different points of time. The counts in the intervals are the number of transactions for corresponding intervals. In papers [1] and [4] a one minute time scale is employed and for papers [2] and [3] a five minute scale. The collection of the number of transactions over a time period makes up a time series of count data. The time series of transactions or count data are synchronized between stocks in the sense that all the numbers of transactions are aggregated transactions over the same time interval. An example of real transactions data over a 30 minute period for the stock AstraZeneca is exhibited in Figure 2. Each observation number corresponds to one minute of time. This type of data series comprises frequent zero frequencies and motivates a count data model.

The time series of transactions or count data may have a long memory property. The long memory implies the long range dependence in the time series of counts, i.e. the present information has a persistent impact on future counts. Note that the long memory property is related to the sampling frequency of
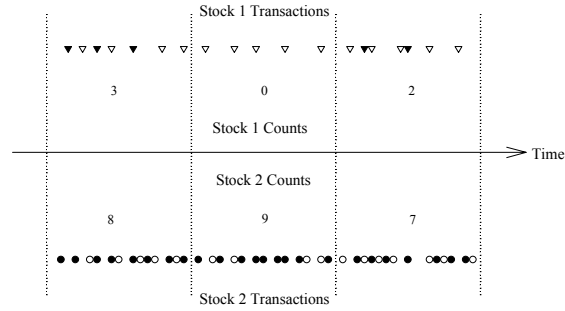
Figure 1: An illustration of how transactions data are generated. The black triangles and circles represent transactions for stock 1 and stock 2, respectively, while the white triangles and circles represent all other activities in an order book. The stock counts record the number of black triangles/circles falling into a time interval, i.e. falling between vertical lines.
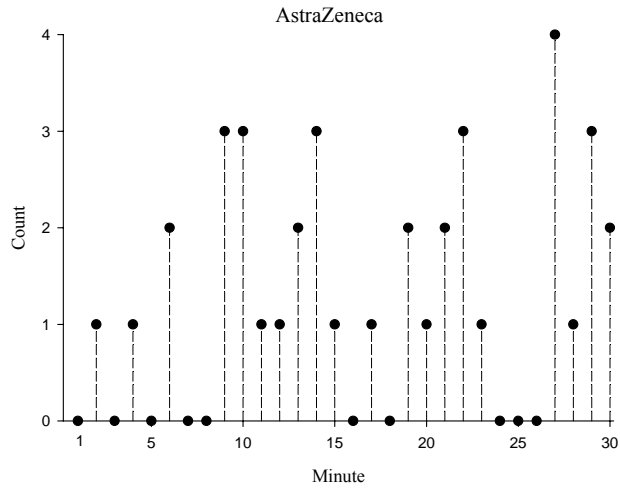
Figure 2: The number of transactions data over minute long intervals for 30 minutes of trading in AstraZeneca.
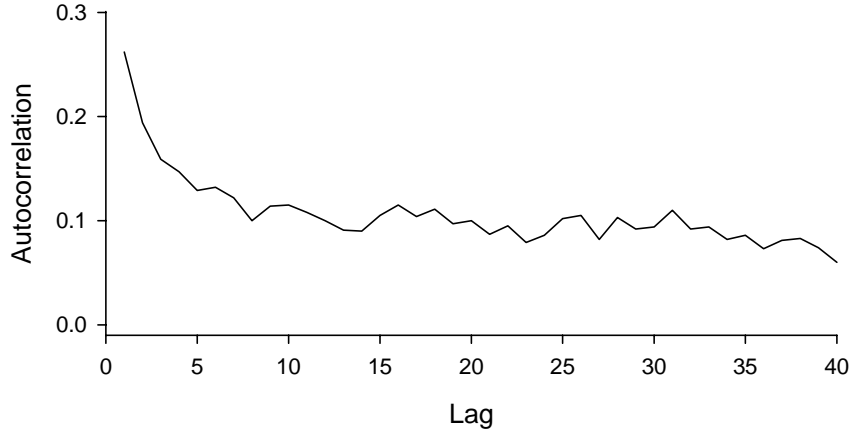
Figure 3: The autocorrelation function for AstraZeneca as an illustration of long range dependence of long memory.

a time series. A manifest long memory may be shorter than one hour if observations are recorded every minute, while stretching over decades for annual data. The time series containing long memory has a very slowly decaying autocorrelation function. The autocorrelation function for stock transactions data aggregated over one minute interval of time for AstraZeneca is illustrated in Figure 3. The autocorrelation function decays sharply in the first few lags but decays very slowly thereafter. Hence, we may expect long memory in stock transactions data for AstraZeneca. Models for long memory and continuous variable time series are not appropriate for integer-valued counts. Therefore, long memory models developed for continuous variables are not automatically of relevance neither with respect to interpretation nor to efficient estimation.

For this thesis the Ecovision system is utilized. Daily downloads are stored to files and count data are calculated from the tick-by-tick data using Matlab programs.

3   The INMA, BINMA and VINMA Models

The INMA model is a special case of the INARMA model. The INMA model of order q, INMA(q), is introduced by Al-Osh and Alzaid (1988) and in a slightly different form by McKenzie (1988). The single thing that most visibly makes the INMA model different from its continuous variable MA counterpart is that multiplication of variables with real valued parameters is no longer a viable operation, when the result is to be integer-valued. Multiplication is therefore replaced by the binomial thinning operator

$$\alpha \circ u = \sum_{i=1}^{u} v_i, \tag{1}$$

where $\{v_i\}_{i=1}^{u}$ is an iid sequence of $0-1$ random variables, such that $\Pr(v_i = 1) = \alpha = 1 - \Pr(v_i = 0)$. Conditionally on the integer-valued u, $\alpha \circ u$ is binomially distributed with $E(\alpha \circ u|u) = \alpha u$ and $V(\alpha \circ u|u) = \alpha(1-\alpha)u$. Unconditionally it holds that $E(\alpha \circ u) = \alpha\lambda$, where $E(u) = \lambda$, and $V(\alpha \circ u) = \alpha^2\sigma^2 + \alpha(1-\alpha)\lambda$, where $V(u) = \sigma^2$. Obviously, $\alpha \circ u$ takes an integer-value in the interval $[0, u]$.

Employing this binomial thinning operator, an INARMA(p, q) model can be written

$$y_t - \alpha_1 \circ y_{t-1} - \ldots - \alpha_p \circ y_{t-p} = u_t + \beta_1 \circ u_{t-1} + \ldots + \beta_q \circ u_{t-q}. \tag{2a}$$

with $\alpha_j, \beta_i \in [0, 1]$, $j = 1, \ldots, p-1$ and $i = 1, \ldots, q-1$, and $\alpha_p, \beta_q \in (0, 1]$. Setting all $\alpha_j = 0$ we obtain the INARMA(q) model

$$y_t = u_t + \beta_1 \circ u_{t-1} + \ldots + \beta_q \circ u_{t-q} \tag{2b}$$

Brännäs and Hall (2001) discuss model generalizations and interpretations resulting from different thinning operator structures, and an empirical study and approaches to estimation are reported by Brännäs et al. (2002). McKenzie (1988), Joe (1996), Jørgensen and Song (1998) and others stress exact distributional results for y_t, while we emphasize in paper [1] only the first two conditional and unconditional moments of the model. Moreover, we discuss and introduce more flexible conditional mean and heteroskedasticity specifications for y_t than implied by the above equation. There is an obvious connection between the introduced count data model and the conditional duration model of, e.g., Engle and Russell (1998) in the sense that long durations in a time interval correspond to a small count and vice versa. Hence, a main use of the

count data models discussed here is also one of measuring reaction times to shocks or news.

In paper [2], we focus on the modelling of bivariate time series of count data that are generated from stock transactions. The used data are aggregates over five minutes intervals and computed from tick-by-tick data. One obvious advantage of the introduced model over the conditional duration model is that there is no synchronization problem between the time series.[1] 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 is employed to capture covariance between stock transactions time series and to measure the reaction time for news or rumors. Moreover, this model is capable of capturing the conditional heteroskedasticity.

In paper [3], we extend the INMA model to a vector INMA (VINMA) model. The VINMA is more general than the BINMA model in paper [2] and enables the study of the spillover effects of transactions from one stock to the other.

A large number of studies have considered the modelling of bivariate or multivariate count data assuming an underlying Poisson distribution (e.g., Gourieroux, Monfort and Trognon, 1984). Heinen and Rengifo (2003) introduce multivariate time series count data models based on the Poisson and the double Poisson distribution. Other extensions to traditional count data regression models are considered by, e.g., Brännäs and Brännäs (2004) and Rydberg and Shephard (1999).

4   Long Memory and the INARFIMA Model

Hurst (1951, 1956) considered first the long memory phenomenon in time series. He explained the long term storage requirements of the Nile River. He showed that the cumulated water flows in a year had a persistent impact on the water flows in the later years. By employing fractional Brownian motion, Mandelbrot and van Ness (1968) explain and advance the Hurst's studies. In analogy with Mandelbrot and van Ness (1968), Granger (1980), Granger and Joyeux (1980) and Hosking (1981) develop Autoregressive Fractionally Integrated Moving Average (ARFIMA) models to account for the long memory in

---

[1]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.

time series data. According to Ding and Granger (1996), a number of other
processes can also have the long memory property. In a recent empirical study,
Bhardwaja and Swanson (2005) found strong evidence in favor of ARFIMA in
absolute, squared and log-squared stock index returns.

Granger and Joyeux (1980) and Hosking (1981) independently propose
ARFIMA processes to account for long memory in continuous variables. We
say that $\{y_t, t = 1, 2, \ldots, T\}$ is an ARFIMA $(0, d, 0)$ process if

$$(1 - L)^d y_t = a_t \tag{3}$$

where L is a lag operator and d is any real number. The $\{a_t\}$ is a white noise
process of random variables with mean $E(a_t) = 0$ and variance $V(a_t) = \sigma_a^2$.
Employing binomial series expansion, we can write

$$(1 - L)^d = \Delta^d = 1 - \sum_{i=1}^{\infty} \frac{(i - 1 - d)!}{i!(-d - 1)!} L^i = 1 - \sum_{i=1}^{\infty} \frac{\Gamma(i - d)}{\Gamma(i + 1)\Gamma(1 - d)} L^i \tag{4}$$

and correspondingly

$$\begin{aligned}
\Delta^{-d} &= 1 + dL + \frac{1}{2}d(1 + d)L^2 + \frac{1}{6}d(1 + d)(2 + d)L^3 - \ldots \\
&= 1 + \sum_{i=1}^{\infty} \frac{(i + d - 1)!}{i!(d - 1)!} L^i = 1 + \sum_{i=1}^{\infty} \frac{\Gamma(i + d)}{\Gamma(i + 1)\Gamma(d)} L^i \tag{5}
\end{aligned}$$

where $\Gamma(n + 1) = n!$ and $i = 1, 2, \ldots$. The Δ^d is needed for AR(∞) and the
Δ^{-d} is needed for MA(∞) representations of the ARFIMA $(0, d, 0)$ model or for
more general ARFIMA(p, d, q) models. If $d < 1/2$, $d \neq 0$, the ARFIMA$(0, d, 0)$
process is called a long memory process, while the process has mean reversion
but is not covariance stationary when $d > 1/2$. A survey of the ARFIMA
literature can be found in Baillie (1996). Note, for instance, that the AR and
MA parameters of an ARFIMA model are less restricted than the corresponding
parameters of the INARFIMA model.

In paper [4], we focus on modelling the long memory property of time
series of count data and on applying the model in a financial setting. Com-
bining the ideas of the INARMA model (2a) with fractional integration is not
quite straightforward. Direct use of (4) or (5) will not give integer-values since
multiplying an integer-valued variable with a real-valued d can not produce an
integer-valued result and this alternative is hence ruled out. Instead, we depart
from the binomial expansion expression and propose in analogy with Granger
and Joyeux (1980) and Hosking (1981) INARFIMA models that accounts for
integer-valued counts and long memory. We apply the INARFIMA models to

stock transactions data for AstraZeneca and Ericsson B. We found evidence for
long memory for the AstraZeneca series while the series for Ericsson B indicates
a process indicates a process that has a mean reversion property.

5   Summary of the Papers

**Paper [1]: Integer-Valued Moving Average Modelling of the Number
of Transactions in Stocks**

The integer-valued moving average model is advanced to model the number of
transactions in intra-day data of stocks. The conditional mean and variance
properties are discussed and model extensions to include, e.g., explanatory vari-
ables are offered. Least squares and generalized method of moment estimators
are presented. In a small Monte Carlo experiment we study the bias and MSE
properties of the CLS, FGLS and GMM estimators for finite-lag specifications,
when data is generated according to an infinite-lag INMA model. In addition,
we study the serial correlation properties of estimated models by the Ljung-
Box statistic as well as the properties of forecasts one and two steps ahead.
In this Monte Carlo study, the feasible least squares estimator comes out as
the best choice. However, the CLS estimator which is the simplest to use of
the three considered estimators is not far behind. The GMM performance is
weaker than that of the CLS estimator. It is also clear that the lag length
should be chosen large and that both under and overparameterization give rise
to detectable serial correlation.

In its practical implementation for the time series of the number of trans-
actions in Ericsson B, we found both promising and less advantageous features
of the model. There is evidence of asymmetric effects of news about prices
on the number of transactions. With the CLS estimator it was relatively easy
to model the conditional mean in a satisfactory way in terms of both inter-
pretation and residual properties. It was more difficult to obtain satisfactory
squared residual properties for the conditional variance specifications that were
tried. The FGLS estimator reversed this picture and we suggest that more em-
pirical research is needed on the interplay between the conditional mean and
heteroskedasticity specifications for count data. Depending on research inter-
est the conditional variance parameters are or are not of particular interest.
For studying reaction times to shocks or news it is the conditional mean that
matters, in much the same way as for conditional duration models. In addition,
the conditional variance has no direct ties to, e.g., risk measures included in,

e.g., option values or portfolios.

Paper [2]: Bivariate Time Series Modelling of Financial Count Data

This study introduces a bivariate integer-valued moving average (BINMA) model 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 mean, variance and covariance are given. The study shows that the correlation between series in the BINMA model is always smaller than one in an absolute sense. Applying the BINMA model 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 is not of particular interest in this study. Further study is required to eliminate such serial correlation. One way of eliminating serial correlation may be to use extended model by letting, e.g., λ_j or σ_j be time-varying.

Paper [3]: A Vector Integer-Valued Moving Average Model for High Frequency Financial Count Data

This paper introduces a Vector Integer-Valued Moving Average (VINMA) model. The VINMA is developed to capture covariance between stock transactions time series. The Model allows for both positive and negative correlation between the count series and the integer-value property of counts is taken into account. The model is capable of capturing the covariance between and within intra-day time series of transaction frequency data due to macroeconomic news and news related to a specific stock. The conditional and unconditional first and second order moments are obtained. The CLS and FGLS estimators are discussed. The FGLS estimator performs better than CLS in terms of eliminating serial correlation. The VINMA model performs better than the BINMA of paper [3] in terms of goodness of fit. Empirically, it is found that the spillover effect from Ericsson B to AstraZeneca is larger than that from AstraZeneca to Ericsson B.

Paper [4]: Long Memory, Count Data, Time Series Modelling for Financial Applications

This paper introduces a model to account for the long memory property in a count data framework. The model emerges from the ARFIMA and INARMA

model classes and hence the model is called INARFIMA. The unconditional and conditional first and second order moments are given. Moreover, we introduce another process by employing an idea introduced by Granger, Joyeux and Hosking but in a different setting. The model is successfully applied to estimate the fractional integration parameter for high frequency financial count data for two stock series for Ericsson B and AstraZeneca.

In order to study residual properties for standardized residual we estimate several INARFIMA models and truncated INMA models. The INMA(70) and INMA(50) for Ericsson B and AstraZeneca, respectively, turns out to be the best in terms of eliminating serial correlation for standardized residuals while INARFIMA$(0, \delta, 0)$ comes in as second best for both series and the estimated parameters are positive. The INARFIMA$(0, \delta, 0)$ is the most parsimonious model in terms of number of parameters. For AstraZeneca, we find evidence of long memory, while the estimated δ for Ericsson B indicates a process that has a mean reversion property. CLS and FGLS estimators perform equally well in terms of residual properties. We also find that the trading intensity increases for both stocks when the macro-economic news or rumors break out and the impact of the macro-economic news remains over a long period and fades away very slowly with time. The reaction due to the macro-economic news on the AstraZeneca series is faster than that of the Ericsson B series.

References

Al-Osh, M.A. and Alzaid, A.A. (1987). First Order Integer-Valued Autoregressive (INAR(1)) Process. *Journal of Time Series Analysis* **8**, 261-275.

Al-Osh M. and Alzaid A. (1988). Integer-Valued Moving Average (INMA) Process. *Statistical Papers* **29**, 281-300.

Bhardwaj, G. and Swanson, N.R. (2005). An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. *Journal of Econometrics*. Article in Press.

Blundell, R., Griffith, R. and Windmeijer, F. (2002). Individual Effects and Dynamics in Count Data Models. *Journal of Econometrics* **108**, 113-131.

Brännäs, K. (1995). Explanatory Variables in the AR(1) Model. *Umeå Economic Studies* **381**.

Brännäs, K., Hellström, J. and Nordström, J. (2002). A New Approach to Modelling and Forecasting Monthly Guest Nights in Hotels. *International Journal of Forecasting* **18**, 19-30.

Brännäs, K. and Hall, A. (2001). Estimation in Integer-Valued Moving Average Models. *Applies Stochastic Models in Business and Industry* **17**, 277-291.

Brännäs, K. and Brännäs, E. (2004). Conditional Variance in Count Data Regression. *Communication in Statistics; Theory and Methods* **33**, 2745-2758.

Cameron, A.C. and Trivedi, P.K. (1998). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Dacorogna, M.M., Gencay, R., Muller, U., Olsen, R.B. and Pictet, O.V. (2001). *An Introduction to High-Frequency Finance*. San Diego: Elsevier.

Demsetz, H. (1968). The Cost of Transacting. *Quarterly Journal of Economics* **82**, 33-53.

Diamond, D.W. and Verrecchia, R.E. (1987). Constraints on Short-Selling and Asset Price Adjustments to Private Information. *Journal of Financial Economics* **18**, 277-311.

Ding, Z. and Granger, C.W.J. (1996). Varieties of Long Memory Models. *Journal of Econometrics* **73**, 61-77.

Easley, D. and O'Hara, M. (1992). Time and the Process of Security Price Adjustment. *Journal of Finance* **47**, 905-927.

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

Engle, R.F. and Russell, J.R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* **66**, 1127-1162.

Gourieroux, C and Jasiak, J. (2001). *Financial Econometrics.* Princeton: Princeton University Press.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Application to Poisson Models. *Econometrica* **52**, 701-720.

Granger, C.W.J. (1980). Long Memory Relationships and the Aggregation of Dynamic Models. *Journal of Econometrics* **14**, 227-238.

Granger, C.W.J. and Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis* **1**, 15-29.

Heinen, A. and Rengifo, E. (2003). Multivariate Modelling of Time Series Count Data: An Autoregressive Conditional Poisson Model. CORE Discussion Paper 2003/25, Université Catholique de Louvain.

Hosking, J.R.M. (1981). Fractional Differencing. *Biometrika*, **68**, 165-176.

Hurst, H.E. (1951). Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers* **116**, 770-808.

Hurst, H.E. (1956). Methods of Using Long-Term Storage in Reservoirs. *Proceedings of the Institute of Civil Engineers* **1**, 519-543.

Jacobs, P.A. and Lewis, P.A.W. (1978a). Discrete Time Series Generalized by Mixtures I: Correlational and Runs Properties. *Journal of the Royal Statistical Society* **B40**, 94-105.

Jacobs, P.A. and Lewis, P.A.W. (1978b). Discrete Time Series Generalized by Mixtures II: Asymptotic Properties. *Journal of the Royal Statistical Society* **B40**, 222-228.

Jacobs, P.A. and Lewis, P.A.W. (1983). Stationary Discrete Autoregressive Moving Average Time Series Generated by Mixtures. *Journal of Time Series Analysis* **4**, 19-36.

Joe, H. (1996). Time Series Models with Univariate Margins in the Convolution-Closed Infinitely Divisible Class. *Journal of Applied Probability* **33**, 664-677.

Jørgensen, B. and Song, P.X-K. (1998). Stationary Time Series Models with Exponential Dispersion Model Margins. *Journal of Applied Probability* **35**, 78-92.

Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review* **10**, 422-437.

McKenzie, E. (1986). Autoregressive Moving-Average Processes with Negative Binomial and Geometric Marginal Distributions. *Advances in Applied Probability* **18**, 679-705.

McKenzie, E. (2003). Discrete Variate Time Series. In Shanbhag, D.N. and

Rao, C.R. (eds.) *Handbook of Statistics*, Volume 21, pp. 573-606. Amsterdam: Elsevier Sciences.

McKenzie, E. (1988). Some ARMA models for Dependent Sequences of Poisson Counts. *Advances in Applied Probability* **20**, 822-835.

Rudholm, N. (2001). Entry and the Number of Firms in the Swedish Pharmaceuticals Market. *Review of Industrial Organization* **19**, 351-364.

Rydberg, T.H. and Shephard, N. (1999). BIN Models for Trade-by-trade Data. Modelling the Number of Trades in a Fixed Interval of Time. Working Paper Series W23. Nuffield College, Oxford.

Tsay, R.S. (2002). *Analysis of Financial Time Series*. New York: John Wiley & Sons, Inc.

Working, H. (1953). Futures Trading and Hedging. *American Economic Review* **43**, 314-343.

Integer-Valued Moving Average Modelling of the Number of Transactions in Stocks*

Kurt Brännäs and A.M.M. Shahiduzzaman Quoreshi
Department of Economics, Umeå University
SE-901 87 Umeå, Sweden

Abstract

The integer-valued moving average model is advanced to model the number of transactions in intra-day data of stocks. The conditional mean and variance properties are discussed and model extensions to include, e.g., explanatory variables are offered. Least squares and generalized method of moment estimators are presented. In a small Monte Carlo study a feasible least squares estimator comes out as the best choice. Empirically we find support for the use of long-lag moving average models in a Swedish stock series. There is evidence of asymmetric effects of news about prices on the number of transactions.

Key Words: Count data, Intra-day, High frequency, Time series, Estimation, Finance.
JEL Classification: C13, C22, C25, C51, G12, G14.

1 Introduction

The paper focuses on modelling time series of the number of intra-day transactions in stocks. Such count data time series typically take on small numbers even for frequently traded stocks if the counts are recorded in short time intervals of, for instance, one minute length. There is an obvious connection between the current count data model and the conditional duration model of, e.g., Engle and Russell (1998) in the sense that long durations in a time interval correspond to a small count and vice versa. Hence, a main use of the count data models discussed here is also one of measuring reaction times to shocks or news. The relationship between durations (and indirectly counts) and the price process is reviewed by Chiang and Wang (2004). In this context straightforward use of the Box-Jenkins methodology for identifying parsimonious time series models raises fundamental questions about the resulting model specification when we wish to adhere to the integer-valued or count data nature of the number of transactions variable.

Previous models for the number of transactions or related variables within the intra-day financial arena have departed from conventional count data regression models or from extended model versions (e.g., Brännäs and Brännäs, 2004; Heinen and Rengifo, 2003). Here, we consider a different approach and start from an integer-valued model corresponding to the conventional ARMA class of Box and Jenkins (1970). An important difference between the continuous variable ARMA model and its corresponding integer-valued version (INARMA) is that the latter contains parameters that are interpreted as probabilities and then take on values in narrower intervals than do the parameters of the ARMA model (e.g., McKenzie, 1986; Al-Osh and Alzaid, 1991; Joe, 1996; Jørgensen and Song, 1998; McKenzie, 2003). This then may make model identification of an appropriate model by correlation methods (Box and Jenkins, 1970) less suitable. While for the ARMA class, specification searches aim at models that leave no serial correlation and satisfy stationarity and invertibility criteria, the INARMA specification should additionally have each parameter estimate of lagged variables in the unit interval.

