

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

I am grateful to many people for their cooperation since September 1, 2001. From the first day, Kurt has been very helpful by providing invaluable information. In the process of writing this thesis he has not only shown endless patience but also been very prompt in reading, correcting and improving the papers. He has always been accessible and generous with time. Moreover, he encouraged and recommended me to visit University of California, Berkeley. I am deeply indebted to him for his excellent supervision. Thank you very much Kurt!

I am grateful to Jörgen Hellström, Mikael Bask, Tomas Sjögren, Ola Simonson, Carl Lönnbark, Thomas Aronsson and Rainer Winkelmann (opponent on my licentiate thesis) for reading the paper(s) and making important comments. I am also grateful to a number of people, whose names have not been mentioned here, for their comments on the paper(s) when presented at a seminar or a conference. I would like to thank all course teachers for the PhD program and fellow colleagues. I am also thankful to Kjell-Göran, Marie Hammarstedt and Eva Cederblad for practical help.

Besides Kurt, I am indebted to specially three persons in my whole educational life. I would like to thank Raul Cano for his excellent thesis supervision in undergraduate studies and encouraging me to pursue a Ph.D. education. I am grateful to Gauranga Chandra Mandal for his caring and devotion in teaching while I was in class six. I am also grateful to my friend Moin (Toy) who explained me mathematics especially geometry in class ten. He had an amazing capacity of explaining mathematics that I still remember.

I would like to thank my parents, brother, sisters, maternal grand father, aunt Shahnaz and friends. Special thanks to my sons, Arvin (7 years) and Nivan (5 years), and my wife Sakira. Arvin has always been understanding and usually does whatever he is asked to do while Nivan questions everything. One evening Nivan questioned "pappa, where are you going". I answered "to the work". He became angry and said "ohaaaaaa you had already been to the work". Thanks to Sakira, I could do the second shifts. Without her help at home it would have been impossible to finish this thesis by this time. She deserves very special thanks for taking care of two kids alone while I was visiting Berkeley. Finally, I want to remember and to express my gratitude to uncle Parvez who submitted himself to and was murdered by the Pakistani army due to saving our lives during the war between Bangladesh and Pakistan.

Umeå, May 10, 2006

A.M.M. Shahiduzzaman Quoreshi

**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 Gouriéroux 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



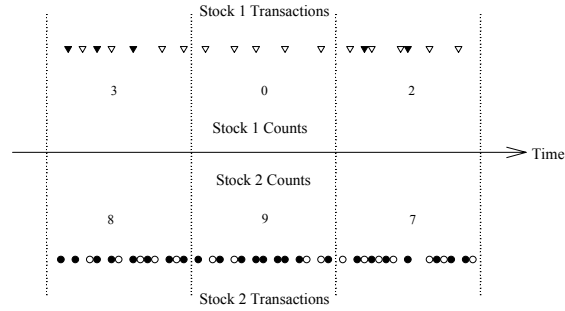


Figure 1: An illustration of how transactions data are generated. The black triangles and circles represent transactions for stock 1 and stock 2, respectively, while the white triangles and circles represent all other activities in an order book. The stock counts record the number of black triangles/circles falling into a time interval, i.e. falling between vertical lines.

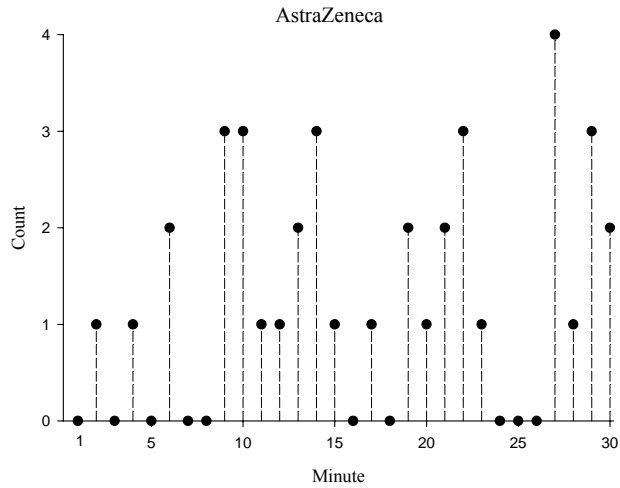


Figure 2: The number of transactions data over minute long intervals for 30 minutes of trading in AstraZeneca.

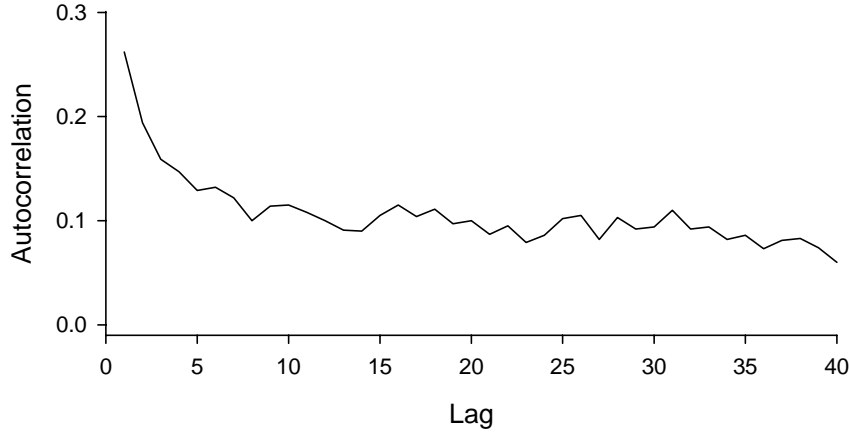


Figure 3: The autocorrelation function for AstraZeneca as an illustration of long range dependence of long memory.

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  $q$ ,  $\text{INMA}(q)$ , 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

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

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  $\text{INARMA}(p, q)$  model can be written

$$y_t - \alpha_1 \circ y_{t-1} - \dots - \alpha_p \circ y_{t-p} = u_t + \beta_1 \circ u_{t-1} + \dots + \beta_q \circ u_{t-q}. \quad (2a)$$

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

$$y_t = u_t + \beta_1 \circ u_{t-1} + \dots + \beta_q \circ u_{t-q} \quad (2b)$$

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.<sup>1</sup> 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., Gourioux, 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

---

<sup>1</sup>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  $\{y_t, t = 1, 2, \dots, T\}$  is an ARFIMA  $(0, d, 0)$  process if

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

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

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

and correspondingly

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

where  $\Gamma(n+1) = n!$  and  $i = 1, 2, \dots$ . The  $\Delta^d$  is needed for AR( $\infty$ ) and the  $\Delta^{-d}$  is needed for MA( $\infty$ ) 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. Combining 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 variables 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 transactions 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 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.

#### **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.,  $\lambda_j$  or  $\sigma_j$  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  $\delta$  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.