In this paper the empirical results indicate that long-lag INMA models satisfy such restrictions, while mixed INARMA models do not. The INMA model class has been studied by, e.g., Al-Osh and Alzaid (1988), McKenzie (1988) and Brännäs and Hall (2001). As far as we are aware the only published, empirical application is due to Brännäs, Hellström and Nordström (2002), who estimated a nonlinear INMA(1) model for tourism demand. This model had time dependent parameters that were functions of explanatory variables. The

present empirical study is focused on a stock transaction series (Ericsson) registered at the order driven Stockholmsbörsen stock exchange and emphasizes different specification issues.

The INMA model is introduced in Section 2, where we also give some moment results and discuss conditional heteroskedasticity. General expressions for conditional and unconditional moments are obtained. Extensions to include explanatory variables in the conditional mean and a more flexible conditional heteroskedasticity specification are also discussed. Section 3 discusses estimation of unknown parameters and gives least squares and GMM estimators for the model class. The section also considers aspects of model evaluation and forecasting. A small Monte Carlo experiment to study some key characteristics of the estimators and of forecasts is included. Section 4 contains the empirical results for the stock series. The final section offers some concluding comments.

2   Model

The single thing that most visibly makes the integer-valued MA (INMA) model different from its continuous variable MA counterpart is that multiplication of variables with real valued parameters is no longer a viable operation, when the result is to be integer-valued. Multiplication is therefore replaced by the binomial thinning operator

$$\alpha \circ u = \sum_{i=1}^{u} v_i,$$

where $\{v_i\}_{i=1}^{u}$ is an iid sequence of $0-1$ random variables, such that $\Pr(v_i = 1) = \alpha = 1 - \Pr(v_i = 0)$. Conditionally on the integer-valued u, $\alpha \circ u$ is binomially distributed with $E(\alpha \circ u|u) = \alpha u$ and $V(\alpha \circ u|u) = \alpha(1-\alpha)u$. Unconditionally it holds that $E(\alpha \circ u) = \alpha\lambda$, where $E(u) = \lambda$, and $V(\alpha \circ u) = \alpha^2\sigma^2 + \alpha(1-\alpha)\lambda$, where $V(u) = \sigma^2$. Obviously, $\alpha \circ u \in [0, u]$.

Employing this binomial thinning operator, the INMA(∞) model can be written

$$y_t = \sum_{i=0}^{\infty} \beta_i \circ u_{t-i}, \tag{1}$$

with mostly $\beta_0 = 1$.[1] The $\{u_t\}$ is an iid sequence of non-negative and integer-valued random variables with, as above, $E(u) = \lambda$ and $V(u) = \sigma^2$.

---

[1] The INMA(∞) can, e.g., be obtained from the INAR(1), i.e. $y_t = \alpha \circ y_{t-1} + \varepsilon_t$ and $y_t = \alpha^t \circ y_0 + \sum_{i=1}^{t} \alpha^{t-i} \circ \varepsilon_i$ are equal in distribution. Assuming $\alpha \in [0, 1)$ and t large gives that $\alpha^t \approx 0$ and $\beta_i = \alpha^i$.

McKenzie (1988), Joe (1996), Jørgensen and Song (1998) and others stress exact distributional results for y_t, while we emphasize only the first two conditional and unconditional moments of the model. One reason for our choice will become apparent below when we discuss and introduce more flexible conditional mean and heteroskedasticity specifications for y_t than implied by (1). As a consequence exact maximum likelihood (ML) estimation is beyond reach though otherwise a desirable candidate for estimation. We could potentially use ML estimation by directly specifying a conditional density for y_t given its history Y_{t-1}; cf. the conditional duration model approach of Engle and Russell (1998).

The finite-lag INMA(q) model

$$y_t = u_t + \beta_1 \circ u_{t-1} + \ldots + \beta_q \circ u_{t-q} \tag{2}$$

was introduced by McKenzie (1986). Brännäs and Hall (2001) discuss model generalizations and interpretations resulting from different thinning operator structures, and an empirical study and approaches to estimation are reported by Brännäs et al. (2002).

For the INMA(∞) model in (1), with independence between and within thinning operations,[2] and with $\{u_t\}$ an iid Poisson sequence with $\sigma^2 = \lambda$, and $\beta_0 = 1$, the moment expressions are:

$$E(y_t) = V(y_t) = \lambda \left(1 + \sum_{i=1}^{\infty} \beta_i \right) \tag{3a}$$

$$\rho_k = \lambda \left(\beta_k + \sum_{i=1}^{\infty} \beta_i \beta_{k+i} \right) / V(y_t), \qquad k \geq 1. \tag{3b}$$

It is obvious from these moments that they only generate positive values and that $\sum_{i=0}^{\infty} \beta_i < \infty$ is required for $\{y_t\}$ to be a stationary sequence. Assuming instead, e.g., an iid distribution with mean λ and variance σ^2 changes the variance and the autocorrelation function to

$$V(y_t) = \lambda \sum_{i=1}^{\infty} \beta_i (1 - \beta_i) + \sigma^2 \left(1 + \sum_{i=1}^{\infty} \beta_i^2 \right) \tag{3c}$$

$$\rho_k = \sigma^2 \left(\beta_k + \sum_{i=1}^{\infty} \beta_i \beta_{k+i} \right) / V(y_t), \qquad k \geq 1, \tag{3d}$$

while leaving the mean unaffected and as in (3a).

---

[2] Pairs of thinning operations of the type $\theta_i \circ u_t$ and $\theta_j \circ u_t$, for $i \neq j$, are independent (McKenzie, 1988). Assumptions of this type can be relaxed (cf. Brännäs and Hall, 2001).

In case the lag length q is finite, summing to infinity is replaced by summing to q in (3c)-(3d) and to $q - k$ in the numerator of (3d). In that case $\rho_k = 0$, for $k > q$. Figure 1 gives an illustration of two autocorrelation functions when $\beta_i, i \geq 1$, are set as in the Monte Carlo experiment, below.
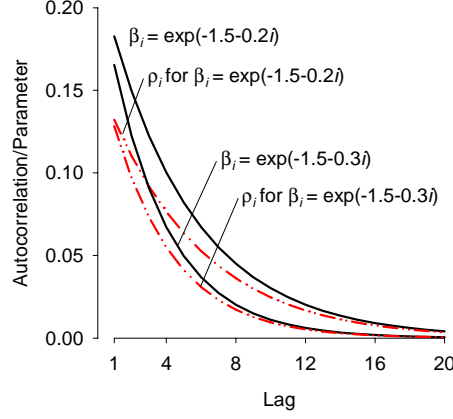
Figure 1: Autocorrelation functions in the Poisson case (dot-dashed lines) and β_i parameters (solid lines) when parameters are as in the Monte Carlo study, below.

The conditional moments for the INMA(∞) model are:

$$E(y_t|Y_{t-1}) = \lambda + \sum_{i=1}^{\infty} \beta_i u_{t-i} \tag{4a}$$

$$V(y_t|Y_{t-1}) = \sigma^2 + \sum_{i=1}^{\infty} \beta_i (1 - \beta_i) u_{t-i}, \tag{4b}$$

where Y_{t-1} is the information set available at time $t-1$.

As the conditional variance varies with $u_{t-i}, i \geq 1$, there is conditional heteroskedasticity of a moving average type. We can call this property MACH(∞) and for finite q, MACH(q) (cf. the notation ARCH(p)). Note that the response to lagged u_{t-i}s is weaker in the conditional variance (4b) than in (4a) as $\beta_i \in [0, 1]$. The relative size of the two moments will largely depend on the sizes of λ and σ^2. Note also that even though a Poisson distribution for u_t implies a Poisson distributed y_t, the conditional distribution of y_t given its

past is not the Poisson, as indicated, e.g., by the difference between (4a) and (4b).

As measures of reaction times to shocks/news in the $\{u_t\}$ sequence we may use the mean lag $\sum_{i=0}^{\infty} i\beta_i/w$, where $w = \sum_{i=0}^{\infty} \beta_i$. Alternatively, we may use the median lag, which is the smallest k such that $\sum_{i=0}^{k} \beta_i/w \geq 0.5$ (e.g., Greene, 2003, ch. 19).

2.1 Extensions

There are two obvious extensions to the model that appear of high empirical relevance. First, we may let λ become time-varying and a function of explanatory variables. The natural specification is

$$\lambda_t = \exp(\mathbf{x}_t \boldsymbol{\theta}) \geq 0, \tag{5}$$

where in $\mathbf{x}_t$ we may include k variables related to previous prices, etc. A consequence of the time-varying λ_t is that moment expressions become time dependent, but the additional difficulty with respect to estimation is marginal.

Another obvious extension in order to obtain more flexible conditional variance specifications in (4b) is to let σ^2 become time dependent. We may let σ_t^2 depend on past values on σ_t^2, u_t and explanatory variables in, e.g., the following exponential way

$$
\begin{aligned}
\sigma_t^2 &= \exp\Big[\phi_0 + \phi_1 \ln \sigma_{t-1}^2 + \ldots + \phi_P \ln \sigma_{t-P}^2 \\
&\quad + \gamma_1 (u_{t-1} - \lambda)^2 + \ldots + \gamma_Q (u_{t-Q} - \lambda)^2 + \mathbf{x}_t \boldsymbol{\alpha}\Big]
\end{aligned}
\tag{6}
$$

(cf. Nelson, 1991). There could also be additional contributions, or at least different ones, to the conditional variance if, e.g., the different thinning operations were dependent (cf. Brännäs and Hall, 2001; Brännäs and Hellström, 2001).

It is also possible to let β_i be, e.g., a logistic function of explanatory variables (Brännäs et al., 2002) and to reduce the number of β_is by specifying a distributed lag distribution of, e.g., the form $\beta_i = \delta_0 \exp(-\delta_1 i)$ with $\delta_0 \in (0, 1]$ and $\delta_1 \geq 0$ for $i \geq 1$.

3 Estimation

In this section we discuss approaches to the estimation of the unknown parameters of the conditional mean and variance functions. Both the conditional

mean and the conditional variance will contain time dependent λ_t and σ_t^2 specifications, respectively, unless otherwise stated. As we do not assume a full density specification the proposed estimators can be viewed as semiparametric ones.

The conditional least squares (CLS) estimator is focused on the unknown parameters of the conditional mean function and requires an additional step to estimate the unknown parameters of the conditional variance expression, in particular in its σ_t^2 part. Typically, CLS estimates of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)'$, and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)'$ when λ_t is time-varying, are used and kept fixed when estimating the variance function. In a feasible generalized least squares (FGLS) estimator, these two steps facilitate GLS estimation of the conditional mean function in a final step. Note that for small β_i parameters we may expect the conditional variance to be almost constant if $\sigma_t^2 = \sigma^2$ holds (cf. (4b)). In such an instance we expect the CLS and FGLS estimators of the parameters to be numerically close.

By a GMM estimator (Hansen, 1982), all parameters can be estimated jointly. For the GMM estimator weighting is with respect to moment conditions and not with respect to individual observations as in FGLS. We may anticipate better performance of the FGLS than of the GMM estimator for the parameters of the conditional mean function (Brännäs, 1995). Brännäs and Hall (2001) found the CLS estimator to have weaker bias/mean square error (MSE) performance than a GMM estimator based on probability generating functions. This type of GMM estimator will not be considered here as it is computationally more intricate and currently rests on arbitrarily setting values on the argument of the generating function.

Common to the considered estimators is their reliance on the prediction error

$$e_{1t} = y_t - E(y_t|Y_{t-1}). \tag{7}$$

The CLS estimator minimizes the criterion function $S_{CLS} = \sum_{t=r}^{T} e_{1t}^2$, where $r = q+1$ and T is the time series length, with respect to the unknown parameter vector $\boldsymbol{\psi}' = (\boldsymbol{\theta}', \boldsymbol{\beta}')$. To calculate the e_{1t} sequence we consider $e_{1t} = u_t - \lambda_t + \sum_{i=1}^{\infty} (\beta_i \circ u_{t-i} - \beta_i u_{t-i})$ and advocate setting the sum to zero rather than using some randomization device. When standard software based on the assumption $E(u_t) = \lambda = 0$ is used to CLS estimate an INMA(q) model, the obtained estimate of the constant term is an estimate of $\lambda \sum_{i=0}^{q} \beta_i$. Given estimates of β_i it is hence possible to obtain an estimate of λ manually.

Obviously, there are alternative estimators, such as Durbin's (1959) estimator extended to handle λ, for the estimation of $\boldsymbol{\psi}$. A recent summary of

least squares and related estimators (interpreted as GMM) of low order MA(q) models utilizing dual AR representations is given in Harris (1999). Method of moment (or GMM) estimation based on the unconditional first and second moments requires the solution of a system of nonlinear equations. Hence, a simplicity argument does not apply and moreover the properties of the estimator are for the MA(q) model not very satisfactory.

For the second step, normal equations based on the conditional variance prediction error

$$e_{2t} = (y_t - E(y_t|Y_{t-1}))^2 - V(y_t|Y_{t-1}) \qquad (8)$$

are used for FGLS estimation, and incorporated as moment conditions for GMM estimation. Here too the conditional variance specification should have a say on the choice of instrumental variables. For FGLS $S_2 = \sum_{t=s}^{T} e_{2t}^2$, where $s = \max(q, P, Q) + 1$, is minimized with respect to the parameters of the σ_t^2 function, i.e. $\boldsymbol{\omega}' = (\sigma^2, \phi_0, \ldots, \phi_P, \gamma_1, \ldots, \gamma_Q, \boldsymbol{\alpha}')$ and with the CLS estimates $\hat{\boldsymbol{\psi}}$ and $\{\hat{u}_t\}$ kept fixed. In case σ^2 is time invariant an obvious estimator is of the simple form

$$\hat{\sigma}^2 = (T - s)^{-1} \sum_{t=s}^{T} \left[\hat{e}_{1t}^2 - \sum_{i=1}^{q} \hat{\beta}_i (1 - \hat{\beta}_i) \hat{u}_{t-i} \right].$$

For the third step of FGLS, minimizing the criterion

$$S_{FGLS} = \sum_{t=s}^{T} e_{1t}^2 / \hat{V}(y_t|Y_{t-1})$$

with $\hat{V}$ taken as given, gives the FGLS estimates of the parameter vector $\boldsymbol{\psi}$ of the conditional mean function. The covariance matrix estimators are:

$$Cov(\hat{\boldsymbol{\psi}}_{CLS}) = \left(\sum_{t=r}^{T} \frac{\partial e_{1t}}{\partial \boldsymbol{\psi}} \frac{\partial e_{1t}}{\partial \boldsymbol{\psi}'} \right)^{-1} = \mathbf{A}^{-1}$$

$$Cov(\hat{\boldsymbol{\psi}}_{FGLS}) = \left(\sum_{t=r}^{T} \hat{V}^{-1}(y_t|Y_{t-1}) \frac{\partial^2 e_{1t}}{\partial \boldsymbol{\psi} \partial \boldsymbol{\psi}'} \right)^{-1} = \mathbf{B}^{-1}.$$

A robust estimator for the CLS estimator is of the form $\mathbf{A}^{-1}\mathbf{B}_c\mathbf{A}^{-1}$, where $\mathbf{B}_c$ is as $\mathbf{B}$ with $\hat{V}^{-1}(y_t|Y_{t-1})$ replaced by $\hat{e}_{1t}^2$.

Using a finite maximum lag q for a true INMA(∞) model can be expected to have a biasing effect on the estimator of the constant term λ of the conditional mean function of the INMA(q) model. The conditional expectation of the difference between the infinite and finite models is $\sum_{i=q+1}^{\infty} \beta_i u_{t-i}$, which has

expectation $\lambda \sum_{i=q+1}^{\infty} \beta_i$. For large q the latter sum will be close to zero as we expect $\beta_i, i \geq q+1$, to approach zero for large q. An estimate of the constant term can then be expected to be only moderately too large. In analogy with a linear regression model and the OLS estimator the conventional analysis of the consequences of omitted variables, i.e. $u_{t-q-1}, \ldots, u_{t-\infty}$, suggests that we should expect a positive bias in the estimates of β_i for small q. This is so, since all $\beta_i > 0$ and the covariance between included and incorrectly excluded lags of u_t is always λ^2 as $E(u_t u_{t-j}) = \lambda^2, j \neq 0$, under independence. By a related argument, we expect no additional bias if q is chosen larger than some true value q^*. An immediate consequence of these intuitive arguments is that q should, at least, initially be chosen large. Subsequent testing could later be used to reduce the initial q.

The GMM criterion function

$$q = \mathbf{m}' \hat{\mathbf{W}}^{-1} \mathbf{m}$$

has the vector of moment conditions $\mathbf{m}$ depending on the specification and is minimized with respect to $\boldsymbol{\eta}' = (\boldsymbol{\psi}', \boldsymbol{\omega}')$. The moment conditions corresponding to the conditional mean are collected into

$$\mathbf{m}_1 = (T - m)^{-1} \sum_{t=m+1}^{T} \mathbf{m}_{1t},$$

where $m \geq q$. In the Monte Carlo study below we use $e_{1t}, e_{1t}e_{1t-1}, \ldots, e_{1t}e_{1t-m}$ for $\mathbf{m}_1$, while for the empirical results we use the first order condition of the CLS estimator, i.e. $e_{1t} \partial e_{1t}/\partial \psi_k = 0$, to give the conditions. For the conditional variance the moment conditions are formed from $E(e_{1t}^2) - \sigma_t^2 - \sum_{i=1}^{q} \beta_i(1-\beta_i)u_{t-i}$ with instruments from the collection of $\sigma_{t-i}^2, e_{1t-i}^2$ and x_{kt}. The conditions are collected into a vector $\mathbf{m}_2$. Finally, $\mathbf{m} = (\mathbf{m}_1', \mathbf{m}_2')'$. As a consistent estimator of the weight matrix $\mathbf{W}$ we use

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

The covariance matrix of the parameter estimator is, when $\mathbf{W}$ is set equal to an identity matrix, estimated by $Cov(\hat{\boldsymbol{\eta}}) = (T-m)^{-1}(\hat{\mathbf{G}}'\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}'\hat{\boldsymbol{\Gamma}}\hat{\mathbf{G}}(\hat{\mathbf{G}}'\hat{\mathbf{G}})^{-1}$, where $\hat{\mathbf{G}} = \partial \hat{\mathbf{m}}/\partial \hat{\boldsymbol{\eta}}'$. When the numbers of moment conditions and parameters are equal; $Cov(\hat{\boldsymbol{\eta}}) = (T-m)^{-1}\hat{\mathbf{G}}^{-1}\hat{\boldsymbol{\Gamma}}(\hat{\mathbf{G}}')^{-1}$.

3.1 Model Evaluation

To test against serial correlation in standardized residuals $\hat{e}_{1t}/\hat{V}^{1/2}(y_t|Y_{t-1})$ and squared standardized residuals we may use the Ljung-Box test statistic $LB_K = T(T+2)\sum_{i=1}^{K} r_i^2/\,(T-i)$, where r_i is the lag i autocorrelation of the standardized residual. Under homoskedasticity and independence the test statistic is asymptotically $\chi^2(K)$ distributed. Davis and Dunsmuir (2000) and Lobato, Nankervis and Savin (2001) recently considered corrections to the Ljung-Box statistic when heteroskedasticity and serial correlation are present.

3.2 Forecasting

For an INMA(∞) model the forecasts $\mu_{T+h|T} = E(y_{T+h}|Y_T), h \geq 1$, are

$$\mu_{T+h|T} = \lambda \sum_{i=0}^{h-1} \beta_i + \sum_{i=h}^{\infty} \beta_i u_{T+h-i}, \qquad h \geq 1.$$

The limiting value of the forecast as $h \to \infty$ is $\lambda \sum_{i=0}^{\infty} \beta_i$, which is the mean of the process. For finite q the sum to infinity in the second term is replaced by summing up to q for $h \leq q$. For $h > q$, the forecast is again equal to the mean of the process, i.e. $\lambda \sum_{i=0}^{q} \beta_i$.

The variance of the forecast error $e_{T+h} = y_{T+h} - \mu_{T+h|T}$ is for the INMA(∞) model with known parameters

$$s_{T+h|T} = \sigma^2 \sum_{i=0}^{h-1} \beta_i^2 + \lambda \sum_{i=1}^{\infty} \beta_i(1 - \beta_i).$$

For $h \leq q$, the forecast error variance for the INMA(q) model is $s_{T+h|T} = \sigma^2 \sum_{i=0}^{h-1} \beta_i^2 + \lambda \sum_{i=1}^{q} \beta_i(1 - \beta_i)$.

Obviously, the uncertainty in estimated parameters will increase these variances. To obtain expressions for such variances we could consider various approximations to $V(e_{T+h}) = E_{\hat{\eta}}(s_{T+h|T,\hat{\eta}}) + V_{\hat{\eta}}(\mu_{T+h|T,\hat{\eta}})$. Using first order Taylor expansions of the two terms around the true parameter vector $\boldsymbol{\eta}$ we get the approximative variance $V(e_{T+h}) \approx s_{T+h|h} + \mathbf{g}'\left[E(\hat{\boldsymbol{\eta}}) - \boldsymbol{\eta}\right] + \mathbf{h}'Cov(\hat{\boldsymbol{\eta}})\mathbf{h}$, where $\mathbf{g} = \partial s_{T+h|h}/\partial\boldsymbol{\eta}$ and $\mathbf{h} = \partial \mu_{T+h|h}/\partial\boldsymbol{\eta}$. Most often a consistent estimator $(E(\hat{\boldsymbol{\eta}}) - \boldsymbol{\eta} = \mathbf{0})$ and a large sample $(Cov(\hat{\boldsymbol{\eta}}) \approx \mathbf{0})$ are assumed in which case $s_{T+h|T}$ evaluated at estimates is the expression to employ.

3.3   A Small Monte Carlo Experiment

In this small Monte Carlo experiment we study the bias and MSE properties of the CLS, FGLS and GMM estimators for finite-lag specifications, when data is generated according to an infinite-lag INMA model. In addition, we study the serial correlation properties of estimated models by the Ljung-Box statistic as well as the properties of forecasts one and two steps ahead.

The data generating process is as in (1), with $\beta_i = \exp(\gamma_0 + \gamma_1 i), i \geq 1$, and $\beta_0 = 1$. The $\{u_t\}$ sequence is generated as Poisson with parameter λ, so that $\sigma^2 = \lambda$ is time invariant in the conditional variance (4b).[3] We set $\lambda = 5$, $\gamma_0 = -1.5, \gamma_1 = -0.1, -0.2, -0.3$ and -0.4, and $T = 1000$ and 10 000. The number of replications is 1000 in each design point. In generating the data the first 50 observations are discarded, which appears safe as β_i is effectively zero at lag 50 (the truncation point) for the used γ_0 and γ_1 combinations.

For the estimators we choose $q = 10, 20$ and 30. By the q choices we will effectively study under as well as overparameterized model version.[4] We use a simplex algorithm (the AMOEBA routine of Press et al., 1992) to minimize the criterion function of each estimator. For the GMM estimator we set **W** equal to the identity matrix. The Ljung-Box statistic is based on 10 autocorrelations, and the forecast horizon is $h = 2$.

We report bias and MSE results for β_i after accumulation over $i = 1, \ldots, 10$ for all employed q values. The full results are summarized in Table A1 of the Appendix, while Figure 2 contains the results for the CLS estimator. Starting with the CLS estimator we find that as sample size increases there is a decline in MSE throughout. For the bias there is a decline only for the most overparameterized case of $q = 30$. Biases and MSEs drop as q increases. Biases are negative for $q = 10$ and 20, with the exception of the $\gamma_1 = -0.1$ case for $q = 10$. This is the most underparameterized model. For $q = 10$ the absolute bias drops from 14.6 percent (or on average 1.46 percent for the individual parameter) for $\gamma_1 = -0.1$ to 4.4 percent and less for $\gamma_1 \leq -0.2$. The FGLS and GMM estimators also focus on the σ^2 parameter, albeit in different ways. The biases and MSEs of the GMM estimator are in most cases the poorest. The FGLS estimator has smaller biases and MSEs than the CLS estimator for $q \leq 20$ and all estimators are quite close for $q = 30$. Setting $\sigma^2 = \lambda$ and letting all other parts of the conditional variance be known improves on the performance of the

---

[3] The experiments are performed using Fortran codes. Poisson random deviates are generated by the POIDEV function (Press et al., 1992), while the binomial thinning is performed by the BNLDEV function.

[4] $\sum_{i=0}^{10} \beta_i = 2.34$ and $\beta_k < 0.01$ for $k \geq 32$ for $\gamma_1 = -0.1$, the sum is 1.87 for $k \geq 16$ and $\gamma_1 = -0.2, 1.61$ for $k \geq 11$ and $\gamma_1 = -0.3$, and 1.45 for $k \geq 8$ and $\gamma_1 = -0.4$.
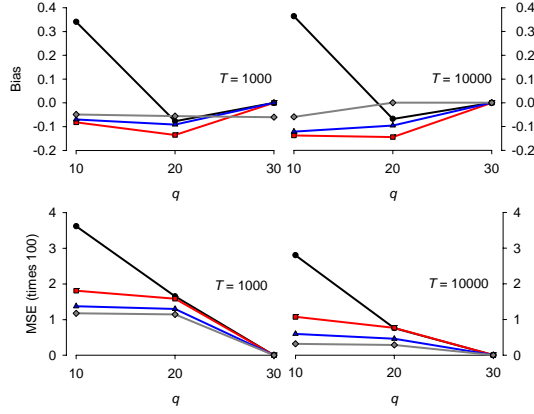
Figure 2: Bias and MSE properties in the Monte Carlo experiment of the CLS estimator of $\sum_{i=1}^{10} \beta_i$ for $T = 1000$ and 10 000 and different values of the lag length q in the adopted INMA models. The circle marker indicates that $\gamma_1 = -0.1$ is used in generating the data, square ($\gamma_1 = -0.2$), triangle ($\gamma_1 = -0.3$) and diamond ($\gamma_1 = -0.4$).

GMM estimator though not by much. Note that the use of a weight matrix different from the identity matrix may change these outcomes.

In summary, the FGLS estimator comes out as the best estimator of $\sum_{i=1}^{10} \beta_i$. However, the CLS estimator which is the simplest to use of the three considered estimators is not far behind. It is also clear that q should be chosen large.

The ability of the Ljung-Box statistic to detect remaining serial correlation was also studied by counting the number of replications exceeding a critical value of $\chi_{0.95}^2(10)$. In brief, both under and overparameterization give rise to detectable serial correlation. With respect to the forecasting performance both in terms of bias (or mean error) and MSE the FGLS estimator performs better than the other two estimators. In addition, the GMM performance is weaker than that of the CLS estimator. For $q = 30$ the differences between the performances of estimators are small.
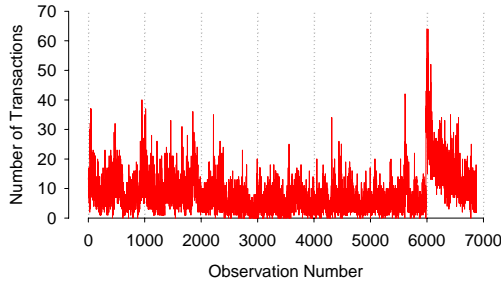
Figure 3: Time series plots for Ericsson.

4   Data and Descriptives

The time series for the number of transactions per minute in Ericsson B, in the period July 2 – 22, 2002, are displayed in Figure 3. There are frequent zero frequencies and hence the use of a count data modelling approach is called for. In producing the graph and for the analyses of this paper we have deleted all trading before 0935 (trading opens at 0930) and after 1714 (order book is closed at 1720). The reason for these deletions is that our main focus is on ordinary transactions and the first few minutes are likely to be subject to a different mechanism with considerably higher trading frequencies. The final minutes of the trading day have practically no trading. The basic data were downloaded from the Ecovision system and later filtered by the authors. Due to a technical problem in capturing transactions the first captured minute of July 19 is 0959. There are altogether 6875 observations for the Ericsson series. Descriptive statistics and a histogram for the series are given in Figure 4.

Autocorrelation functions for the series and its first difference are given in Figure 5. For the level series, the function indicates long memory. The autocorrelations after lag one of the first difference series are practically zero. The partial autocorrelation functions die out gradually for both the level and the difference series. Taken together the functions for the first difference series signal that a model for such a series should include a MA(1) component. The autocorrelations for the level series suggest that a low order AR-part is required together with a low order MA-part. With respect to the mean pattern over the day there is more trading during the first two hours than later.

When specifying and estimating INARMA models according to the conven-
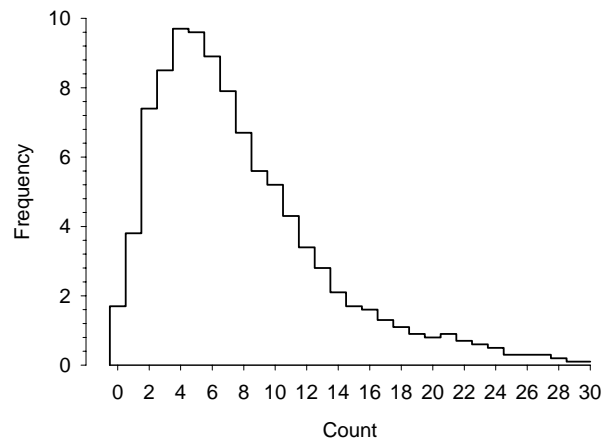
Figure 4: Histogram (in percent) for Ericsson (mean 8.13, variance 40.4, maximum 64).
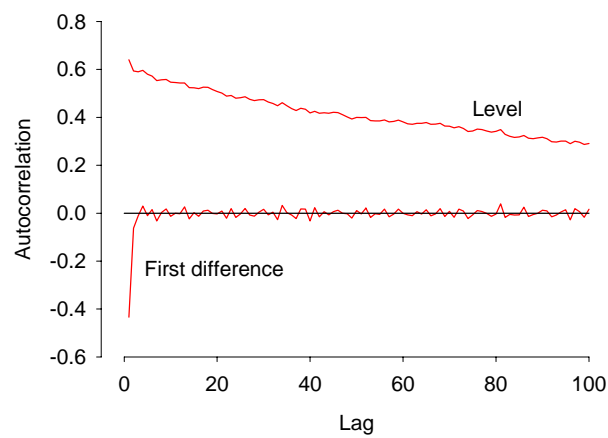
Figure 5: Autocorrelation functions from lag one for the Ericsson series and its first differences.

Table 1: ARMA estimation results (standard errors in parentheses) for the Ericsson series.

$$\hat{y}_t = \underset{(0.002)}{0.991}\,\hat{y}_{t-1} + \underset{(0.870)}{8.179} - \underset{(0.012)}{0.757}\,\hat{u}_{t-1} - \underset{(0.012)}{0.088}\,\hat{u}_{t-2}$$

$$s^2 = 18.25, LB_{20} = 28.16 \ (p = 0.11)$$

$$\nabla\hat{y}_t = \underset{(0.014)}{0.110}\,\nabla\hat{y}_{t-1} - \underset{(0.007)}{0.001} - \underset{(0.007)}{0.874}\,\hat{u}_{t-1}$$

$$s^2 = 18.32, LB_{20} = 29.79 \ (p = 0.07)$$

Note: Models estimated in SPSS.

tional Box-Jenkins methodology, in both level and first difference forms, it is obvious that, at least, the $\hat{\beta}_1$ estimates are large and have negative signs and that the INAR(1) parameter is suspiciously close to one (cf. Table 1).[5] Hence there is a violation of the probability interpretation of β_i. It also appears that pure but higher order INAR models are not successful in eliminating serial correlation. Therefore, there is empirical justification for considering INMA models with long lags even though such models are less parsimoniously parameterized.

For the purpose of evaluating the impact of explanatory variables on λ_t and σ_t^2 we also have access to price, volume and spread. These are recorded at the last transaction of the previous minute, i.e. at $t - 1$. In case there is no transaction in the previous minute(s), we use the most recent previous notations. Instead of including variables in levels we use first differences to reflect news: $\nabla p_t = p_{t-1} - p_{t-2}$ reflects the news (innovation when p_t follows a random walk) value in the price, $\nabla v_t = v_{t-1} - v_{t-2}$ in volume and $\nabla s_t = s_{t-1} - s_{t-2}$ in the spread. For volume we divide through by 10 000 000.

5   Empirical Results

The empirical results are presented in terms of finite-lag INMA(q) models. Estimation is carried out by the conditional and feasible generalized least squares (CLS(FGLS) estimators for the standard INMA(q) model as well as when a

---

[5]Note that for a count data INAR(1) model with a unit root the observed sequence of observations can not decline. Adding a MA part to the INAR(1) does not alter this feature. As is obvious from Figure 3 there are ups and downs in the present time series, so that a unit root can not logically be supported by the data.

time-varying λ_t is allowed for. GMM estimation is employed for full models including conditional heteroskedasticity and time-varying λ_t specifications.

AIC and SBIC criteria are used to find the lag length q, allowing for no gaps in the β_i sequence. For Ericsson AIC is minimized at $q = 50$, while SBIC indicates an order $q = 47$. Both criterion functions are quite flat indicating some uncertainty with respect to a true q-value.[6] CLS estimation results for pure INMA models are presented in Table 2. Though differences between estimates are quite small, the smaller q-value gives estimates that are larger for low lags and smaller for large lags. The standard errors of the estimates (based on a numerical derivative version of $\mathbf{A}^{-1}$) are throughout small and for both q-values individual hypotheses of $\beta_i = 0, i = 1, \ldots, q$, are rejected throughout. Note also that $\hat{\beta}_i$ estimates are larger than zero throughout even if an unrestricted estimator is used. The confidence intervals at the same lag overlap in most instances.

For $q = 47$, the $\hat{\beta}_i$ estimates give a mean lag of 15.8 minutes and a median lag of 14 minutes, while for $q = 50$ the mean lag is 16.0 minutes and the median 14 minutes. Hence, for the measurement of reaction time the q-choice does not matter much in this case. In additional estimations with 5 minute interval lengths, the mean lag remained unchanged but the median lag increased to 16. Therefore, we conclude that a one-unit size increase in u will have about half its effect in the first 14-16 minutes after the change.

For both models $R^2 = 0.54$, while the fit of models containing an INAR(1) parameter (cf. Table 1) is better than for pure INMA models. Note also that there are no strong correlations between estimates in this case. Table 2 suggests that the $\hat{\beta}_i$s are roughly linear in i. A linear regression gives $\tilde{\beta}_i = 0.175 - 0.0029i$ ($R^2 = 0.84$) for $q = 47$ and $\tilde{\beta}_i = 0.174 - 0.0029i$ ($R^2 = 0.85$) for $q = 50$. The goodness-of-fit improves further if $\hat{\beta}_1$ is dropped for these regressions.

There is no remaining serial correlation in the standardized residual, for either q-value, when for the conditional variance we use that of the Poisson distributed INMA(q), i.e. we set $\hat{\sigma}^2 = \hat{\lambda}$. However, the squared standardized residuals indicate remaining conditional heteroskedasticity for both models. The largest autocorrelation coefficient is 0.057 for the squared and -0.018 for the standardized residual. As a first step of attempting to find a remedy for the squared residual problem, a time invariant σ^2 is estimated in a second step of the FGLS estimator. For both q-values the $\hat{\sigma}^2$s are substantially larger than the corresponding $\hat{\lambda}$ estimates. Using $\hat{\sigma}^2$ instead of $\hat{\lambda}$ gives roughly the

---

[6]In some experimentation with an AstraZeneca series lower order model representations ($q = 18$ and 30) are found.

Table 2: Estimation results for INMA(q) models for Ericsson (s.e. times 100).

	$q = 47$						$q = 50$				
Lag	$\hat{\beta}_i$	s.e.	Lag	$\hat{\beta}_i$	s.e.	Lag	$\hat{\beta}_i$	s.e.	Lag	$\hat{\beta}_i$	s.e.
1	.2407	.28	26	.1146	.30	1	.2378	.28	26	.1102	.30
2	.1509	.29	27	.0919	.30	2	.1466	.29	27	.0899	.30
3	.1539	.29	28	.0917	.30	3	.1520	.29	28	.0891	.30
4	.1776	.29	29	.1025	.30	4	.1734	.29	29	.0991	.30
5	.1545	.30	30	.1135	.30	5	.1498	.30	30	.1114	.30
6	.1559	.30	31	.0905	.30	6	.1512	.30	31	.0869	.30
7	.1175	.30	32	.0901	.30	7	.1141	.30	32	.0863	.30
8	.1310	.30	33	.0737	.30	8	.1260	.30	33	.0696	.30
9	.1414	.30	34	.1060	.30	9	.1382	.30	34	.1031	.30
10	.1365	.30	35	.0874	.30	10	.1309	.30	35	.0855	.30
11	.1377	.30	36	.0683	.30	11	.1321	.30	36	.0676	.30
12	.1358	.30	37	.0476	.30	12	.1320	.30	37	.0494	.30
13	.1457	.30	38	.0669	.30	13	.1426	.30	38	.0665	.30
14	.1120	.30	39	.0775	.30	14	.1090	.30	39	.0780	.30
15	.1200	.30	40	.0407	.30	15	.1139	.30	40	.0380	.30
16	.1203	.30	41	.0578	.30	16	.1147	.30	41	.0545	.30
17	.1333	.30	42	.0313	.30	17	.1269	.30	42	.0293	.30
18	.1470	.30	43	.0412	.29	18	.1412	.30	43	.0395	.30
19	.1393	.30	44	.0368	.29	19	.1347	.30	44	.0369	.30
20	.1252	.30	45	.0528	.29	20	.1200	.31	45	.0540	.29
21	.1245	.30	46	.0583	.29	21	.1198	.30	46	.0633	.29
22	.0990	.30	47	.0289	.28	22	.0938	.30	47	.0324	.29
23	.1127	.30				23	.1090	.30	48	.0200	.29
24	.0936	.30				24	.0899	.30	49	.0029	.29
25	.1020	.30				25	.0974	.30	50	.0179	.28
λ	1.341	1.67					1.362	1.72			
LB_{20}	14.37						14.80				
LB_{20}^2	81.74						84.06				

Note: The Ljung-Box statistics LB_{20} and LB_{20}^2 are obtained with $\hat{\sigma}^2 = \hat{\lambda}$.

same LB_{20} statistic, while for the squared standardized residual values increase substantially. Estimating models with $\beta_i = \kappa_0 + \kappa_1 i$, where κ_0 and κ_1 are the unknown parameters, leads to more severe serial correlation.

Next, we consider the impact of explanatory variables on λ_t and on σ_t^2. Table 3 reports FGLS and CLS estimates for INMA(q) with λ_t specifications for $q = 50$. The FGLS estimates are obtained by using the CLS estimates $\hat{e}_{1t}$ and $\sum_{i=1}^q \hat{\beta}_i (1 - \hat{\beta}_i) \hat{u}_{t-i}$ to estimate a σ_t^2 model. Corresponding to the $\hat{\lambda}_t$ specification we have $\hat{\sigma}_t^2 = \exp(1.17 + 0.03 \ln \hat{\sigma}_{t-1}^2 - 3.70 \nabla p_t + 0.08 \nabla p_t^+ + 8.58 \nabla s_t)$, $R^2 = 0.10$. This suggests that a negative price change increases volatility while a positive change reduces volatility. A negative spread change lowers volatility while a widening spread increases volatility. A dummy variable for trading before 1101 AM had no significant effect even though trading is more frequent in the early hours of the day.

On comparison with Table 2, the CLS $\hat{\beta}_i$ estimates are marginally smaller for all lags. In addition, the FGLS estimates are sometimes smaller than the CLS estimates. For CLS the lag 49 estimate has a negative sign but is insignificant. It remains insignificant and small when estimated by the FGLS estimator but then the sign is correct.

In the λ_t function the lagged mean level, λ_{t-1}, has a rather small but significant effect when estimated by the FGLS estimator, while it is insignificant in the CLS estimated model. In terms of the CLS estimates there are significant asymmetric but not very different effects for the price change variable; with a tick size of 0.1 SEK we expect an enhancing average effect of 0.63 for a positive and 0.73 for a negative one tick change. The asymmetry is insignificant for FGLS but the corresponding estimated effects are larger and equal to -0.81 and 0.39, respectively. News about spread increase the frequencies. These are expected signs when compared to duration models for the same underlying data (Brännäs and Simonsen, 2003). The current effects are more significant as for the duration data only the news about prices came out with a significant effect. To account, at least, partly for seasonal within days effects we included a dummy variable $\mathbf{1}(t \leq 1100)$ which takes value one for transactions before 1101 AM and zero otherwise. The estimated effect is positive and significant for both estimators.

There is no practical change in the serial correlation properties for the CLS estimated model with $\hat{\sigma}_t^2 = \hat{\lambda}_t$, though the Ljung-Box statistic for the squared residuals is much smaller and not very far from a p-value of 0.05. In this case the largest autocorrelations are -0.038 and 0.032 (squared standardized residual) for $q = 50$. The effect of news in volume also come out significantly but implies substantial serial correlation in both $\hat{e}_{1t}/\hat{V}^{1/2}(y_t | Y_{t-1})$ and its square. When

Table 3: FGLS and CLS estimation results for INMA(50) with exponential λ_t models for Ericsson (s.e. times 100). For the σ_t^2 specification, see the text.

	FGLS						CLS				
Lag	$\hat{\beta}_i$	s.e.	Lag	$\hat{\beta}_i$	s.e.	Lag	$\hat{\beta}_i$	s.e.	Lag	$\hat{\beta}_i$	s.e.
1	.2257	.80	26	.1035	.84	1	.2312	.28	26	.1069	.30
2	.1481	.82	27	.0839	.85	2	.1412	.29	27	.0854	.30
3	.1466	.82	28	.0868	.85	3	.1511	.29	28	.0853	.30
4	.1641	.83	29	.0941	.85	4	.1701	.29	29	.0943	.30
5	.1502	.83	30	.1110	.85	5	.1463	.30	30	.1078	.30
6	.1468	.83	31	.0914	.85	6	.1462	.30	31	.0845	.30
7	.1124	.84	32	.0844	.85	7	.1116	.30	32	.0826	.30
8	.1207	.84	33	.0631	.85	8	.1215	.30	33	.0666	.30
9	.1301	.84	34	.0993	.85	9	.1295	.30	34	.1004	.30
10	.1301	.84	35	.0855	.85	10	.1255	.30	35	.0835	.30
11	.1296	.84	36	.0594	.85	11	.1273	.30	36	.0625	.30
12	.1251	.85	37	.0417	.84	12	.1283	.30	37	.0436	.30
13	.1435	.85	38	.0607	.84	13	.1401	.30	38	.0606	.30
14	.1034	.85	39	.0689	.84	14	.1012	.30	39	.0718	.30
15	.1159	.85	40	.0302	.84	15	.1126	.30	40	.0333	.30
16	.1096	.86	41	.0500	.84	16	.1080	.30	41	.0540	.30
17	.1254	.85	42	.0316	.84	17	.1252	.30	42	.0281	.30
18	.1402	.85	43	.0390	.84	18	.1374	.30	43	.0381	.30
19	.1317	.86	44	.0331	.84	19	.1292	.30	44	.0328	.30
20	.1160	.86	45	.0548	.84	20	.1173	.31	45	.0500	.29
21	.1193	.85	46	.0616	.82	21	.1172	.30	46	.0599	.29
22	.0910	.85	47	.0308	.82	22	.0930	.30	47	.0301	.29
23	.0991	.85	48	.0172	.81	23	.1057	.30	48	.0174	.29
24	.0847	.85	49	.0051	.81	24	.0883	.30	49	-.0001	.29
25	.0927	.85	50	.0235	.80	25	.0973	.30	50	.0158	.28
θ_0	.2483	31.54					-.5802	4.01			
λ_{t-1}^*	.3281	3.07					-.0151	1.33			
∇p_t	-3.892	142.6					-7.276	51.59			
∇p_t^+	-4.154	314.4					13.60	18.79			
∇s_t	11.128	54.85					4.987	35.40			
1_t	.2420	2.86					.3613	1.30			
LB_{20}	43.95						20.66				
LB_{20}^2	12.60						33.28				

Note: $\lambda_t = \exp(\theta_0 + \theta_1 \lambda_{t-1}^* + \theta_2 \nabla p_t + \theta_3 \nabla p_t^+ + \theta_4 \nabla s_t + \theta_5 1_t)$, where $\lambda_{t-1}^* = \ln \lambda_{t-1}$, $\nabla p_t^+ = 0$ for $\nabla p_t \leq 0$ and $\nabla p_t^+ = \nabla p_t$ for positive news, and $1_t = 1(t \leq 1100)$.

λ_t is changed to have a linear form there are no serial correlation problems. Unfortunately, $\hat{\lambda}_t$ is then negative for some 30 percent of the observations. Obviously, this is a logically unappealing feature. For the FGLS estimated model we note that conditional heteroskedasticity no longer appears a problem, while the standardized residual now signals trouble. The largest autocorrelation coefficients are 0.042 and 0.021 (squared standardized residual).

Full models including σ_t^2 of the exponential type in (6) have been estimated by GMM. In each instance there is no serial correlation in the standardized residual, but there is serial correlation in the squared standardized residuals of about the same magnitude as the CLS of Table 3.

6 Concluding Remarks

The suggested integer-valued moving average model has relatively straightforward moment properties and estimating the unknown parameters by well-known techniques is relatively simple. In addition, both the conditional least squares and feasible least squares estimators are readily available in many standard statistical packages and have good statistical properties.

The current paper focused on modelling a time series of the intra-day number of transactions per time unit using the integer-valued moving average model class. In its practical implementation for the time series of the number of transactions in Ericsson B, we found both promising and less advantageous features of the model. With the CLS estimator it was relatively easy to model the conditional mean in a satisfactory way in terms of both interpretation and residual properties. It was more difficult to obtain satisfactory squared residual properties for the conditional variance specifications that were tried. The FGLS estimator reversed this picture and we suggest that more empirical research is needed on the interplay between the conditional mean and heteroskedasticity specifications for count data. Depending on research interest the conditional variance parameters are or are not of particular interest. For studying reaction times to shocks or news it is the conditional mean that matters, in much the same way as for conditional duration models. In addition, the conditional variance has no direct ties to, e.g., risk measures included in, e.g., option values or portfolios.

Table A1: Bias and MSE properties in the Monte Carlo experiment of estimators of $\sum_{i=1}^{10} \beta_i$ when the model is estimated as INMA(q), for $q = 10, 20$ and 30. The results are based on 1000 replications, and the bias is multiplied by 100 and the MSE by 100 000.

| | $q = 10$ | | | | $q = 20$ | | | | $q = 30$ | | | |
| | 1000 | | 10000 | | 1000 | | 10000 | | 1000 | | 10000 | |
γ_1	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
						CLS						
-0.1	34.073	3615.148	36.390	2800.233	-7.714	1650.969	-6.778	757.487	0.011	0.016	0.007	0.008
-0.2	-8.169	1807.777	-13.723	1073.082	-13.525	1586.333	-14.437	766.568	0.011	0.017	0.009	0.011
-0.3	-6.948	1374.710	-12.111	597.983	-9.124	1300.164	-9.569	460.360	0.016	0.019	0.015	0.015
-0.4	-4.867	1175.993	-5.883	319.916	-5.562	1142.259	-6.033	286.904	0.020	0.021	0.019	0.018
						FGLS						
-0.1	35.365	3816.869	37.921	3059.045	-6.767	1646.517	-5.520	743.225	0.014	0.016	0.007	0.008
-0.2	-7.803	1792.105	-11.186	986.607	-13.399	1585.772	-14.098	755.514	0.011	0.017	0.012	0.012
-0.3	-7.529	1393.705	-12.157	597.938	-8.926	1294.023	-9.370	456.195	0.013	0.020	0.017	0.017
-0.4	-4.387	1163.534	-6.404	327.260	-5.583	1137.357	-5.917	285.665	0.021	0.022	0.020	0.019
						GMM						
-0.1	21.976	2714.975	21.707	2041.617	-15.382	1770.271	-17.114	1046.574	0.002	0.017	-0.004	0.010
-0.2	-12.272	1829.799	-19.636	1255.089	-14.906	1667.040	-17.836	934.515	0.011	0.018	0.011	0.015
-0.3	-8.401	1407.453	-12.394	653.398	-9.081	1308.663	-10.594	524.195	0.016	0.020	0.014	0.016
-0.4	-5.232	1216.969	-7.897	400.151	-6.001	1142.354	-6.721	329.992	0.020	0.023	0.017	0.017

References

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control* **19**, 716-723.

Al-Osh M. and Alzaid A. (1988). Integer-Valued Moving Average (INMA) Process. *Statistical Papers* **29**, 281-300.

Al-Osh M. and Alzaid A. (1991). Binomial Autoregressive Moving Average Models. *Communications in Statistics: Stochastic Models* **7**, 261-282.

Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Brännäs, K. (1995). Explanatory Variables in the AR(1) Count Data Model. *Umeå Economic Studies* **381**.

Brännäs, K. and Brännäs, E. (2004). Conditional Variance in Count Data Regression. *Communications in Statistics: Theory and Methods* **33**, 2745-2758.

Brännäs, K. and Hall, A. (2001). Estimation in Integer-Valued Moving Average Models. *Applied Stochastic Models in Business and Industry* **17**, 277-291.

Brännäs, K. and Hellström, J. (2001). Generalizations to the Integer-Valued AR(1) Model. *Econometric Reviews* **20**, 425-443.

Brännäs, K., Hellström, J. and Nordström, J. (2002). A New Approach to Modelling and Forecasting Monthly Guest Nights in Hotels. *International Journal of Forecasting* **18**, 19-30.

Brännäs, K. and Simonsen, O. (2003). Discretized Time and Conditional Duration Modelling for Stock Transaction Data. *Umeå Economic Studies* **610**.

Chiang, M-H. and Wang, C-H. (2004). Intradaily Relationship between Information Revelation and Trading Duration under Market Trends: The Evidence of MSCI Taiwan Stock Index Futures. *Applied Economics Letters*, **11**, 495-501.

Davis, R.A. and Dunsmuir, W.T.M. (2000). On Autocorrelation in a Poisson Regression Model. *Biometrika* **87**, 491-505.

Durbin, J. (1959). Efficient Estimation of Parameters in Moving-Average Models. *Biometrika* **46**, 306-316.

Engle, R.F. and Russell, J.R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* **66**, 1127-1162.

Greene, W.H. (2003). *Econometric Analysis*. 5th edition, Upper Saddle River: Prentice-Hall.

Hansen, L.P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* **50**, 1029-1054.

Harris, D. (1999). GMM Estimation of Time Series Models. In Mátyás, L. (ed.) *Generalized Method of Moments Estimation.* Cambridge: Cambridge University Press.

Heinen, A. and Rengifo, E. (2003). Multivariate Modelling of Time Series Count Data: An Autoregressive Conditional Poisson Model. CORE Discussion Paper 2003/25, Université Catholique de Louvain.

Joe, H. (1996). Time Series Models with Univariate Margins in the Convolution-Closed Infinitely Divisible Class. *Journal of Applied Probability* **33**, 664-677.

Jørgensen, B. and Song, P.X-K. (1998). Stationary Time Series Models with Exponential Dispersion Model Margins. *Journal of Applied Probability* **35**, 78-92.

Lobato, I., Nankervis, J.C. and Savin, N.E. (2001). Testing for Autocorrelation using a Modified Box-Pierce Q Test. *International Economic Review* **42**, 187-205.

McKenzie E. (1986). Autoregressive-Moving-Average Processes with Negative Binomial and Geometric Marginal Distribution. *Advances in Applied Probability* **18**, 679-705.

McKenzie, E. (1988). Some ARMA models for Dependent Sequences of Poisson Counts. *Advances in Applied Probability* **20**, 822-835.

McKenzie, E. (2003). Discrete Variate Time Series. In Shanbhag, D.N. and Rao, C.R. (eds.) *Handbook of Statistics*, Volume 21, pp. 573-606. Amsterdam: Elsevier Sciences.

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* **59**, 347-370.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1992). *Numerical Recipes.* 2nd ed., Cambridge: Cambridge University Press.

Bivariate Time Series Modelling of Financial Count Data*

A.M.M. 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.

Mathematics Subject Classification 62M10; 91B84; 91B24; 91B28.

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.[1] 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

---

[1]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.

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 can be used to measure the reaction time for, e.g., macro-economic news or rumours and how new information spreads through the system. 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. Extensions of the BINMA model are also discussed 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 the possible inclusion 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 intra-day 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 impact 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

$$
\begin{aligned}
y_{1t} &= u_{1t} + \alpha_{11} \circ u_{1t-1} + \ldots + \alpha_{1q_1} \circ u_{1t-q_1} & \text{(1a)} \\
y_{2t} &= u_{2t} + \alpha_{21} \circ u_{2t-1} + \ldots + \alpha_{2q_2} \circ u_{2t-q_2}. & \text{(1b)}
\end{aligned}
$$

The macro-economic news are assumed to be captured by $\{u_{jt}\}$, $j = 1, 2$ and filtered by $\{\alpha_{jt}\}$ through the system. 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 $\{u_{jt}\}$ an iid sequence with mean λ_j and variance $\sigma_j^2 = \upsilon_j \lambda_j$, the unconditional first and second order moments can be given as follows:

$$
\begin{aligned}
E(y_{jt}) &= \lambda_j(1 + \sum_{i=1}^{q_j} \alpha_{ji}) & \text{(3a)} \\
V(y_{jt}) &= \lambda_j[(\upsilon_j + \sum_{i=1}^{q_j} \alpha_{ji}) + (\upsilon_j - 1)\sum_{i=1}^{q_j} \alpha_{ji}^2] & \text{(3b)} \\
\gamma_{jk} &= \sigma_j^2(\alpha_{jk} + \sum_{i=1}^{q_j} \alpha_{ji}\alpha_{jk+i}), \quad k \geq 1 & \text{(3c)}
\end{aligned}
$$

where γ_{jk} denotes the autocovariance function at lag k and $\upsilon_j > 0$. It is clear from (3) that the mean, variance and autocovariance take only positive values since λ_j, σ_j^2 and α_{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 $\upsilon_j > 1$, $\upsilon_j \in (0, 1)$ or $\upsilon_j = 1$, 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 $[t-1, t)$ 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 BINMA(q_1, q_2) can be given as follows:

$$
\gamma_k = \begin{cases} \Lambda\left(\alpha_{1k} + \sum_{i=1}^{q_j-k} \alpha_{1k+i}\alpha_{2i}\right), & 0 \le k \le \min(q_1, q_2) \\ 0, & k > \min(q_1, q_2) > 0 \end{cases} \quad (4a)
$$

$$
\gamma_{-k} = \begin{cases} \Lambda\left(\alpha_{2k} + \sum_{i=1}^{q_j-k} \alpha_{2k+i}\alpha_{1i}\right), & 0 \le k \le \min(q_1, q_2) \\ 0, & k > \min(q_1, q_2) > 0 \end{cases} \quad (4b)
$$

where $\gamma_k = Cov(y_{1t}, y_{2t-k})$, $\gamma_{-k} = Cov(y_{1t-k}, y_{2t})$, and $Cov(u_{1t}, u_{2t}) = \Lambda = \varphi - \lambda_1\lambda_2$ where $\varphi = E(u_{1t}u_{2t})$. Note that there is no cross-lag dependence among u_{jt} and the covariances $Cov(u_{1t}, u_{2t})$ are assumed constant over time. Note also that the sign of the covariance function in (4a−b) depends on the relative sizes of φ 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

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

$$
V(y_{jt}|Y_{jt-1}) = V_{j|t-1} = v_j\lambda_j + \sum_{i=1}^{q_j} \alpha_{ji}(1 - \alpha_{ji})u_{jt-i}. \qquad (5b)
$$

Note that the mean and variance are conditioned on only the previous observations, $Y_{jt-1} = y_{jt-1}, y_{jt-2}....$ Since the conditional variance varies with $u_{jt-i}, i \ge 1$, 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 $v_j > 1 + \sum_{i=1}^{q_j} \alpha_{ji}^2/\lambda_j$, $v_j \in (0, 1 + \sum_{i=1}^{q_j} \alpha_{ji}^2/\lambda_j)$ or $v_j = 1 + \sum_{i=1}^{q_j} \alpha_{ji}^2/\lambda_j$, respectively. The conditional covariance function for

BINMA(q_1, q_2) can be written

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

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. $v_j = 1$, 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

$$
\rho_0 = \frac{(1 + \sum_1^{\min(q_1, q_2)} \alpha_{1i}\alpha_{2i})\Lambda}{V^{1/2}(y_{1t})V^{1/2}(y_{2t})}. \quad (7)
$$

Note that this correlation can take a positive or a negative sign depending on the size of φ 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 ρ_0 that

$$
\mid \varphi - \lambda_1\lambda_2 \mid \leq \lambda_1^{1/2}\lambda_2^{1/2}.
$$

The correlation $\mid \rho_0 \mid = 1$ if and only if $\mid \varphi - \lambda_1\lambda_2 \mid = \lambda_1^{1/2}\lambda_2^{1/2}$ 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 $\mid \rho_0 \mid < 1$ (see appendix for proof). This also holds for the over- and underdispersion cases. As v_j deviates from 1, $\mid \rho_0 \mid$ decreases. It can be shown that this result gets support from the coherence function for the BINMA(q_1, q_2).[2]

2.2 Extensions of the BINMA Model

The Multivariate INMA model follows directly from BINMA(q_1, q_2) and can be written

$$
\mathbf{y}_t = \mathbf{u}_t + \mathbf{A}_1 \circ \mathbf{u}_{t-1} + \ldots + \mathbf{A}_q \circ \mathbf{u}_{t-q} \quad (8)
$$

---

[2] A detailed description of coherence function can be found, e.g., in Brockwell and Davis (1991, ch. 4, 10 and 11).

where the $\mathbf{A}_i, i = 1, \ldots, q$, are diagonal matrices and $\mathbf{u}_t \sim (\boldsymbol{\lambda}, \boldsymbol{\Sigma})$. Covariation between $\mathbf{y}_t$ elements can also arise, even if $\boldsymbol{\Sigma}$ is a diagonal matrix but then, at least, one of the $\mathbf{A}_i$ 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 λ as a function of explanatory variables. This can be specified as

$$\lambda_{jt} = \exp(\mathbf{x}_{jt}\boldsymbol{\theta}_j) \geq 0 \tag{9}$$

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

$$\begin{aligned}
\sigma_{jt}^2 &= \exp[\phi_0 + \phi_1 \ln \sigma_{jt-1}^2 + \ldots + \phi_{P_j} \ln \sigma_{t-1}^2 \\
&\quad + \omega_1(u_{jt-1} - \lambda_j)^2 + \ldots + \omega_{Q_j}(u_{jt-1} - \lambda_j)^2 + \mathbf{x}_{jt}\boldsymbol{\xi}_j].
\end{aligned} \tag{10}$$

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

$$e_{j1t} = y_{jt} - E_{jt|t-1}, \quad j = 1, 2. \tag{11}$$

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 $E(e_{j1t}) = 0$ and $E(e_{j1t}e_{j1t-i}) = 0$, $i \geq 1$ could be used. The CLS estimator minimizes the criterion function $S_{CLS} = \sum_{t=r}^{T} e_{j1t}^2$, where $r = q_j + 1$ and T is the length of the time series, with respect to the unknown parameter vector $\boldsymbol{\psi}_j' = (\lambda_j, \boldsymbol{\alpha}_j')$ or $\boldsymbol{\psi}_j' = (\boldsymbol{\theta}_j', \boldsymbol{\alpha}_j')$ when a time-varying λ_{jt} is employed. To calculate the sequence e_{j1t} we write the prediction error on the form

$$e_{j1t} = u_{jt} - \lambda_{jt} + \sum_{i=1}^{q_j}(\alpha_{ji} \circ u_{jt-i} - \alpha_{ji}u_{jt-i}). \tag{12}$$

Instead of using, say, some randomization device to evaluate the sum we advocate using its expected value zero and so employ $e_{j1t} = u_{jt} - \lambda_{jt}$.

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

$$e_{j2t} = (y_{jt} - E_{jt|t-1})^2 - V_{j|t-1} \tag{13}$$

$$e_{3t} = e_{11t}e_{21t} - \gamma_{0|t-1} \tag{14}$$

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

For FGLS, $S_{j2} = \sum_{t=s_j}^T e_{j2t}^2$ and $S_3 = \sum_{t=s_j}^T e_{3t}^2$, where $s_j = \max(q_j, P_j, Q_j)$ $+1$, are minimized with respect to the respective parameters of the function σ_{jt}^2 and with the CLS estimates $\widehat{\psi}_j$ and $\widehat{u}_{jt}$ kept fixed. Simpler and obvious moment estimators for time invariant $\sigma_j^2 = \upsilon_j \lambda_j$ and φ can be written on the following forms

$$\hat{\sigma}_j^2 = (T-r)^{-1} \sum_{t=s_j}^T \left[e_{j1t}^2 - \sum_{i=1}^{q_j} \widehat{\alpha}_{ji}(1 - \widehat{\alpha}_{ji})\widehat{u}_{jt-i} \right]$$

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

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

$$S_{FGLS} = \sum_{t=s_j}^T (\widehat{V}_{2|t-1}e_{11t}^2 + \widehat{V}_{1|t-1}e_{21t}^2 - 2\widehat{\gamma}_{0|t-1}e_{11t}e_{21t})/\widehat{D}_t \tag{15}$$

is minimized with respect to ψ_j. In (15) $\widehat{V}_{j|t-1}$, $\widehat{\gamma}_{0|t-1}$ and $\widehat{D}_t = \widehat{V}_{1|t-1}\widehat{V}_{2|t-1} - \widehat{\gamma}_{0|t-1}^2$ 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

$$Cov(\widehat{\psi}_{FGLS}) = \left(\sum_{t=s+1}^T \frac{\partial e_t}{\partial \psi'} \widehat{\Omega}^{-1} \frac{\partial e_t}{\partial \psi} \right)^{-1}$$

where $e_t = (e_{11t}, e_{21t})'$ and $\widehat{\Omega}$ 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 (Brännäs, 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 $\boldsymbol{\eta}' = (\boldsymbol{\psi}', \boldsymbol{\omega}')$ where $\boldsymbol{\omega}' = (\sigma_j^2, \phi_0, \ldots, +\phi_{P_j}, \omega_1, \ldots, \omega_{Q_j}, \boldsymbol{\xi}_j)$. 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 $\mathbf{m}_{1t} = e_{j1t}\partial e_{j1t}/\partial \psi_k$ 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 $\mathbf{W}$

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

If we set $\widehat{\mathbf{W}} = \mathbf{I}$, the covariance matrix of the parameter estimator is esti-
mated by $Cov(\boldsymbol{\eta}) = (T-n)^{-1}(\hat{\mathbf{G}}'\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}'\hat{\boldsymbol{\Gamma}}\hat{\mathbf{G}}(\hat{\mathbf{G}}'\hat{\mathbf{G}})^{-1}$, where $\hat{\mathbf{G}} = \partial\hat{\mathbf{m}}/\partial\boldsymbol{\eta}'$.
The covariance matrix of the parameter estimator becomes $Cov(\boldsymbol{\eta}) = (T-n)^{-1}\hat{\mathbf{G}}^{-1}\hat{\boldsymbol{\Gamma}}(\hat{\mathbf{G}}')^{-1}$ 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 fre-
quently 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 Novem-
ber 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 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 IN-ARMA 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 auto-correlation 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 $\sigma_k^2 = V(y_t - y_{t-k})$ and $\sigma_k^2 \sim O(k^{2d-1})$ for an I(d) process, a test for the presence of I(d) can be conducted (Diebold, 1989). The growth in $R(k)$ is
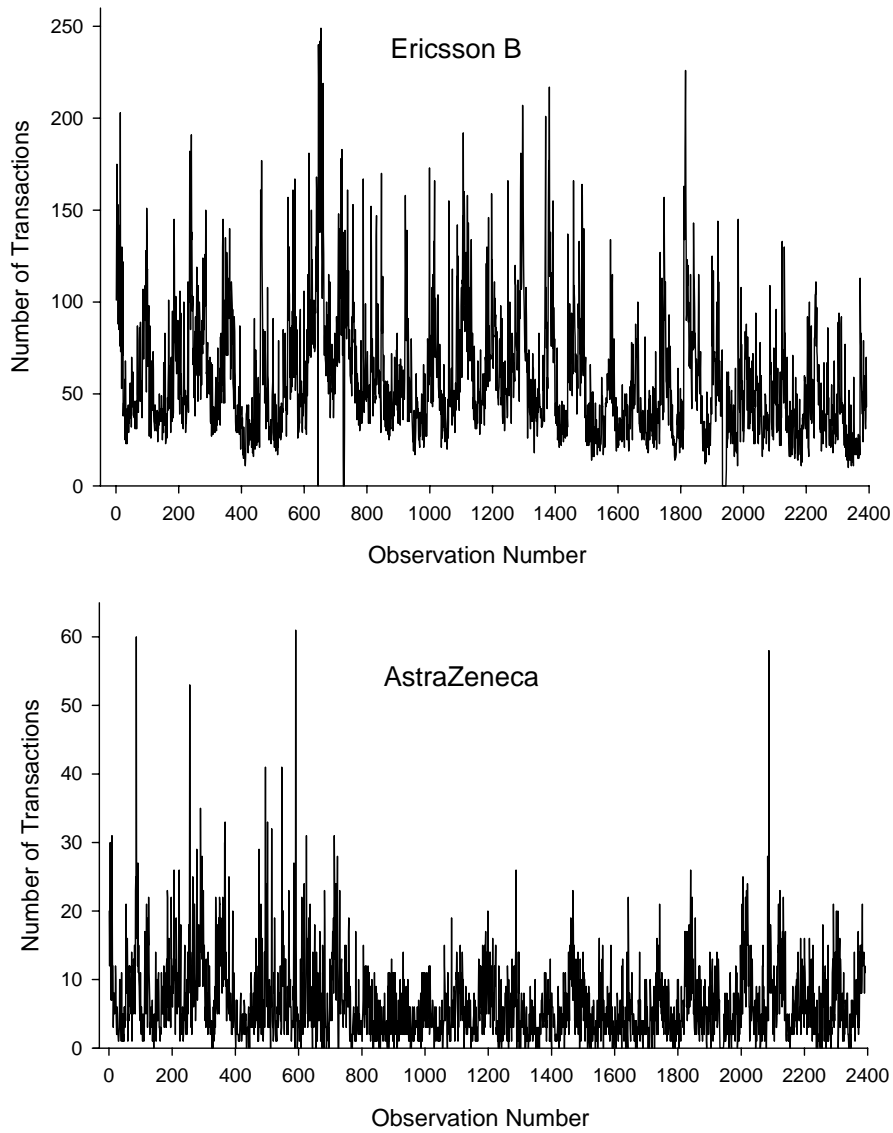
Figure 1: Time series plots for Ericsson B (mean 58.64, variance 1193.70, maximum 249) and AstraZeneca (mean 6.64, variance 34.38, maximum 64) for the period November 5-December 12, 2002.
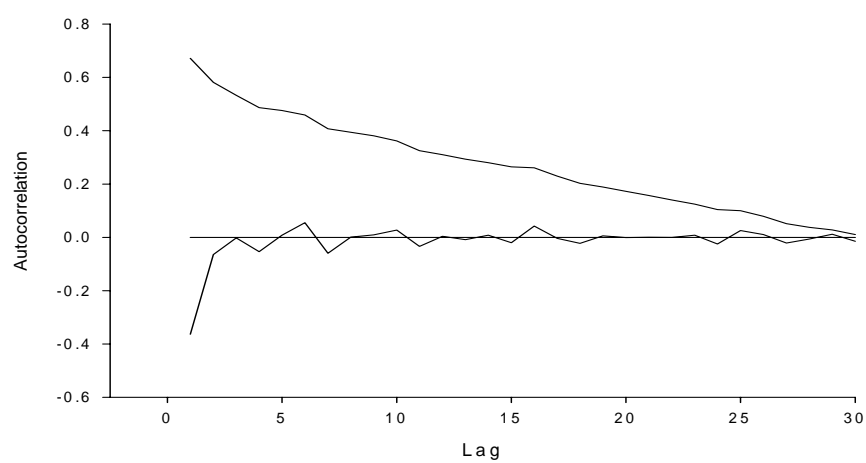
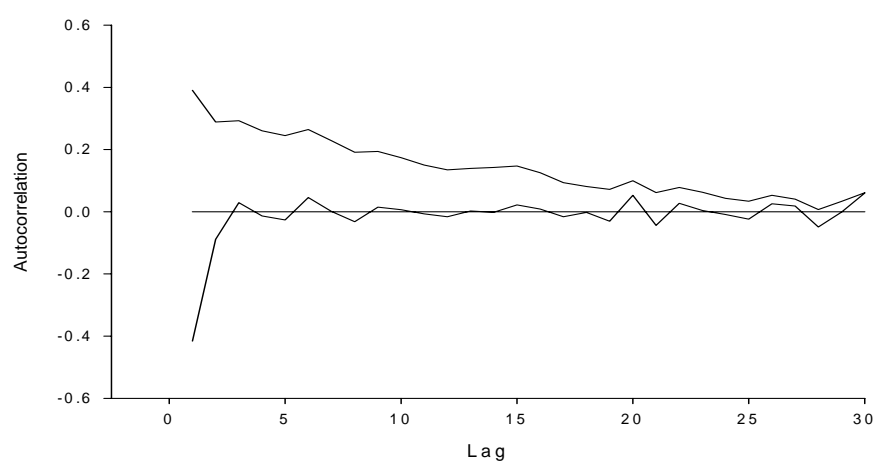Figure 2: Autocorrelation functions from lag one for the Ericsson B series and its first difference.

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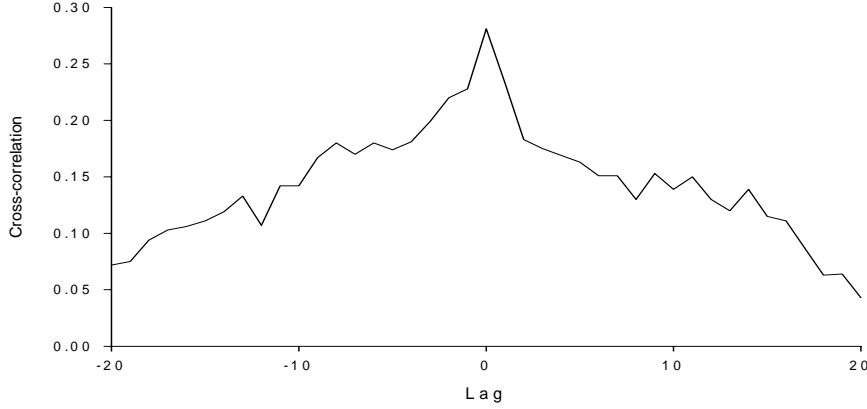
Figure 4: Cross-correlation function between Ericsson B and AstraZeneca from lag 0 to ± 20.

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 $\lim_{n\to\infty} \sum_{-n}^{n} \mid \rho_k \mid = \infty$, where ρ_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 seems to have 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.
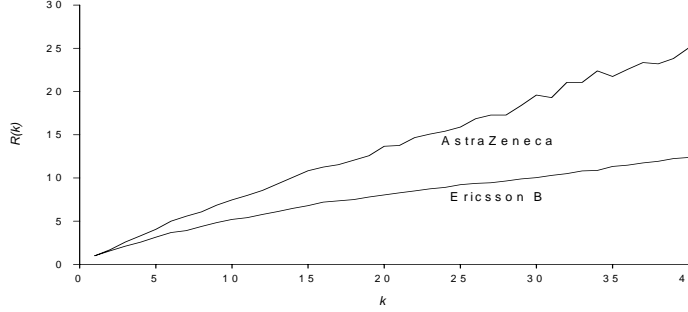
Figure 5: $R(k)$ for the AstraZeneca series indicates a long memory process, while $R(k)$ for the Ericsson B series does not.

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, while the standardized residuals based on FGLS are serially uncorrelated by the Ljung-Box test statistic. 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 remaining cross-correlation in 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 $\widehat{\alpha}_{1i}$s decrease, with some exceptions, all the way as the lag increases, while $\widehat{\alpha}_{2i}$s decrease in the same way until lag 12. After lag 12, the $\widehat{\alpha}_{2i}$s fluctuate around 0.06. This downward trend implies that the impact of macro-economic and other common news on both stocks are similar i.e., the intensity of trading for both stocks increases as the news breaks out and fades away with time. We can also say that the probability for increase in intensity of trading due to macro-economic news is higher on both stocks as the news breaks out and decreases gradually as time elapses. For FGLS and $\widehat{\alpha}_{1i}$, a linear regression gives $\widetilde{\alpha}_{1i} = 0.358 - 0.019i$ ($R^2 = 0.87$). We get a better result in terms of goodness-of-fit when $\widetilde{\alpha}_{11}$ 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 $\widehat{\rho}_{0|t-1}$ 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 $\widehat{\alpha}_{1i}$ estimates give a mean lag of 5.18 (5.52) and a median lag of 4 (5), while the $\widehat{\alpha}_{2i}$ estimates give a mean lag of 4.93 (3.78) and a median lag of 3 (3).[3] Since the used data are aggregates over five minutes intervals, the mean lags must be multiplied by 5 to express them in terms of minute. Hence, for FGLS, the $\widehat{\alpha}_{1i}$ 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.[4] 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

---

[3]As measures of reaction times to macroeconomic news/rumors in the $\{u_{jt}\}$ sequence we use the mean lag $\sum_{i=0}^{q_j} i\alpha_{ji}/w$, where $w = \sum_{i=0}^{q_j} \alpha_{ji}$ and where $\alpha_{j0} = 1$. Alternatively, we use the median lag, which is the smallest k such that $\sum_{i=0}^{k} \alpha_{ji}/w \geq 0.5$.

[4]For aggregated data into 10 minute intervals of time we also estimate a BINMA$(13, 11)$ model employing FGLS. The estimates give a mean reaction time of $28, 38$ minutes and a median reaction time of 21.66 minutes for Ericsson B The corresponding mean and median for AstraZeneca are 16.66 and 12 minutes, respectively.

Table 1: Results for BINMA models for Ericsson B and AstraZeneca estimated by FGLS and GMM.

| | Ericsson B | | | | | AstraZeneca | | | |
| | FGLS | | GMM | | | FGLS | | GMM | |
Lag	$\widehat{\alpha}_{1i}$	s.e.	$\widehat{\alpha}_{1i}$	s.e.	Lag	$\widehat{\alpha}_{2i}$	s.e.	$\widehat{\alpha}_{2i}$	s.e.
1	0.453	0.029	0.576	0.032	1	0.253	0.025	0.218	0.053
2	0.335	0.025	0.501	0.033	2	0.128	0.021	0.142	0.030
3	0.283	0.024	0.446	0.029	3	0.158	0.022	0.376	0.064
4	0.231	0.022	0.382	0.029	4	0.137	0.021	0.108	0.033
5	0.234	0.022	0.436	0.028	5	0.119	0.021	0.096	0.027
6	0.266	0.023	0.455	0.033	6	0.158	0.022	0.125	0.036
7	0.192	0.021	0.370	0.027	7	0.124	0.021	0.098	0.026
8	0.180	0.021	0.356	0.028	8	0.081	0.020	0.087	0.025
9	0.189	0.021	0.269	0.024	9	0.105	0.020	0.097	0.026
10	0.188	0.021	0.306	0.024	10	0.085	0.020	0.104	0.024
11	0.118	0.020	0.196	0.022	11	0.065	0.020	0.072	0.021
12	0.123	0.020	0.179	0.020	12	0.046	0.020	0.044	0.028
13	0.103	0.020	0.119	0.021	13	0.066	0.020	0.045	0.023
14	0.092	0.020	0.132	0.022	14	0.061	0.020	0.053	0.023
15	0.084	0.020	0.118	0.025	15	0.097	0.020	0.076	0.022
16	0.098	0.020	0.150	0.024	16	0.072	0.020		
17	0.070	0.020	0.081	0.022	20	0.054	0.020		
18	0.040	0.019			22	0.047	0.020		
$\widehat{\lambda}_1$	13.54	0.320	9.95	3.18	$\widehat{\lambda}_2$	2.31	0.032	2.867	0.91
$\widehat{\sigma}_1^2$	550.96		746.53	38.71	$\widehat{\sigma}_2^2$	22.75		10.57	2.81
$\widehat{\varphi}$	50.17		98.79	28.56	$\widehat{\rho}_{0|t-1}$	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	

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 this paper are aggregated into five minute intervals of time while the corresponding interval is one minute in Brännäs and Quoreshi (2004). Third, 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 λ_j or σ_j^2 be time-varying. In a univariate setting, by letting λ 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.

Appendix

Proposition: If $\alpha_{ji} \in (0, 1)$, then

$$| \rho_0 | < 1.$$

Proof: By applying the Cauchy-Schwarz inequality we can show for ρ_0 that

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

Hence it will be sufficient to proof the proposition, if we can show that

$$(1 + \sum_1^{\min(q_1,q_2)} \alpha_{1i}\alpha_{2i}) < (1 + \sum_1^{q_1} \alpha_{1i})^{1/2}(1 + \sum_1^{q_2} \alpha_{2i})^{1/2}.$$

Assume that a, b, c and d are all positive integers. Assume further that $b > a$, $d > c$ and $a/b > c/d$. Then we can write the following inequalities

$$\frac{c}{d}a \quad < \quad \frac{c}{d}b \tag{1}$$

$$\Rightarrow \frac{ac}{bd} \quad < \quad \frac{c}{d} < \frac{a}{b} \tag{2}$$

$$\Rightarrow 1 + \frac{ac}{bd} \quad < \quad 1 + \frac{c}{d} < 1 + \frac{a}{b} \tag{3}$$

$$\Rightarrow (1 + \frac{ac}{bd})^2 \quad < \quad (1 + \frac{c}{d})(1 + \frac{c}{d}) \tag{4}$$

$$\Rightarrow (1 + \frac{ac}{bd})^2 \quad < \quad (1 + \frac{c}{d})(1 + \frac{a}{b}) \tag{5}$$

$$\Rightarrow (1 + \frac{ac}{bd}) \quad < \quad (1 + \frac{c}{d})^{1/2}(1 + \frac{a}{b})^{1/2}. \tag{6}$$

Since a/b, $c/d \in (0, 1)$, we can replace a/b and c/d by α_1 and α_2. By relaxing the assumption that $a/b > c/d$, we can write

$$\alpha_1\alpha_2 \quad < \quad \min(\alpha_1, \alpha_2) \tag{7}$$

$$\Rightarrow (1 + \alpha_1\alpha_2) \quad < \quad (1 + \alpha_1)^{1/2}(1 + \alpha_2)^{1/2}. \tag{8}$$

Using $(2), (3), (6), (8)$ and (7) and by introducing index for α, we can generalize the idea and show that

$$\alpha_{1i}\alpha_{2i} \quad < \quad \min(\alpha_{1i}, \alpha_{2i})$$

$$\Rightarrow (1 + \sum_1^{\min(q_1,q_2)} \alpha_{1i}\alpha_{2i}) \quad < \quad (1 + \sum_1^{\min(q_1,q_2)} \alpha_{1i})^{1/2}(1 + \sum_1^{\min(q_1,q_2)} \alpha_{2i})^{1/2}$$

$$\Rightarrow (1 + \sum_1^{\min(q_1,q_2)} \alpha_{1i}\alpha_{2i}) \quad < \quad (1 + \sum_1^{q_1} \alpha_{1i})^{1/2}(1 + \sum_1^{q_2} \alpha_{2i})^{1/2}$$

References

Al-Osh, M.A. and Alzaid, A.A. (1987). First Order Integer-Valued Autoregressive INAR(1)) Process. *Journal of Time Series Analysis* **8**, 261-275.

Blundell, R., Griffith, R. and Windmeijer, F. (2002). Individual Effects and Dynamics in Count Data Models. *Journal of Econometrics* **108**, 113-131.

Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis: Forecasting and Control.* San-Francisco: Holden-Day.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Method.* New York: Springer-Verlag.

Brännäs, K. (1995). Explanatory Variables in the AR(1) Model. *Umeå Economic Studies* **381**.

Brännäs, K. and Brännäs, E. (2004). Conditional Variance in Count Data Regression. *Communication in Statistics; Theory and Methods* **33**, 2745-2758.

Brännäs, K. and Hall, A. (2001). Estimation in Integer-Valued Moving Average Models. *Applied Stochastic Models in Business and Industry* **17**, 277-291.

Brännäs, K., Hellström, J. and Nordström, J. (2002). A New Approach to Modelling and Forecasting Monthly Guest Nights in Hotels. *International Journal of Forecasting* **18**, 19-30.

Brännäs, K. and Quoreshi, S. (2004). Integer-Valued Moving Average Modelling of the Number of Transactions in Stocks. *Umeå Economic Studies* **637.**

Cameron, A.C. and Trivedi, P.K. (1998). *Regression Analysis of Count Data.* Cambridge: Cambridge University Press.

Diebold, F.X. (1989). Random Walks versus Fractional Integration: Power Comparisons of Scalar and Joint Tests of the Variance-time Function. In: B. Raj, ed., *Advances in Econometrics.* Dordrecht: Kluwer.

Engle, R.F. and Russell, J.R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* **66**, 1127-1162.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Application to Poisson Models. *Econometrica* **52**, 701-720.

Hansen, L.P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* **50**, 1029-1054.

Heinen, A. and Rengifo, E. (2003). Multivariate Modelling of Time Series Count Data: An Autoregressive Conditional Poisson Model. CORE Discussion Paper 25, Catholic University of Louvain.

McKenzie, E. (1986). Autoregressive Moving-Average Processes with Negative Binomial and Geometric Marginal Distributions. *Advances in Applied Probability* **18**, 679-705.

McLeod, A.I. and Hipel, K.W. (1978). Preservation of the Rescaled Adjusted Range, 1: A Reassessment of the Hurst Phenomenon. *Water Resources Research* **14**, 491-509

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* **59**, 347-370.

Rudholm, N. (2001). Entry and the Number of Firms in the Swedish Pharmaceuticals Market. *Review of Industrial Organization* **19**, 351-364.

Rydberg, T.H. and Shephard, N. (1999). BIN Models for Trade-by-trade Data. Modelling the Number of Trades in a Fixed Interval of Time. Working Paper Series W23. Nuffield College, Oxford.

Winkelmann, R. and Zimmermann, K.F. (1995). Recent Developments in Count Data Modelling: Theory and Application. *Journal of Economic Surveys* **9**, 1-24.

A Vector Integer-Valued Moving Average Model for High Frequency Financial Count Data[*]

A.M.M. Shahiduzzaman Quoreshi

Department of Economics, 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 vector integer-valued moving average (VINMA) model is introduced. The VINMA model allows for both positive and negative correlations between the counts. The conditional and unconditional first and second order moments are obtained. The CLS and FGLS estimators are discussed. The model is capable of capturing the covariance between and within intra-day time series of transaction frequency data due to macroeconomic news and news related to a specific stock. Empirically, it is found that the spillover effect from Ericsson B to AstraZeneca is larger than that from AstraZeneca to Ericsson B.

Key Words: Count data, Intra-day, Time series, Estimation, Reaction time, Finance.

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

1 Introduction

This paper introduces a Vector Integer-Valued Moving Average (VINMA) model. The VINMA is developed to capture covariance in and between stock transactions time series. Each transaction refers to a trade between a buyer and a seller in a volume of stocks for a given price. A transaction is impounded with information such as volume, price and spread. The trading intensity or the number of transactions for a fixed interval of time and the durations can be seen as inversely related since the more time elapses between successive transactions the fewer trades take place. Easley and O'Hara (1992) shows that a low trading intensity implies no news. Engle (2000) models time according to the autoregressive conditional duration (ACD) model of Engle and Russell (1998) and finds that longer durations are associated with lower price volatilities. One obvious advantage of the VINMA model over extensions of the ACD model is that there is no synchronization problem due to different onsets of durations in count data time series. Hence, the spread of shocks and news is more easily studied in the current framework.

The VINMA allows for both negative and positive correlation in the count series and the integer-value property of counts is taken into account. The VINMA model arises from the integer-valued autoregressive moving average (INARMA) model, which is related to the conventional ARMA class of Box and Jenkins (1970). The INARMA model for pure time series is independently introduced by McKenzie (1986) and Al-Osh and Alzaid (1987). An important difference between the continuous variable vector MA (VMA), a special case of vector ARMA model, and the VINMA model is that the latter has parameters that are interpreted as probabilities and hence the values of the parameters are restricted to unit intervals. An introductory treatise of count data is available in, e.g., Cameron and Trivedi (1998).

Until now, the only studies based on the INMA class for intra-day transactions data appear to be Brännäs and Quoreshi (2004) and Quoreshi (2006ab). Quoreshi (2006a) develops a bivariate integer-valued moving average (BINMA) model that satisfies the natural conditions of a count data model. This model is employed to measure the reaction time for news or rumors and how new information is spread through the system. The VINMA model is more general than the BINMA model and enables the study of spillover effects of news from one stock to the other.

2   The VINMA Model

Assume that there are M intra-day series, $y_{1t}, y_{2t}, \ldots, y_{Mt}$, for the number of stock transactions in $t = 1, \ldots, T$ time intervals. Assume further that the dependence between y_{it} and $y_{jt}, i \neq j$, emerges from common underlying factor(s) such as macro-economic news, rumors, etc. Moreover, news related to the y_{jt} series may also have an impact on y_{it} and vice versa.

The covariation within and between the count data variables can be modelled by a VINMA(q) model, with $q = \max(q_1, q_2, \ldots, q_M)$, which can be written on the form

$$
\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Mt} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Mt} \end{pmatrix} + \begin{pmatrix} \alpha_{111} & \alpha_{121} & \cdots & \alpha_{1M1} \\ \alpha_{211} & \alpha_{221} & & \alpha_{2M1} \\ \vdots & \vdots & & \vdots \\ \alpha_{M11} & \alpha_{M21} & & \alpha_{MM1} \end{pmatrix} \circ \begin{pmatrix} u_{1t-1} \\ u_{2t-1} \\ \vdots \\ u_{Mt-1} \end{pmatrix} + \ldots
$$

$$
+ \begin{pmatrix} \alpha_{11q} & \alpha_{12q} & \cdots & \alpha_{1Mq} \\ \alpha_{21q} & \alpha_{22q} & & \alpha_{2Mq} \\ \vdots & \vdots & & \vdots \\ \alpha_{M1q} & \alpha_{M2q} & & \alpha_{MMq} \end{pmatrix} \circ \begin{pmatrix} u_{1t-q} \\ u_{2t-q} \\ \vdots \\ u_{Mt-q} \end{pmatrix} \tag{1a}
$$

or compactly as

$$
\mathbf{y}_t = \mathbf{u}_t + \mathbf{A}_1 \circ \mathbf{u}_{t-1} + \ldots + \mathbf{A}_q \circ \mathbf{u}_{t-q}. \tag{1b}
$$

The integer-valued innovation sequence $\{\mathbf{u}_t\}$ is assumed independent and identically distributed (iid) with $E(\mathbf{u}_t) = \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_M)'$ and $Cov(\mathbf{u}_t) = \boldsymbol{\Omega}$. Obviously, there is reason to expect the $\mathbf{A}_i$ matrices to become sparser as i increases.

The binomial thinning operator distinguishes the VINMA model from the VMA model. By employing the binomial thinning operator in (1a-b) we account for the integer-value property of count data. The operator can be written

$$
\mathbf{A} \circ \mathbf{u} = \begin{pmatrix} \sum_{i=1}^{M} \alpha_{1i} \circ u_i \\ \vdots \\ \sum_{i=1}^{M} \alpha_{Mi} \circ u_i \end{pmatrix} \tag{2}
$$

where $\alpha_{ki} \circ u_i = \sum_{j=1}^{u_i} z_{jki}$. The $\{z_{jki}\}$ is assumed to be an iid sequence of 0-1 random variables with $\Pr(z_{jki} = 1) = \alpha_{ki} = 1 - \Pr(z_{jki} = 0)$ and the z_{jki} and u_i are assumed to be independent. Since $\alpha_{ki} \in [0, 1]$, $\alpha_{ki} \circ u_i \in [0, u_i]$.

Some conditional and unconditional moment properties of $\alpha_{ki} \circ u_i$ and $\mathbf{A} \circ \mathbf{u}$ are given in the Appendix.

For the covariance matrix $\mathbf{\Omega}$ we assume

$$\mathbf{\Omega}_{ij} = Cov(u_{it}, u_{js}) = \begin{cases} \sigma_{ij} - \lambda_i \lambda_j, & \text{for } t = s \\ 0, & \text{for } t \neq s \end{cases}$$

with $\sigma_{ij} = E(u_{it} u_{jt}), i \neq j$, and $\sigma_{ij} = \sigma_i^2 = E(u_{it}^2), i = j$. The off-diagonal elements $\mathbf{\Omega}_{ij}, i \neq j$, can be positive or negative depending on the relative sizes of σ_{ij} versus $\lambda_i \lambda_j$.

Retaining the previous assumptions, the conditional (on the previous observations, $Y_{t-1} = y_{1t-1}, y_{1t-2}, ..., y_{2t-1}, y_{2t-2}, ...$) first and second order moments for the VINMA(q) model are, in analogy with Brännäs and Hall (2001) and Quoreshi (2006a),

$$E(\mathbf{y}_t | Y_{t-1}) \quad = \quad E_{t|t-1} = \boldsymbol{\lambda} + \sum_{i=1}^{q} \mathbf{A}_i \mathbf{u}_{t-i} \tag{3a}$$

$$\boldsymbol{\Gamma}_{t,t-k|t-1} \quad = \quad E[(\mathbf{y}_t - E_{t|t-1})(\mathbf{y}_{t-k} - E_{t-k|t-k-1}) \mid Y_{t-1}] \tag{3b}$$

$$= \quad \begin{cases} \mathbf{\Omega} + \sum_{i=1}^{q} \mathbf{H}_{it}, & k = 0 \\ \mathbf{0}, & \text{otherwise} \end{cases}$$

where $\text{diag}(\mathbf{H}_{it}) = \mathbf{B}_i \mathbf{u}_{t-i}$ with $\mathbf{B}_i$ an $M \times M$ matrix with elements $(\mathbf{B}_i)_{jk} = \alpha_{jki}(1 - \alpha_{jki})$. Since the conditional variance varies with $\mathbf{u}_{t-i}$ there is conditional heteroskedasticity. When $M = 2$ and matrices $\mathbf{A}_i$ are diagonal the VINMA(q) collapses into the BINMA(q_1, q_2) model of Quoreshi (2006a).

The unconditional first and second order moments for VINMA(q) model can be written

$$E\mathbf{y}_t \quad = \quad \left[\mathbf{I} + \sum_{i=1}^{q} \mathbf{A}_i \right] \boldsymbol{\lambda} \tag{4a}$$

$$Cov(\mathbf{y}_t, \mathbf{y}_{t-k}) \quad = \quad \begin{cases} \mathbf{\Omega} + \sum_{i=1}^{q} \mathbf{A}_i \mathbf{\Omega} \mathbf{A}_i' + \sum_{i=1}^{q} \mathbf{G}_i, & \text{for } k = 0 \\ \mathbf{A}_k \mathbf{\Omega} + \sum_{i=1}^{q} \mathbf{A}_{k+i} \mathbf{\Omega} \mathbf{A}_i', & \text{for } k = 1, 2, \ldots, q \\ \mathbf{0}, & \text{for } \mid k \mid > q \end{cases} \tag{4b}$$

with $\text{diag}(\mathbf{G}_i) = \mathbf{B}_i \boldsymbol{\lambda}$.

We may wish to include explanatory variables in the VINMA(q) model setup. This is most easily done by introducing a time-varying $\boldsymbol{\lambda}_t$ (Brännäs, 1995)

$$\lambda_{jt} = \exp(\mathbf{x}_{jt} \boldsymbol{\beta}_j) \geq 0, \quad j = 1, \ldots, M. \tag{5}$$

Previous prices, etc. are included in $\mathbf{x}_{jt}$. To obtain a more flexible conditional variance specification in (3b) we may let σ_j^2 become time dependent σ_{jt}^2. Allowing σ_{jt}^2 to depend on past values of σ_{jt}^2, u_{jt}, σ_{it}^2 and u_{it}, for $i \neq j$, and explanatory variables, using an exponential form, we may specify (cf. Nelson, 1991)

$$\mathrm{diag}\,(\boldsymbol{\Omega}_t) \;\;=\;\; \exp\left[\boldsymbol{\phi}_0 + \sum_{i=1}^{P} \boldsymbol{\Phi}_i \mathrm{diag}\,(\boldsymbol{\Omega}_{t-i}) \right. \tag{6}$$
$$\left. + \sum_{i=1}^{Q} \boldsymbol{\Theta}_i \mathrm{diag}\,\left(\widetilde{\mathbf{u}}_{t-i}\widetilde{\mathbf{u}}'_{t-i}\right) + \sum_{i=1}^{R} \boldsymbol{\Psi}_i \mathbf{x}'_{t-i}\right]$$

where $\widetilde{\mathbf{u}}_{t-i} = \mathbf{u}_{t-i} - \boldsymbol{\lambda}_{t-i}$. The $\boldsymbol{\phi}_0$ is an M vector with elements ϕ_{j0}, $\boldsymbol{\Phi}_i$, $\boldsymbol{\Theta}_i$ and $\boldsymbol{\Psi}_i$ are $M \times M$ matrices.

3   Estimation

As we specify the model with first and second order moment conditions the conditional least squares (CLS), the feasible generalized least squares (FGLS) and the generalized method of moments (GMM) estimators are first hand candidates for estimation. Here, we only consider the CLS and FGLS estimators. The CLS and FGLS have the residual

$$\mathbf{e}_{1t} = \mathbf{y}_t - E(\mathbf{y}_t | Y_{t-1}) \tag{7}$$

in common and both the CLS and FGLS estimators of $\boldsymbol{\psi} = (\boldsymbol{\psi}'_1, \boldsymbol{\psi}'_2, \ldots, \boldsymbol{\psi}'_M)'$ with $\boldsymbol{\psi}_j$ containing the α-parameters of the jth equation minimize a criterion function of the form

$$S = \sum_{t=q+1}^{T} \mathbf{e}'_{1t} \widehat{\mathbf{W}}_t^{-1} \mathbf{e}_{1t}, \tag{8}$$

where $\mathbf{e}_{1t} = (e_{11t}, e_{21t}, \ldots, e_{M1t})'$, with respect to the unknown parameters. For CLS, $\widehat{\mathbf{W}}_t = \mathbf{I}$, while for FGLS $\widehat{\mathbf{W}}_t = \widehat{\boldsymbol{\Gamma}}_{t,t-k|t-1}$, where $\widehat{\boldsymbol{\Gamma}}_{t,t-k|t-1}$ is a estimate of the conditional covariance matrix in (3b). To calculate the sequences $\{\mathbf{e}_{1t}\}$ we employ $\mathbf{e}_{1t} = \mathbf{u}_t - \boldsymbol{\lambda}_t$.

The conditional variance and the covariance prediction errors

$$e_{j2t} \;\;=\;\; e_{j1t}^2 - \sigma_{jt}^2 - \sum_{i=1}^{q}\sum_{k=1}^{M} \alpha_{jki}(1 - \alpha_{jki})u_{kt-i} \tag{9}$$
$$e_{j3t} \;\;=\;\; e_{i1t}e_{j1t} - \boldsymbol{\Omega}_{ij}, \quad \text{for } i \neq j \tag{10}$$

are used for FGLS estimation. $S_{j2} = \sum_{t=s}^{T} e_{j2t}^2$ and $S_{j3} = \sum_{t=s}^{T} e_{j3t}^2$, where $s = \max(q, P, Q, R)+1$, are minimized with respect to the parameters of the function σ_{jt}^2 and σ_{ij} and with the CLS estimates for the jth equation $\widehat{\psi}_j$ and $\widehat{u}_{jt}$ kept fixed. For time invariant $\mathbf{\Omega}$ a simple and obvious moment estimator

$$\widehat{\mathbf{\Omega}} = (T-s)^{-1} \sum_{t=s}^{T} \left[(\mathbf{e}'_{1t}\mathbf{e}_{1t}) - \sum_{i=1}^{q} \mathbf{H}_{it} \right]$$

follows from (3b). The covariance matrix estimators for CLS and FGLS are

$$Cov(\widehat{\psi}_{CLS}) = (\sum_{t=s}^{T} \frac{\partial \mathbf{e}_{1t}}{\partial \psi'} \frac{\partial \mathbf{e}_{1t}}{\partial \psi})^{-1}$$

$$Cov(\widehat{\psi}_{FGLS}) = (\sum_{t=s}^{T} \frac{\partial \mathbf{e}_{1t}}{\partial \psi'} \widetilde{\mathbf{\Gamma}}_{t,t-k|t-1}^{-1} \frac{\partial \mathbf{e}_{1t}}{\partial \psi})^{-1}.$$

The $\widetilde{\mathbf{\Gamma}}_{t,t-k|t-1}$ is the covariance matrix for the residual series from FGLS estimation.

4 Empirical Results

Tick-by-tick data for Ericsson B and AstraZeneca are aggregated over five minute intervals of time. The covered period is November 5-December 12, 2002. Since our intention is to capture ordinary transactions we have deleted trading before 0935 (trading opens at 0930) and after 1714 (order book closes at 1720). There are altogether 2392 observations for each stock series. Both CLS and FGLS estimators are employed for the VINMA model and the AIC criterion is used to find lag lengths for the VINMA(q) model. The FGLS estimator turns out to be the better one in terms of eliminating serial correlations. The parameters for Ericsson B (α_{11i}) and AstraZeneca (α_{22i}) estimated by FGLS are presented in Figure 1 (left panel). All estimates are positive and significant at the 5 percent level. The parameters to capture spillover effects from AstraZeneca to Ericsson B (α_{12i}) and from Ericsson B to AstraZeneca (α_{21i}) are presented in Figure 1 (right panel). About 19 percent of the estimated mean transactions for AstraZeneca is due to spillover effects while about 2 percent of the estimated mean transactions for Ericsson B is due to spillover.[1] The α_{21i} are all significant until lag 16 except for lags $6-8$ and 13, while the α_{12i} are all

---

[1]To calculate the percent of mean transactions for y_j due to spillover effect from y_k we employ $100 \cdot (\sum_{i=0}^{q} \alpha_{jki} u_{kt-i}/E(y_{jt}|Y_{t-1}))$.
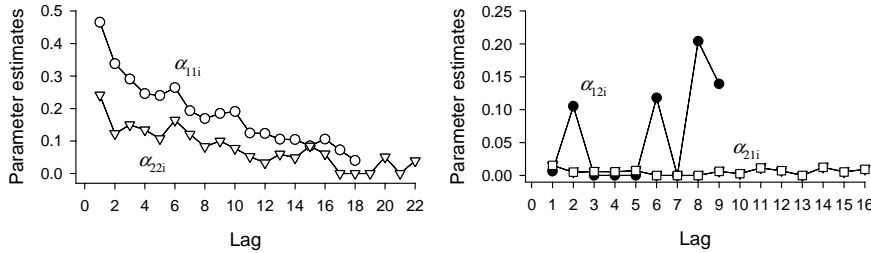
Figure 1: The circles and the triangles are the moving average parameters for Ericsson B $(\widehat{\lambda}_1 = 13.16)$ and AstraZeneca $(\widehat{\lambda}_2 = 1.97)$ (left figure). The solid circles capture the impact of AstraZeneca on Ericsson B while the squares capture the impact of Ericsson B on AstraZeneca (right figure).

significant until lag 9 except for lags $1, 3 - 5$ and 7. This implies that Ericsson B influences AstraZeneca for a longer period of time than is the case in the other direction.

The estimation results for the VINMA and BINMA (see Quoreshi, 2006a) models are summarized in Table 1. For FGLS, the VINMA model is marginally better than the BINMA model in terms of goodness of fit. For the VINMA model, the adjusted R^2 for Ericsson B increases from the BINMA model by 1.1 percent while it increases by 4.9 percent for AstraZeneca. It is found that news related to AstraZeneca Granger-causes Ericsson B and vice versa. The conditional correlations between the stock series at lag zero estimated with VINMA and BINMA models are 0.16 are 0.15, respectively. This implies that the intensity of trading for both stocks moves in the same direction, i.e. increases or decreases, due to macroeconomic news and news related to a specific stock. The corresponding estimated unconditional correlation for the VINMA model is 0.21 while the correlation between the two stock series in the sample is 0.28.

For FGLS (CLS), the α_{11i} and α_{22i} estimates give a mean reaction time (RT_m) of 26.07 (26.21) and 23.23 (17.23) minutes, respectively.[2] For FGLS (CLS), the α_{11i} and α_{22i} estimates give a median reaction (RT_{me}) time of 20 (20) and 15 (10), respectively. Hence, for the measurement of reaction time, the choice of mean or median reaction time matters.

---

[2] As measures of reaction times to macroeconomic news/rumors in the $\{u_{jt}\}$ sequence on y_{jt} we use the mean lag $\sum_{i=0}^{q_j} i\alpha_{jji}/w$, where $w = \sum_{i=0}^{q_j} \alpha_{jji}$ and where $\alpha_{jj0} = 1$. Alternatively, we use the median lag, which is the smallest k such that $\sum_{i=0}^{k} \alpha_{jji}/w \geq 0.5$.

Table 1: Results for VINMA and BINMA models for Ericsson B (Stock 1) and AstraZeneca (Stock 2).

| | VINMA | | | | BINMA | |
| | CLS | | FGLS | | FGLS | |
	Stock 1	Stock 2	Stock 1	Stock 2	Stock 1	Stock 2
$\overline{R}^2$	0.505	0.209	0.504	0.231	0.498	0.221
RT_m	26.21	17.12	26.07	23.23	25.91	24.64
RT_{me}	20.00	10.00	20.00	15.00	20.00	15.00
LB_{30}	31.58	70.26	32.87	17.96	34.67	18.41
$\widehat{\rho}_{0|t-1}$	0.159		0.160		0.150	

5   Concluding Remarks

This study introduces a vector integer-valued moving average (VINMA) model. The conditional and unconditional first and second order moments are obtained. The VINMA model allows for both positive and negative correlations between the counts. The model is capable of capturing the covariance between and within intra-day time series of transaction frequency data due to macroeconomic news and news related to a specific stock. In its empirical application, we found that the spillover effect from Ericsson B to AstraZeneca is larger than that from AstraZeneca to Ericsson B. The FGLS estimator performs better than the CLS estimator in terms of eliminating serial correlations.

Appendix

Conditionally on the integer-valued u_i, $\alpha_{ki} \circ u_i$ is binomially distributed with $E(\alpha_{ki} \circ u_i \mid u_i) = \alpha_{ki} u_i$, $V(\alpha_{ki} \circ u_i \mid u_i) = \alpha_{ki}(1 - \alpha_{ki})u_i$ and $E[(\alpha_{ki} \circ u_i)(\alpha_{kj} \circ u_j)\mid u_i, u_j] = \alpha_{ki}\alpha_{kj}u_i u_j$, for $i \neq j$. Unconditionally it holds that $E(\alpha_{ki} \circ u_i) = \alpha_{ki}\lambda_i$, $V(\alpha_{ki} \circ u_i) = \alpha_{ki}^2 \sigma_i^2 + \alpha_{ki}(1 - \alpha_{ki})\lambda_i$ and $E[(\alpha_{ki} \circ u_i)(\alpha_{kj} \circ u_j)] = \alpha_{ki}\alpha_{kj}E(u_i u_j)$, for $i \neq j$, where $E(u_i) = \lambda_i$ and $V(u_i) = \sigma_i^2$.

Assuming independence between and within the thinning operations, conditionally on $M \times 1$ integer-valued vector $\mathbf{u}$, $\mathbf{A} \circ \mathbf{u}$ has

$$E(\mathbf{A} \circ \mathbf{u} \mid \mathbf{u}) = \mathbf{A}\mathbf{u}$$

$$E[(\mathbf{A}_i \circ \mathbf{u}_{t-i})(\mathbf{A}_j \circ \mathbf{u}_{t-j})' \mid \mathbf{u}_{t-i}, \mathbf{u}_{t-j}] = \begin{cases} \mathbf{A}_i \mathbf{u}_{t-i}\mathbf{u}_{t-i}' \mathbf{A}_i' + \mathbf{H}_{it}, & \text{for } i = j \\ \mathbf{A}_i \mathbf{u}_{t-i}\mathbf{u}_{t-j}' \mathbf{A}_j', & \text{for } i \neq j \end{cases}$$

where the $\mathbf{A}$ is a $M \times M$ matrix with elements $\alpha_{ki} \in [0,1]$ and diag($\mathbf{H}_{it}$) = $\mathbf{B}\mathbf{u}_{t-i}$. The $\mathbf{B}$ is an $M \times M$ matrix with elements $\alpha_{ki}(1 - \alpha_{ki})$. The corresponding unconditional first and second order moments are

$$E(\mathbf{A} \circ \mathbf{u}) = \mathbf{A}E(\mathbf{u}) = \mathbf{A}\boldsymbol{\lambda}$$

$$E[(\mathbf{A}_i \circ \mathbf{u}_{t-i})(\mathbf{A}_j \circ \mathbf{u}_{t-j})'] = \begin{cases} \mathbf{A}_i E\left(\mathbf{u}_{t-i}\mathbf{u}_{t-i}'\right)\mathbf{A}_i' + \mathbf{G}, & \text{for } i = j \\ \mathbf{A}_i E\left(\mathbf{u}_{t-i}\mathbf{u}_{t-j}'\right)\mathbf{A}_j', & \text{for } i \neq j \end{cases}$$

where diag($\mathbf{G}$) = $\mathbf{B}\boldsymbol{\lambda}$.

References

Al-Osh, M.A. and Alzaid, A.A. (1987). First Order Integer-Valued Autoregressive INAR(1)) Process. *Journal of Time Series Analysis* **8**, 261-275.

Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis: Forecasting and Control.* San-Francisco: Holden-Day.

Brännäs, K. (1995). Explanatory Variables in the AR(1) Model. *Umeå Economic Studies* **381**.

Brännäs, K. and Hall, A. (2001). Estimation in Integer-Valued Moving Average Models. *Applied Stochastic Models in Business and Industry* **17**, 277-291.

Brännäs, K. and Quoreshi, A.M.M.S. (2004). Integer-Valued Moving Average Modelling of the Number of Transactions in Stocks. *Umeå Economic Studies* **637**.

Cameron, A.C. and Trivedi, P.K. (1998). *Regression Analysis of Count Data.* Cambridge: Cambridge University Press.

Easley, D. and O'Hara, M. (1992). Time and the Process of Security Price Adjustment. *Journal of Finance* **47**, 905-927.

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

Engle, R.F. and Russell, J.R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* **66**, 1127-1162.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Application to Poisson Models. *Econometrica* **52**, 701-720.

McKenzie, E. (1986). Autoregressive Moving-Average Processes with Negative Binomial and Geometric Marginal Distributions. *Advances in Applied Probability* **18**, 679-705.

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* **59**, 347-370.

Quoreshi, A.M.M.S. (2006a). Bivariate Time Series Modelling of Financial Count Data. *Communications in Statistics–Theory and Methods* **35**, Issue 7. Forthcoming.

Quoreshi, A.M.M.S. (2006b). A Long Memory, Count Data, Time Series Model for Financial Application. *Umeå Economic Studies* **673**.

Long Memory, Count Data, Time Series Modelling for Financial Application[*]

A.M.M. Shahiduzzaman Quoreshi
Department of Economics, 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 model to account for the long memory property in a count data framework is proposed and applied to high frequency stock transactions data. The unconditional and conditional first and second order moments are given. The CLS and FGLS estimators are discussed. In its empirical application to two stock series for AstraZeneca and Ericsson B, we find that both series have a fractional integration property.

Key Words: Intra-day, High frequency, Estimation, Fractional integration, Reaction time.

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

1 Introduction

This paper focuses on modelling the long memory property of time series of count data and on applying the model in a financial setting. The long range dependence or the long memory implies that the present information has a persistent impact on future counts. Note that the long memory property is related to the sampling frequency of a time series. A manifest long memory may be shorter than one hour if observations are recorded every minute, while stretching over decades for annual data. A time series of count data is an integer-valued and non-negative sequence of count observations observed at equidistant instants of time. In the current context series typically have small counts and many zeros. Models for long memory, continuous variable time series are not applicable for integer-valued time series. This is so with respect to both interpretation and inference.

The long memory phenomenon in time series was first considered by Hurst (1951, 1956). In these studies, he explained the long term storage requirements of the Nile River. He showed that the cumulated water flows in a year depends not only on the water flows in recent years but also on water flows in years much prior to the present year. Mandelbrot and van Ness (1968) explain and advance the Hurst's studies by employing fractional Brownian motion. In analogy with Mandelbrot and van Ness (1968), Granger (1980), Granger and Joyeux (1980) and Hosking (1981) develop Autoregressive Fractionally Integrated Moving Average (ARFIMA) models to account for the long memory in time series data. Ding and Granger (1996) point out that a number of other processes can also have the long memory property. A recent empirical study regarding the usefulness of ARFIMA models is conducted by Bhardwaja and Swanson (2005), who found strong evidence in favor of ARFIMA in absolute, squared and log-squared stock index returns.

In this paper, we develop a model to account for the long memory property in a count data framework. We propose an integer-valued ARFIMA (INARFIMA) model and apply the model to high frequency stock transaction data. Each transaction refers to a trade between a buyer and a seller in a volume of stocks for a given price. The model can be used to measure the reaction times for, e.g., macro-economic news or rumors and captures information spread through the system.

The paper is organized as follows. The INARMA and ARFIMA models are discussed and INARFIMA models are introduced in Section 2. The conditional and unconditional moment properties of the INARFIMA models are obtained. A discussion on model identification is given in Section 3. The es-

timation procedures, CLS and FGLS for unknown parameters are discussed in Section 4. A detailed description of the empirical data is given in Section 5. The empirical results for the stock series are presented in Section 6 and the concluding comments are included in the final section.

2   Model

Many economic time series, e.g., the number of transactions, the number of car passes during an interval of time, comprise integer-valued count data. It is reasonable to assume that this type of data may also have long memory. However, if employing the previous workhorse, the ARFIMA model, integers can not be generated. By combining features of the INARMA and ARFIMA models, we are able to introduce a count data (integer-valued) autoregressive fractionally integrated moving average (INARFIMA) model that takes account of both the integer-valued property of counts and incorporates the long memory property.

2.1   The INARFIMA Model

The INARMA model for a time series $y_1, \ldots, y_T$ is introduced independently by McKenzie (1986) and Al-Osh and Alzaid (1987). The INARMA model can be written

$$y_t - \alpha_1 \circ y_{t-1} - \ldots - \alpha_p \circ y_{t-p} = u_t + \beta_1 \circ u_{t-1} + \ldots + \beta_q \circ u_{t-q}. \quad (1)$$

Here, the binomial thinning operator is the key device enabling integer-values to arise in the model. The operator can be written

$$\varphi \circ v = \sum_{i=1}^{v} z_i \quad (2)$$

with $\{z_i\}_{i=1}^{v}$ an iid sequence of 0-1 random variables and with z_i and v as independent variables. It holds that $\Pr(z_i = 1) = \varphi = 1 - \Pr(z_i = 0)$. Conditionally on the integer-valued v, $\varphi \circ v$ is binomially distributed with $E(\varphi \circ v \mid v) = \varphi v$ and $V(\varphi \circ v \mid v) = \varphi(1 - \varphi)v$. Unconditionally it holds that $E(\varphi \circ v) = \varphi \mu$ and $V(\varphi \circ v) = \varphi^2 \sigma_v^2 + \varphi(1 - \varphi)\mu$, where $E(v) = \mu$ and $V(v) = \sigma_v^2$. Obviously, $\varphi \circ v \in [0, v]$. In equation (1) the $\{u_t\}$ is an iid sequence of non-negative integer-valued random variables with $E(u_t) = \lambda$ and $V(u_t) = \sigma^2$. Since $\alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_q$ are all thinning probabilities, they are restricted to fall in unit intervals.

Granger and Joyeux (1980) and Hosking (1981) independently propose ARFIMA models. We say that $\{y_t, t = 1, 2, \ldots, T\}$ is an ARFIMA $(0, d, 0)$ model if

$$(1 - L)^d y_t = a_t \qquad (3)$$

where L is a lag operator and d is a real number. The $\{a_t\}$ is a white noise process of random variables with mean $E(a_t) = 0$ and variance $V(a_t) = \sigma_a^2$. Employing binomial series expansion, we can write

$$(1 - L)^d = \Delta^d = 1 - \sum_{i=1}^{\infty} \frac{(i - 1 - d)!}{i!(-d - 1)!} L^i = 1 - \sum_{i=1}^{\infty} \frac{\Gamma(i - d)}{\Gamma(i + 1)\Gamma(1 - d)} L^i \qquad (4)$$

and correspondingly

$$
\begin{aligned}
\Delta^{-d} &= 1 + dL + \frac{1}{2} d(1 + d) L^2 + \frac{1}{6} d(1 + d)(2 + d) L^3 - \ldots \\
&= 1 + \sum_{i=1}^{\infty} \frac{(i + d - 1)!}{i!(d - 1)!} L^i = 1 + \sum_{i=1}^{\infty} \frac{\Gamma(i + d)}{\Gamma(i + 1)\Gamma(d)} L^i \qquad (5)
\end{aligned}
$$

where $\Gamma(n + 1) = n!$ and $i = 1, 2, \ldots$. The Δ^d is needed for AR(∞) and the Δ^{-d} is needed for MA(∞) representations of the ARFIMA $(0, d, 0)$ model or for more general ARFIMA(p, d, q) models. If $d < 1/2$, $d \neq 0$, the ARFIMA$(0, d, 0)$ model is said to have long memory. The model has mean reversion when $d < 1$, while the model has mean reversion but is not covariance stationary when $d \in (1/2, 1)$. A survey of the ARFIMA literature can be found in Baillie (1996).

Combining the ideas of the INARMA model with fractional integration is not quite straightforward. Direct use of (4) or (5) will not give integer-values since multiplying an integer-valued variable with a real-valued d can not produce an integer-valued result and this alternative is hence ruled out. In order to set up an operational model we may instead depart from the binomial expansion expression and be careful with placing the thinning operator properly. Importantly, d and functions of d will in this setting be part of the binomial thinning operations and hence there will be a shift in interpretation. In analogy with Granger and Joyeux (1980) and Hosking (1981) we can consider the following INMA(∞) representation of the INARFIMA $(0, d, 0)$ model

$$
\begin{aligned}
y_t &= u_t + d_1 \circ u_{t-1} + d_2 \circ u_{t-2} + d_3 \circ u_{t-3} + \ldots \\
y_t &= (1 + d_1 \circ L + d_2 \circ L^2 + d_3 \circ L^3 + \ldots) u_t \qquad (6) \\
y_t &= (1 - L^\circ)^{-d} u_t
\end{aligned}
$$

where $d_i = \Gamma(i + d)/\left[\Gamma(i + 1)\Gamma(d)\right], i \geq 1$ and the notation $(L^\circ)^i = \circ (L)^i$, for $i > 0$ is introduced. The $(1 - L^\circ)^{-d}$ is a slight alteration of (5). By this, we take account of the integer-valued property. The coefficients d_i in expression (6) are considered thinning probabilities and hence we require $d \in [0, 1]$.

Employing the same idea, we can write

$$(1 - L^\circ)^d y_t = u_t \tag{7}$$

for an INAR(∞) representation of the INARFIMA $(0, d, 0)$ model. Here, $(1 - L^\circ)^d$ is a slight rearrangement of (4). Note that the models in (6) and (7) are two different representations of the INARFIMA $(0, d, 0)$ model and can not be considered as inversely related due to the thinning operations. The models have similar first order moments but differ in second order moments. Since the second order moments for an INMA(∞) representation of the INARFIMA are less complicated to obtain than those of a corresponding INAR(∞) representation the former is employed throughout the paper.

We say that $\{y_t, t = 1, 2, \ldots, T\}$ is an INARFIMA (p, d, q) model when

$$\alpha(L^\circ)y_t = \beta(L^\circ)(1 - L^\circ)^{-d}u_t. \tag{8}$$

In (8) $\alpha(L^\circ) = 1 - \alpha_1 \circ L - \alpha_2 \circ L^2 - \ldots - \alpha_p \circ L^p$ and $\beta(L^\circ) = 1 + \beta_1 \circ L + \beta_2 \circ L^2 + \ldots + \beta_q \circ L^q$, are lag polynomials of orders p and q, respectively. Note that we require $\alpha_i, \beta_j, d \in [0, 1]$, for $i > 0$ and $j \geq 0$, for an INARFIMA(p, d, q) model. Hence, the AR, MA and fractional integration parameters of an INARFIMA model are more restricted than the corresponding parameters of the ARFIMA model. When $d = 0$, the INARFIMA(p, d, q) becomes an INARMA(p, q) while for $d = 1$ it turns into an INARIMA$(p, 1, q)$.

In analogy with Brännäs and Hall (2001), who give the conditional moments for an INMA model, we can write the conditional first and second order moments for the INARFIMA(p, d, q)

$$E(y_t|Y_{t-1}) = \sum_{j=1}^{p}\alpha_j y_{t-j} + \lambda + \sum_{j=1}^{q}\beta_j u_{t-j} + \sum_{j=0}^{q}\beta_j\sum_{i=1}^{\infty}d_i u_{t-i-j} \tag{9a}$$

$$V(y_t|Y_{t-1}) = \sum_{j=1}^{p}\alpha_j(1 - \alpha_j)y_{t-j} + \sigma^2 + \sum_{j=1}^{q}\beta_j(1 - \beta_j)u_{t-j}$$

$$+ \sum_{j=0}^{q}\beta_j\sum_{i=1}^{\infty}d_i(1 - \beta_j d_i)u_{t-i-j}. \tag{9b}$$

Note that the moments are conditioned on only the previous observations, $Y_{t-1} = (y_{t-1}, y_{t-2}, \ldots)$. Whether the conditional variance is overdispersed,

underdispersed or equidispersed depends on the relative sizes of σ^2 and λ. The effects of y_{t-j} and u_{t-i}, $j, i \geq 1$ are larger on the mean than on the variance since $\alpha_j > \alpha_j(1 - \alpha_j)$, $\beta_j > \beta_j(1 - \beta_j)$ and $\beta_j \sum_{i=1}^{\infty} d_i > \beta_j \sum_{i=1}^{\infty} d_i(1 - \beta_j d_i)$, $j \geq 0$. Since the conditional variance varies with y_{t-j} and u_{t-i} there is a conditional heteroskedasticity property for which we can use the shorthand notation ARFIMACH(p, d, q) (cf. the notations ARCH(p) and MACH(q)).

The second order unconditional moments of a general INARFIMA(p, d, q) model are quite complicated. In empirical applications, INARFIMA(p, d, q) models with low p and q orders are likely to be of most interest. So, instead of studying the general INARFIMA(p, d, q), we focus on INARFIMA$(p, d, 0)$ and INARFIMA$(0, d, q)$ models in some detail.

First, consider the INARFIMA$(0, d, q)$ model

$$y_t = \beta(L^\circ)(1 - L^\circ)^{-d} u_t. \tag{10}$$

Assuming independence between and within the thinning operations and that $\{u_t\}$ is an iid sequence with mean λ and variance σ^2, the unconditional mean and variance of an INARFIMA$(0, d, p)$ are

$$E(y_t) = \lambda D \sum_{j=0}^{q} \beta_j \tag{11a}$$

$$V(y_t) = \lambda \sum_{j=0}^{q} \beta_j \left(\sum_{i=1}^{\infty} d_i(1 - d_i\beta_j) + \sum_{j=0}^{q}(1 - \beta_j) \right) \tag{11b}$$

$$+ \sigma^2 D^2 \sum_{j=0}^{q} \beta_j^2$$

with $D = 1 + \sum_{i=1}^{\infty} d_i$, $D^2 = 1 + \sum_{i=1}^{\infty} d_i^2$ and $d_i = \Gamma(i+d)/\left[\Gamma(i+1)\Gamma(d)\right]$, $i \geq 1$. It is clear from (10a-b) that the mean and variance only generate positive values when $d \in [0, 1]$ since all β_j are positive. Since the λ and σ^2 are not functions of time and $\sum_{i=1}^{\infty} d_i \geq \sum_{i=1}^{\infty} d_i^2$ and $\sum_{i=1}^{\infty} d_i > \sum_{i=1}^{\infty} d_i(1 - d_i)$ for $d \in [0, 1]$, it is sufficient that $\sum_{i=1}^{\infty} d_i < \infty$ for $\{y_t\}$ to be a stationary sequence. Note that for $d \in (0, 1)$ the d_i decreases as the lag i increases. Note also that we can not determine values for d_i when i is large since both $\Gamma(i+d)$ and $\Gamma(i+1)$ approach infinity. However, we can approximate d_i for large i with $i^{(d-1)}/\Gamma(d)$ (Granger and Joyeux, 1980 and Hosking, 1981). When $d = 0.6, 0.4$ and 0.2 the approximate values for $d_{9999} - d_{10000} = 6.7e^{-7}, 1.1e^{-7}$ and $1.1e^{-8}$, respectively. Hence, the function of d converges. For an invertible INARMA$(0, d, q)$ model, $d_i < 1$ and $\beta_j < 1$ for $i, j > 0$ are required.

General forms of the autocorrelation function for INARFIMA$(0, d, q)$ can be obtained, but expressions are complicated. For simplicity, we consider the autocorrelation function for an INARFIMA$(0, d, 1)$, which is

$$\rho_k = \sigma^2 V^{-1}(y_t) \sum_{j=0}^{2} B_j \left(d_{k+j-1} + \sum_{i=1}^{\infty} d_i d_{i+k+j-1} \right), \quad k \geq 1 \qquad (12)$$

with $B_0 = B_2 = \beta_1$, $B_1 = \beta_0 + \beta_1^2$ and $d_0 = 1$.

The INARFIMA$(p, d, 0)$ model can be written

$$\alpha(L^\circ) y_t = (1 - L^\circ)^{-d} u_t. \qquad (13)$$

Retaining previous assumptions, we can give the mean and variance as

$$E(y_t) = \lambda D \left(1 - \sum_{j=1}^{p} \alpha_j \right)^{-1} \qquad (14a)$$

$$V(y_t) = \left(1 - \sum_{j=1}^{p} \alpha_j^2 \right)^{-1} \left[\lambda \sum_{i=1}^{\infty} d_i(1 - d_i) + \sigma^2 D^2 + \qquad (14b) \right.$$

$$\left. + \sum_{j=1}^{p} \alpha_j(1 - \alpha_j) E(y_t) \right]$$

with $d_i = \Gamma(i + d)/\left[\Gamma(i+1)\Gamma(d)\right]$. Note that all d_i are positive for $d \in [0, 1]$. Hence, the conditions $\sum_{j=1}^{p} \alpha_j < 1$ and $\sum_{i=1}^{\infty} d_i < \infty$ must be fulfilled in order to generate a finite and positive expected value. Note that the $E(y_t)$ and σ^2 are not functions of time and $\sum_{i=1}^{\infty} d_i \geq \sum_{i=1}^{\infty} d_i^2$, $\sum_{i=1}^{\infty} d_i(1 - d_i)$ for $d \in [0, 1]$. Therefore, it is sufficient that $\sum_{i=1}^{\infty} d_i < \infty$ for $\{y_t\}$ to be a stationary sequence.

The parametric expression of the autocorrelation function of the INARFIMA $(1, d, 0)$ model is

$$\rho_k = \begin{cases} V^{-1}(y_t) \left[\alpha_1^k V(y_t) + \sigma^2 \left\{ \sum_{j=1}^{k} \alpha_1^{k-j} \left(d_j + \sum_{i=1}^{\infty} d_i d_{i+j} \right) \right. \right. \\ \left. \left. + (d_k + \sum_{i=1}^{\infty} d_i d_{i+k}) \right\} \right], \quad k \geq 1 \\ \\ 0, \quad \text{otherwise} \end{cases} \qquad (15)$$

where we set α_1^0 equal to zero.

Autocorrelation functions for INARFIMA$(1, d, 0)$ and INARFIMA$(0, d, 1)$ with different d values are exhibited in Figure 1. All autocorrelation functions
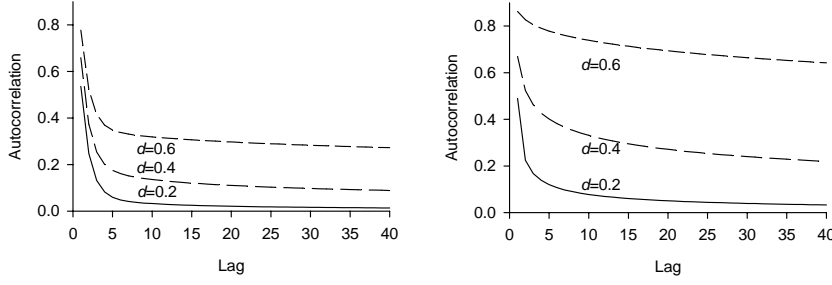
Figure 1: The autocorrelation functions for INARFIMA$(1, d, 0)$ (left figure) with $\alpha = 0.4, \lambda = 2$ and $\sigma^2 = 50$ and INARFIMA$(0, d, 1)$ (right figure) with $\beta = 0.4$ and the same values for the λ and σ^2.

have slower decay after lag 1 and approach zero very slowly. Given λ and σ^2, the higher the value of d the larger is the autocorrelation and the autocorrelations remain higher at all lags. We conjecture, but have been unable to prove, that for $d \in (0, 1)$ the process has long memory if $d < 1/2$ and the process has mean reversion but is not covariance stationary when $d \in (1/2, 1)$, while the process has mean reversion when $d < 1$ Baillie (1996).

2.2 Model Extension

Ding and Granger (1996) point out that a number of other processes have long memory. Here, we extend the model in (8) to a more general form. Consider the following representation

$$(b - Z)^{-d} = b^{-d}(1 - Z/b)^{-d}. \tag{16}$$

The $(1 - Z/b)^d$ can be called a binomial series if $b > 0$ and Z is a number so that $\mid Z/b \mid < 1$. Denoting $b^{-d} = \theta(d)$ and $Z/b = L$ and employing the same idea as in (6) we can write

$$
\begin{aligned}
y_t &= (\theta(d) \circ L^0 + \theta(d)d_1 \circ L^1 + \theta(d)d_2 \circ L^2 + \theta(d)d_3 \circ L^3 + \ldots)u_t \\
&= \theta(d) \circ (1 + d_1 \circ L + d_2 \circ L^2 + d_3 \circ L^3 + \ldots)u_t \\
&= \theta(d) \circ (1 - L^\circ)^{-d}u_t
\end{aligned}
\tag{17}
$$

where $d_i = \Gamma(i + d)/\left[\Gamma(i + 1)\Gamma(d)\right], i \geq 1$ and the property $\varphi_1 \circ (\varphi_2 \circ v) \stackrel{\mathrm{d}}{=} (\varphi_1\varphi_2) \circ v$ is employed. The coefficients in this expression are considered thin-

ning probabilities and hence we require $\theta(d), d \in [0, 1]$. Note that the parameter $\theta(d)$ rescales d_i, for $i \geq 0$. Here, we also use the definition of long memory and the mean recursion property as before, i.e. for $d \in (0, 1)$ we say that the model has long memory if $d < 1/2$ and the model has mean reversion but is not covariance stationary when $d \in (1/2, 1)$, while the model has mean reversion when $d < 1$. We denote the model in (17) INARFIMA$(0, \delta, 0)$. We say that $\{y_t, t = 1, 2, \ldots, T\}$ is an INARFIMA (p, δ, q) model when

$$\alpha(L^\circ)y_t = \theta(\delta) \circ (1 - L^\circ)^{-\delta}\beta(L^\circ)u_t \tag{18}$$

where $\alpha(L^\circ)$ and $\beta(L^\circ)$ are defined as in (8). Note that we require $\alpha_j, \beta_j, \theta(\delta), \delta \in [0, 1]$, for $j > 0$, for an INARFIMA(p, δ, q). The unconditional first and second moments can be given in a similar way to INARFIMA(p, d, q). The conditional first and the second order moments for an INARFIMA(p, δ, q) are

$$
\begin{aligned}
E(y_t|Y_{t-1}) &= \sum_{j=1}^{p} \alpha_j y_{t-j} + \lambda\theta(\delta) + \theta(\delta)\sum_{j=1}^{q}\beta_j u_{t-j} \tag{19a} \\
&\quad + \theta(\delta)\sum_{j=0}^{q}\beta_j\sum_{i=1}^{\infty}\delta_i u_{t-i-j} \\
V(y_t|Y_{t-1}) &= \sum_{j=1}^{p}\alpha_j(1-\alpha_j)y_{t-j} + \theta^2(\delta)\sigma^2 + \sum_{j=1}^{q}\theta(\delta)\beta_j(1-\theta(\delta)\beta_j)u_{t-j} \\
&\quad + \theta(\delta)(1-\theta(\delta))\lambda \tag{19b} \\
&\quad + \sum_{j=0}^{q}\beta_j\sum_{i=1}^{\infty}\theta(\delta)\delta_i(1-\theta(\delta)\beta_j\delta_i)u_{t-i-j}.
\end{aligned}
$$

It is sufficient that $\theta(\delta)\sum_{i=1}^{\infty}\delta_i < \infty$ for $\{y_t\}$ to be a stationary sequence. When $\theta(\delta) = 1$, the INARFIMA(p, δ, q) becomes an INARFIMA(p, d, q) and hence for an invertible INARFIMA$(0, \delta, 0)$, $\delta_i < 1$ for $i > 0$ is required.

3   Model Identification

In this section, we discuss the problem of finding an appropriate INARFIMA model for a given time series. In Figure 2, the autocorrelation and partial autocorrelation functions for integer-valued long memory, INARFIMA(p, d, q), processes are illustrated. The data are generated in accordance with (8) and Matlab codes for generating Poisson and binomial random number are used.

Since the autocorrelations after lag 100 are very small when $d, \delta = 0.4$ the lag length is chosen to be 100 for the INMA(∞) representation of the INARFIMA $(0, d, 0)$ or the corresponding part of the INARFIMA(p, d, q) model in generating the data. The first 500 observations are discarded from 10500 observations in order to avoid start up effects.

The identification of INARFIMA(p, d, q) models is not straightforward. The autocorrelation function for INARFIMA$(1, d, 0)$ and INARFIMA$(1, d, 1)$ look almost alike while the partial-autocorrelation functions are different. The autocorrelation and the partial-autocorrelation functions of the INARFIMA$(0, \delta, 0)$ are quite similar to those of an INARFIMA$(0, d, 1)$. In general, an INARFIMA (p, d, q) model and an INARFIMA(p, δ, q) model have more slowly decaying autocorrelation functions than an INARMA(p, q). Hence, whether a time series has a fractional integration property or not can be identified by the autocorrelation function. But identifying the $\theta(d)$ and the p and/or q lag(s) for an INARFIMA process is difficult by studying the autocorrelation and the partial autocorrelation functions.

There are a number of estimation methods for and tests of long memory, e.g., variance time function $(R(k))$ (Diebold, 1989), rescaled range (RR) (Hurts, 1951), modified rescaled range (MRR) (Lo, 1991), GPH (introduced by Geweke and Porter-Hudak, 1983) and WHI tests (proposed by Künsch, 1987 and modified by Robinson 1995). The GPH and WHI are based on first estimating d, while RR and MRR do not estimate d, to assess whether a series has long memory or not.

By employing AIC and $SBIC$ criteria

$$
\begin{aligned}
AIC &= T \ln \widehat{\sigma}^2 + 2M \\
SBIC &= T \ln \widehat{\sigma}^2 + M \ln T
\end{aligned}
$$

we can choose the lag length of the model. Here, $\widehat{\sigma}^2$ is the variance estimate based on the residuals from the INARFIMA(p, d, q) model, T is the number of observations and $M = p + m + 1$, with m the chosen lag length for estimating d.

4 Estimation

Here, we discuss methods for the estimation of the unknown parameters of the conditional mean and variance functions for the INARFIMA(p, δ, q) model. Since we do not assume a full density function the maximum likelihood estimator is not considered. As we specify the model with first and second moment
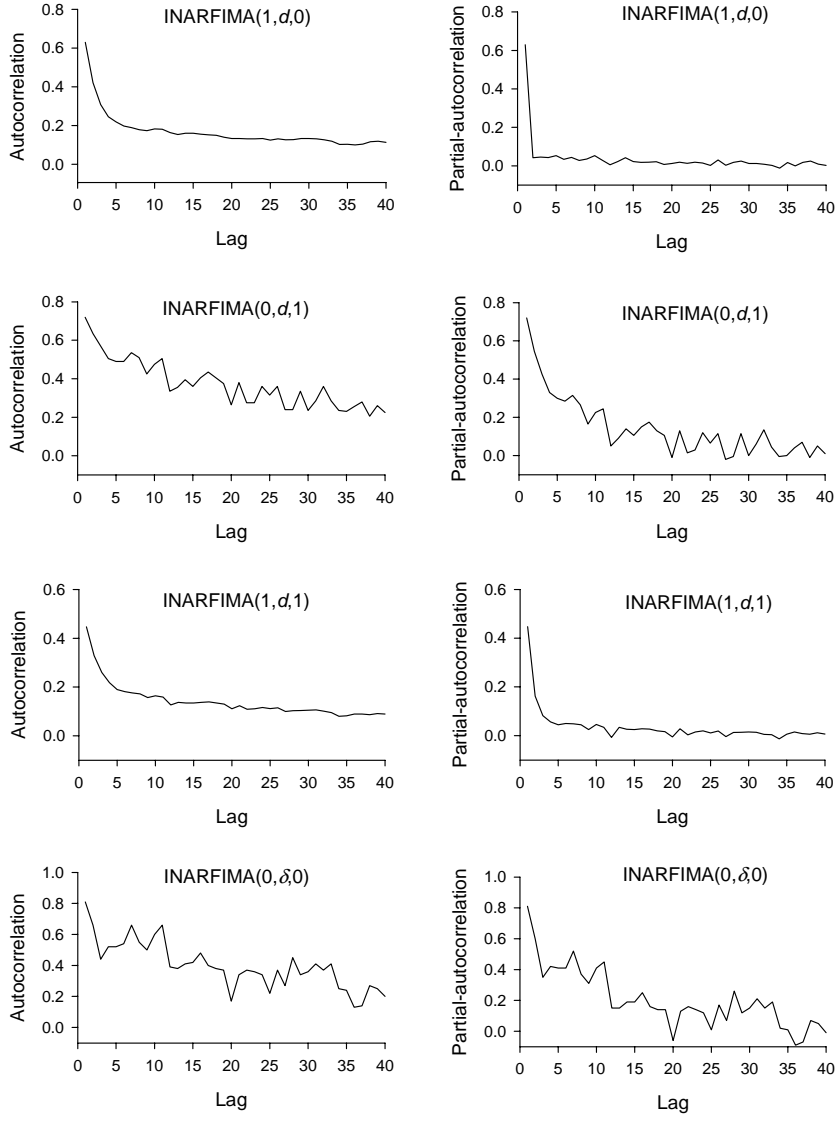
Figure 2: The autocorrelation and partial-autocorrelation functions for the different INARFIMA models with $\alpha = 0.5$, $d = 0.4$, $\theta(\delta) = 0.5$ and $\beta = 0.5$ when applicable. The autocorrelations and partial-autocorrelations for INARFIMA$(0, d, 1)$ and INARFIMA$(0, \delta, 0)$ are multiplied by 5 and 10, respectively.

conditions the conditional least squares (CLS), the feasible generalized least square (FGLS), the generalized method of moments (GMM) estimators and possibly others are candidates for estimation. Here, we only consider the CLS and FGLS estimators. The choice of CLS is obvious since it is easy to estimate and readily available in standard statistical softwares like SPSS. The reason for choosing FGLS instead of GMM is that we may anticipate a better performance of the FGLS than of the GMM estimator (Brännäs, 1995).

Brännäs and Quoreshi (2004) propose CLS and FGLS for INMA(q) and Quoreshi (2006) for bivariate INMA with possibly a large q. Here, the moment conditions are specified in analogy with Brännäs and Quoreshi (2004). To employ the CLS estimator, we need to specify the first moment condition for an INARFIMA(p, δ, q), while we need both the first and the second moment conditions for the FGLS estimator. The CLS estimator for an INARFIMA(p, δ, q) has the following residual

$$
\begin{aligned}
e_{1t} &= y_t - E(y_t | Y_{t-1}) \\
&= y_t - \sum_{j=1}^{p} \alpha_j y_{t-j} - \lambda \theta(\delta) - \theta(\delta) \sum_{j=1}^{q} \beta_j u_{t-j} \\
&\quad - \theta(\delta) \sum_{j=0}^{q} \beta_j \sum_{i=1}^{\infty} \delta_i u_{t-i-j}
\end{aligned}
\tag{20}
$$

and the criterion function $S_1 = \sum_{t=m+1}^{T} e_{1t}^2$ is minimized with respect to the unknown parameters, i.e. $\psi = (\lambda, \alpha', \beta', \theta(\delta)$ and $\delta)$. Using a finite maximum lag m in (20) instead of infinite lags may have biasing effects. Due to the omitted variables, i.e. $u_{t-m-1}, \ldots, u_{t-\infty}$ we may expect a positive biasing effect on the parameters $\alpha', \beta', \theta(\delta)$ and δ (Brännäs and Quoreshi, 2004). Hence, the m should be chosen large. Alternatively and equivalently, the properties $E(e_{1t}) = 0$ and $E(e_{1t} e_{1t-i}) = 0$, $i \geq 1$ could be used. To calculate e_{1t}, we employ $e_{1t} = u_t - \lambda \theta(\delta)$. Note that the moment conditions for an INARFIMA(p, d, q) can be obtained by setting $\theta(\delta) = 1$.

The parameters estimated by CLS are considered a first step of the FGLS estimator. For the next step, the conditional variance prediction errors for INARFIMA(p, δ, q)

$$
e_{2t} = (y_t - E(y_t | Y_{t-1}))^2 - V(y_t | Y_{t-1})
\tag{21}
$$

are used. An obvious least squares estimator for σ^2 is then

$$
\sigma^2 \;=\; \theta^{-2}(\widehat{\delta})\,(T-m)^{-1} \sum_{t=m+1}^{T} \left[\widehat{e}_{1t}^2 - \sum_{j=0}^{p} \widehat{\alpha}_j (1-\widehat{\alpha}_j) y_{t-j} - \theta(\widehat{\delta})(1-\theta(\widehat{\delta}))\widehat{\lambda} \right.
$$

$$
\left. - \sum_{j=1}^{q} \theta(\widehat{\delta})\widehat{\beta}_j (1-\theta(\widehat{\delta})\widehat{\beta}_j) u_{t-j} - \sum_{j=0}^{q} \widehat{\beta}_j \sum_{i=1}^{\infty} \theta(\widehat{\delta})\widehat{\delta}_i (1-\theta(\widehat{\delta})\widehat{\beta}_j\widehat{\delta}_i) u_{t-i-j} \right]
$$

Finally, the FGLS estimator minimizes

$$
S_2 = \sum_{t=m+1}^{T} e_{1t}^2 \widehat{V}^{-1}(y_t|Y_{t-1}) \tag{22}
$$

with $\widehat{V}(y_t|Y_{t-1})$ taken as given. The covariance matrix estimators for CLS and FGLS are:

$$
\begin{aligned}
Cov(\widehat{\psi}_{CLS}) &= \left(\sum_{t=m+1}^{T} \frac{\partial e_{1t}}{\partial \psi} \frac{\partial e_{1t}}{\partial \psi'} \right)^{-1} \\
Cov(\widehat{\psi}_{FGLS}) &= \left(\sum_{t=m+1}^{T} \widehat{V}^{-1}(y_t|Y_{t-1}) \frac{\partial e_{1t}}{\partial \psi} \frac{\partial e_{1t}}{\partial \psi'} \right)^{-1}.
\end{aligned}
$$

5   Data and Descriptives

The tick-by-tick data for Ericsson B and AstraZeneca have been downloaded from the Ecovision system and are later filtered by the author. The stocks are frequently traded and have the highest turnovers at the Stockholmsbörsen. The two 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 minutes of December 5 are 0959 and 1037, respectively. Since we are interested in capturing the number of ordinary transactions, we have deleted all trading before 0935 (trading opens at 0930) and after 1714 (order book closes at 1720). 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 one minute intervals of time. For high frequency data, researchers usually use one, two, five or ten minute intervals of time and the choice is rather arbitrary. There are altogether 11960 observations for both the Ericsson B and AstraZeneca series.

The series together with their autocorrelation and partial-autocorrelation functions are exhibited in Figure 3. There are frequent zero frequencies in both
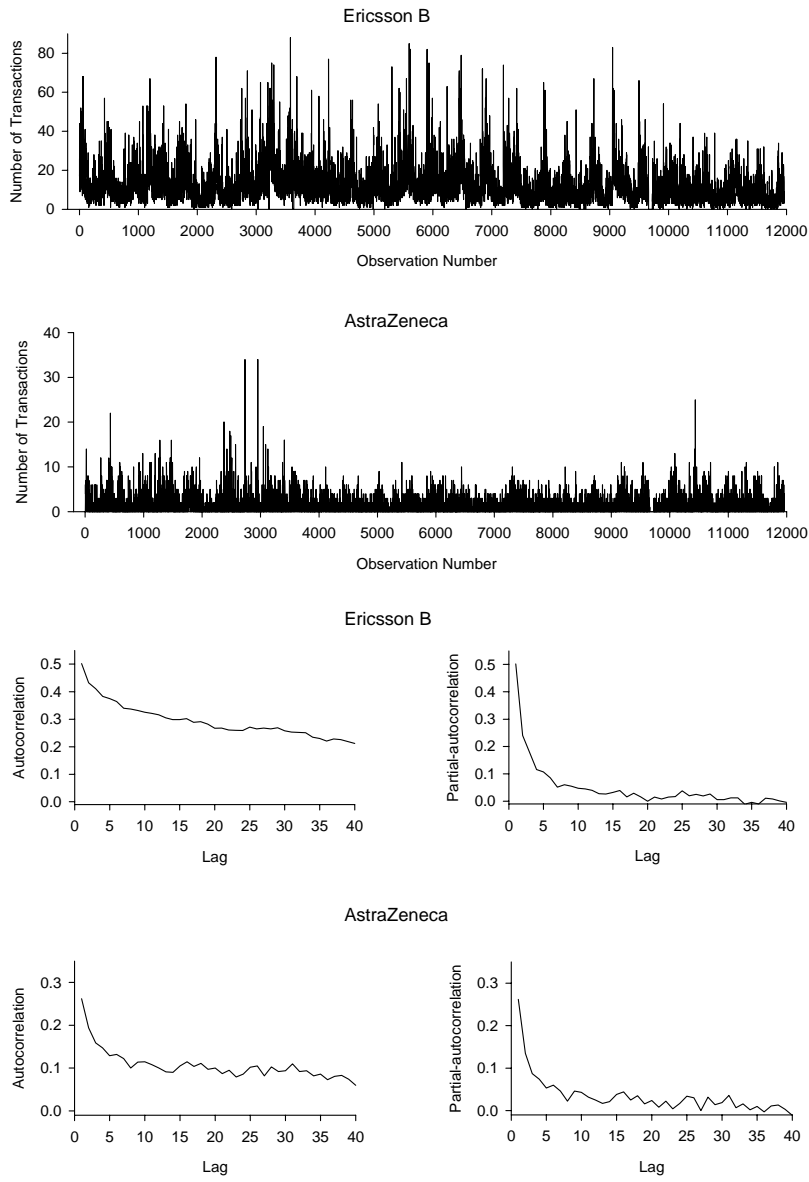
Figure 3: Time series plots for the Ericsson B (mean 11.73 variance 84.86, maximum 88) and AstraZeneca series (mean 1.33 variance 3.75, maximum 34) and their autocorrelation and partial-autocorrelation functions.

series, specially in the AstraZeneca series and hence the application of count data modelling is called for. The counts in both series fluctuate around their means which is an indication of mean reverting processes. The autocorrelation functions for both series suggest fractional integration.

6   Empirical Results

Both CLS and FGLS methods are employed for estimation and the AIC criterion is used to select the lag lengths of the INARFIMA models. Employing CLS and FGLS for Ericsson B an INARFIMA$(0, d, 0)$ with $m = 70$ is chosen while the corresponding model for AstraZeneca is INARFIMA$(0, d, 0)$ with $m = 50$. Serial correlations for the standardized residuals could however not be eliminated. An INARFIMA$(0, d, 1)$ gives a better result in terms of eliminating serial correlations but the estimates of β for both series turn out negative. The estimates of α for the INARFIMA$(1, d, 0)$ for both series also turn out negative. The INMA(70) and INMA(50) for Ericsson B and AstraZeneca, respectively, turns out to be the best in terms of eliminating serial correlation for standardized residuals while INARFIMA$(0, \delta, 0)$ becomes the second best for both series and the estimated parameters are positive. The INARFIMA$(0, \delta, 0)$ is the most parsimonious model in terms of number of parameters.[1]

The empirical results for INARFIMA$(0, \delta, 0)$ for both series are presented in Table 1. For AstraZeneca, we find empirical support for long memory $(\delta < 0.5)$ which implies that the macro-economic news or rumors have persistence impact on the number of transactions. The impact of news on the Ericsson B series can be interpreted in a related way. The series has a mean reversion property but not long memory since the confidence interval for δ includes 0.5. CLS and FGLS perform almost equally well in terms of eliminating serial correlation from standardized residuals. The Ljung-Box statistics, LB_{100} and LB_{200}, for both stocks are larger than the critical values. The reason behind the large values is that we could not eliminate serial correlation at a few of the lags. For AstraZeneca we have remaining serial correlation at lags 31, 57, 59, 70, 154 and 172. The corresponding lags for Ericsson B are 49, 72, 73 and 80. We are not able to provide an explanation to the large correlations at these lags.

In Figure 4 the functions of the fractional integration parameters and the

---

[1] We also estimate truncated INMA(∞) models with $\beta_i = \theta_0 \exp(-\theta_1 i)$ and $\theta_0 \exp(-\theta_1 i - \theta_1 (i - q/2)^2)$ for $i \geq 1$. Though the truncated INMA(∞) models are also parsimonious they performed very poorly for both series in terms of eliminating serial correlations. The Ljung-Box statistic, LB$_{200}$, for Ericsson B for the former model is 370 while for the latter model it is 8740.

Table 1: Results for INARFIMA(0,δ,0) models for Ericsson B and AstraZeneca estimated by CLS and FGLS.

| | Ericsson B | | | | AstraZeneca | | | |
| | CLS | | FGLS | | CLS | | FGLS | |
	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.
δ	0.490	0.021	0.490	0.019	0.372	0.041	0.372	0.046
b	3.336	0.253	3.338	0.321	7.578	1.968	7.565	2.619
$\theta(\delta)$	0.554	0.054	0.554	0.042	0.471	0.097	0.471	0.040
λ	2.309	0.104	2.309	0.093	0.578	0.042	0.578	0.038
σ^2	193.1		193.2		15.82		15.82	
LB_{100}	185.5		185.5		201.6		201.6	
LB_{200}	264.2		264.2		321.2		321.2	

corresponding parameters estimated with INMA(70) and INMA(50) for Ericsson B and AstraZeneca, respectively, are exhibited. The parameters estimated with the INARFIMA(0, δ, 0) models for both stocks look like fitted lines for the corresponding parameters for the INMA models. Hence, we may expect that the reaction times measured by either INMA or INARFIMA would be almost the same. The mean lags for Ericsson B and AstraZeneca measured with the INARFIMA(0, δ, 0) parameters are 22.75 and 13.96 minutes, respectively, while the corresponding mean lags with the truncated INMA are 20.28 and 12.40 minutes.[2] It may appear surprising that the reaction time for Ericsson B is longer than that of AstraZeneca despite the intensity of trading for Ericsson B is almost 9 times larger than that of AstraZeneca. But the result is not that surprising if we consider the price ratio between the two stocks and the turnover (volume times price). During the sample period, the stocks for Ericsson B are traded at a price between SEK 7.10 and 10.40, while the stocks for AstraZeneca are traded at a price between SEK 316.50 and 365.00. The turnovers for the sample period for Ericsson B and AstraZeneca are $5.7 \cdot 10^{13}$ and $3.4 \cdot 10^{12}$, respectively.

The median lags for Ericsson B are 16 and 15 with the INARFIMA and INMA, respectively while the corresponding median lags for AstraZeneca are

---

[2] As measures of reaction times to macroeconomic news/rumors in the $\{u_t\}$ sequence we use the mean lag $\theta(\delta) \sum_{i=0}^{q} i\delta_i/w$, where $w = \theta(\delta) \sum_{i=0}^{q} \delta_i$ and where $\delta_0 = 1$ for an INARFIMA(0, δ, 0). We set $\theta(\delta) = 1$ when the parameters are estimated with an INMA(q). Alternatively, we use the median lag, which is the smallest k such that $\theta(\delta) \sum_{i=0}^{k} \delta_i/w \geq 0.5$.
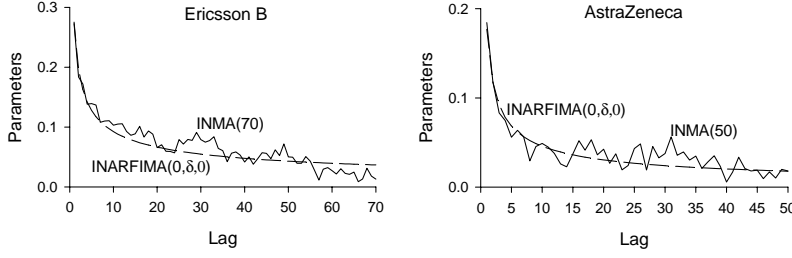
Figure 4: The parameters estimated with INARFIMA$(0, \delta, 0)$ (dashed line) and truncated INMA(∞) models for Ericsson B and AstraZeneca.

8 and 4 minutes. Hence, in estimating mean reaction time, it does not matter much which method we employ. But in estimating median reaction time it may matter more. The large difference in the medians for AstraZeneca is due to the parameter at lag 0. For an INMA(q) the parameter at lag 0 is always 1 while the parameter for an INARFIMA$(0, \delta, 0)$ at lag 0 is $\theta(\delta)$. We see in Figure 4 that the fractional integration functions start with high values and decrease rather sharply in the beginning but decay very slowly afterward. This implies that the trading intensity increases as the news breaks out and fades away very slowly with time.

7   Concluding Remarks

This paper concerns modelling the long memory property in a count data framework. The introduced models emerge from the ARFIMA and INARMA model classes and hence the model is called INARFIMA. The unconditional and conditional first and second moments are given. Moreover, we introduce another process by employing an idea introduced by Granger, Joyeux and Hosking but in a different setting. In its empirical application we find evidence of long memory in the AstraZeneca series, while the estimated δ for Ericsson B indicates a process that has a mean reversion property. CLS and FGLS estimators perform equally well in terms of residual properties. We also find that the trading intensity increases for both stocks when the macro-economic news or rumors break out and the impact remains over a long period and fades away very slowly with time. The reaction due to the macro-economic news on the AstraZeneca series is faster than that of the Ericsson B series.

References

Al-Osh, M.A. and Alzaid, A.A. (1987). First Order Integer-Valued Autoregressive INAR(1)) Process. *Journal of Time Series Analysis* **8**, 261-275.

Baillie, R.T. (1996). Long Memory Process and Fractional Integration in Econometrics. *Journal of Econometrics* **71**, 5-59.

Bhardwaj, G. and Swanson, N.R. (2005). An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. *Journal of Econometrics*. Article in Press.

Brännäs, K. (1995). Explanatory Variables in the AR(1) Model. *Umeå Economic Studies* **381**.

Brännäs, K. and Hall, A. (2001). Estimation in Integer-Valued Moving Average Models. *Applied Stochastic Models in Business and Industry* **17**, 277-291.

Brännäs, K. and Quoreshi, A.M.M.S. (2004). Integer-Valued Moving Average Modelling of the Number of Transactions in Stocks. *Umeå Economic Studies* **637**.

Diebold, F.X. (1989). Random Walks versus Fractional Integration: Power Comparisons of Scalar and Joint Tests of the Variance-time Function. In: B. Raj, ed., *Advances in Econometrics*. Dordrecht: Kluwer.

Ding, Z. and Granger, C.W.J. (1996). Varieties of Long Memory Models. *Journal of Econometrics* **73**, 61-77.

Geweke, J.A. and Porter-Hudak, S. (1983). The Estimation and Application of Long-Memory Time Series Models. *Journal of Time Series Analysis* **4**, S221-S238.

Granger, C.W.J. (1980). Long Memory Relationships and the Aggregation of Dynamic Models. *Journal of Econometrics* **14**, 227-238.

Granger, C.W.J. and Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis* **1**, 15-29.

Hosking, J.R.M. (1981). Fractional Differencing. *Biometrika*, **68**, 165-176.

Hurst, H.E. (1951). Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers* **116**, 770-808.

Hurst, H.E. (1956). Methods of Using Long-Term Storage in Reservoirs. *Proceedings of the Institute of Civil Engineers* **1**, 519-543.

Künsch, H.R. (1987). Statistical Aspects of Self-Similar Processes. In: Prohorov, Y., Sasanov, V.V. (Eds.). *Proceedings of the First World Congress of the Bernoulli Society*. VNU Science Press, Utrecht.

Lo, A. (1991). Long-Term Memory in Stock Market Prices. *Econometrica* **59**, 1279-1313.

Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review* **10**, 422-437.

McKenzie E. (1986). Autoregressive-Moving-Average Processes with Negative Binomial and Geometric Marginal Distribution. *Advances in Applied Probability* **18**, 679-705.

Quoreshi, A.M.M.S. (2006). Bivariate Time Series Modelling of Financial Count Data. *Communications in Statistics–Theory and Methods* **35**, Issue 7. Forthcoming.

Robinson, P. (1995). Gaussian Semiparametric Estimation of Long Rang Dependence. *The Annals of Statistics* **23**, 1630-1661.

Avhandlingar framlagda vid Institutionen för nationalekonomi, Umeå universitet

List of dissertations at the Department of Economics, Umeå University

Holmström, Leif (1972) Teorin för företagens lokaliseringsval. UES 1. PhLic thesis

Löfgren, Karl-Gustaf (1972) Studier i teorin för prisdiskriminering. UES 2. PhLic thesis

Dahlberg, Åke (1972) Arbetsmarknadsutbildning - verkningar för den enskilde och samhället. UES 3. PhD thesis

Stage, Jørn (1973) Verklighetsuppfattning och ekonomisk teori. UES 4. PhLic thesis

Holmlund, Bertil (1976) Arbetslöshet och lönebildning - kvantitativa studier av svensk arbetsmarknad. UES 25. PhD thesis

Löfgren, Karl-Gustaf (1977) En studie i neokeynesiansk arbetslöshets- och inflationsteori. UES 34. PhD thesis

Lundberg, Lars (1976) Handelshinder och handelspolitik - Studier av verkningar på svensk ekonomi. Industriens Utredningsinstitut, Stockholm. PhD thesis

Johansson, Per-Olof (1978) Sysselsättning och samhällsekonomi - En studie av Algots etablering i Västerbotten. UES 53. PhD thesis

Wibe, Sören (1980) Teknik och aggregering i produktionsteorin. Svensk järnhantering 1850-1975; en branschanalys. UES 63. PhD thesis

Ivarson, Lars (1980) Bankers portföljvalsbeteende. En teoretisk studie. UES 64. PhD thesis

Batten, David (1981) Entropy, Information Theory and Spatial Input-output Analysis. UES 92. PhD thesis

Hårsman, Björn (1982) Housing Demand Models and Housing Market Models for Regional and Local Planning. Swedish Council for Building Research, D13:1981. PhD thesis

Holm, Magnus (1983) Regionalekonomiska modeller för planering och samordning i en decentraliserad ekonomi. Byggforskningsrådet, R118:1981 and R5:1983. PhD thesis

Ohlsson, Henry (1986) Cost-Benefit Analysis of Labor Market Programs - Applied to a Temporary Program in Northern Sweden. UES 167. PhLic thesis

Sarafoglou, Nikias (1987) A Contribution to Population Dynamics in Space. UES 179. PhD thesis

Ohlsson, Henry (1988) Cost-Benefit Analysis of Labor Market Programs - Applied to a Temporary Program in Northern Sweden. UES 182. PhD thesis

Anderstig, Christer (1988) Applied Methods for Analysis of Economic Structure and Change. CERUM 1988:2, Umeå University. PhD thesis

Karlsson, Charlie (1988) Innovation Adoption and a Product Life Cycle. UES 185. PhD thesis

Löfström, Åsa (1989) Diskriminering på svensk arbetsmarknad - En analys av löneskillnader mellan kvinnor och män. UES 196. PhD thesis

Axelsson, Roger (1989) Svensk arbetsmarknadsutbildning - En kvantitativ analys av dess effekter. UES 197. PhD thesis

Zhang, Wei-Bin (1989) Theory of Economic Development - Nonlinearity, Instability and Non-equilibrium. UES 198. PhD thesis

Hansson, Pär (1989) Intra-Industry Trade: Measurements, Determinants and Growth - A study of Swedish Foreign Trade. UES 205. PhD thesis

Kriström, Bengt (1990) Valuing Environmental Benefits Using the Contingent Valuation Method: An Econometric Analysis. UES 219. PhD thesis

Aronsson, Thomas (1990) The Short-Run Supply of Roundwood under Nonlinear Income Taxation - Theory, Estimation Methods and Empirical Results Based on Swedish Data. UES 220. PhD thesis

Westin, Lars (1990) Vintage Models of Spatial Structural Change. UES 227. PhD thesis

Wikström, Magnus (1992) Four Papers on Wage Formation in a Unionized Economy. UES 287. PhD thesis

Westerlund, Olle (1993) Internal Migration in Sweden - The Role of Fiscal Variables and Labor Market Conditions. UES 293. PhLic thesis

Bergman, Mats A. (1993) Market Structure and Market Power. The Case of the Swedish Forest Sector. UES 296. PhD thesis

Johansson, Per (1993) Count Data Models - Estimator Performance and Applications. UES 315. PhD thesis

Roson, Roberto (1994) Transport Networks and the Spatial Economy - A General Equilibrium Analysis. UES 340. PhD thesis

Li, Chuan-Zhong (1994) Welfare Evaluations in Contingent Valuation - An Econometric Analysis. UES 341. PhD thesis

Østbye, Stein (1994) Regional Labour and Capital Subsidies - Theory and Evidence of the Impact on Employment under Wage Bargaining. UES 344. PhLic thesis

Westerlund, Olle (1995) Economic Influences on Migration in Sweden. UES 379. PhD thesis

Mortazavi, Reza (1995) Three Papers on the Economics of Recreation, Tourism and Property Rights. UES 396. PhLic thesis

Østbye, Stein (1995) Regional Labour and Capital Subsidies. UES 397. PhD thesis

Hussain-Shahid, Imdad (1996) Benefits of Transport Infrastructure Investments: A Spatial Computable General Equilibrium Approach. UES 409. PhD thesis

Eriksson, Maria (1996) Selektion till arbetsmarknadsutbildning. UES 410. PhLic thesis

Karlsson, Niklas (1996) Testing and Estimation in Labour Supply and Duration Models. UES 413. PhD thesis

Olsson, Christina (1996) Chernobyl Effects and Dental Insurance. UES 428. PhLic thesis

Vredin, Maria (1997) The African Elephant - Existence Value and Determinants of Willingness to Pay. UES 441. PhLic thesis

Eriksson, Maria (1997) To Choose or not to Choose: Choice and Choice Set Models. UES 443. PhD thesis

Widerstedt, Barbro (1997) Employer Change and Migration. Two Papers on Labour Mobility in Sweden. UES 444. PhLic thesis

Lundberg, Sofia (1997) The Economics of Child Auctions in 19th Century Sweden. UES 445. PhLic thesis

Lundberg, Johan (1997) Two Papers on Revenue and Expenditure Decisions in the Swedish Local Public Sector. UES 454. PhLic thesis

Widerstedt, Barbro (1998) Moving or Staying? Job Mobility as a Sorting Process. UES 464. PhD thesis

Bask, Mikael (1998) Essays on Exchange Rates: Deterministic Chaos and Technical Analysis. UES 465. PhD thesis

Löfgren, Curt (1998) Time to Study Students: Two Essays on Student Achievement and Study Effort. UES 466. PhLic thesis

Sjögren, Tomas (1998) Union Wage Setting in a Dynamic Economy. UES 480. PhD thesis

Mortazavi, Reza (1999) Essays on Economic Problems in Recreation, Tourism and Transportation. UES 483. PhD thesis

Rudholm, Niklas (1999) Two Essays on Generic Competition in the Swedish Pharmaceuticals Market. UES 485. PhLic thesis

Olsson, Christina (1999) Essays in the Economics of Dental Insurance and Dental Health. UES 494. PhD thesis

Marklund, Per-Olov (1999) Environmental Regulation and Firm Efficiency. UES 504. PhLic thesis

Berglund, Elisabet (1999) Regional Entry and Exit of Firms. UES 506. PhD thesis

Hellström, Jörgen (1999) Count Data Autoregression Modelling. UES 507. PhLic thesis

Nordström, Jonas (1999) Tourism and Travel: Accounts, Demand and Forecasts. UES 509. PhD thesis

Johansson Vredin, Maria (1999) Economics Without Markets. Four papers on the Contingent Valuation and Stated Preference Methods. UES 517. PhD thesis

Schei, Torbjørn (2000) Natural recreation resources: production and a diversity of interests related to the management of grouse as an outfield resource in Finnmark, Norway, in the Euro-Arctic Barents region. UES 523. PhLic thesis

Backlund, Kenneth (2000) Welfare Measurement, Externalities and Pigouvian Taxation in Dynamic Economies. UES 527. PhD thesis

Andersson, Linda (2000) Job Turnover, Productivity and International Trade. UES 530. PhLic thesis

Ylvinger, Svante (2000) Essays on Production Performance Assessment. UES 531. PhD thesis

Bergkvist, Erik (2001) Freight Transportation. Valuation of Time and Forecasting of Flows. UES 549. PhD thesis

Rudholm, Niklas (2001) The Swedish Pharmaceuticals Market - Essays on Entry, Competition and Antibiotic Resistance. UES 552. PhD thesis

Lundberg, Johan (2001) Local Government Expenditures and Regional Growth in Sweden. UES 554. PhD thesis

Lundberg, Sofia (2001) Going Once, Going Twice, SOLD! The Economics of Past and Present Public Procurement in Sweden. UES 557. PhD thesis

Eliasson, Kent (2001) University Enrollment and Geographical Mobility: The Case of Sweden. UES 558. PhLic thesis

Samakovlis, Eva (2001) Economics of Paper Recycling. Efficiency, policies, and substitution possibilities. UES 563. PhD thesis

Daunfeldt, Sven-Olov (2001) Essays on Intra-Household Allocation and Policy Regime Shifts. UES 570. PhD thesis

Hellström, Jörgen (2002) Count Data Modelling and Tourism Demand. UES 584. PhD thesis

Andersson, Linda (2002) Essays on Job Turnover, Productivity and State-Local Finance. UES 586. PhD thesis

Rashid, Saman (2002) Invandrarinkomster, förvärvsdeltagande och familj. UES 588. PhLic thesis

Hanes, Niklas (2003) Empirical Studies in Local Public Finance: Spillovers, Amalgamations, and Tactical Redistributions. UES 604. PhD thesis

Stenberg, Anders (2003) An Evaluation of the Adult Education Initiative Relative Labor Market Training. UES 609. PhD thesis

Stage, Jesper (2003) Mixing Oil and Water. Studies of the Namibian Economy. UES 611. PhD thesis

Marklund, Per-Olov (2004) Essays on Productive Efficiency, Shadow Prices , and Human Capital. UES 621. PhD thesis

Rashid, Saman (2004) Immigrants' Income and Family Migration. UES 625. PhD thesis

Sandberg, Krister (2004) Hedonic Prices, Economic Growth, and Spatial Dependence. UES 631. PhD thesis

Sjöström, Magnus (2004) Factor Demand and Market Power. UES 633. PhD thesis

Nilsson, William (2005) Equality of Opportunity, Heterogeneity and Poverty. UES 652. PhD thesis

Quoreshi, Shahiduzzaman (2005) Modelling High Frequency Financial Count Data. UES 656. Ph Lic thesis

Ankarhem, Mattias (2005) Bioenergy, Pollution, and Economic Growth. UES 661. PhD thesis

Quoreshi, Shahiduzzaman (2006) Time Series Modelling of High Frequency Stock Transaction Data. UES 675. PhD thesis