

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.

Umeå Economic Studies 673, 2006

*The financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. Special thanks to Kurt Brännäs, Jörgen Hellström and Carl Lönnbark for their comments on the earlier version(s) of this paper.

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-

timination 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 $y_1, \dots, y_T$ is introduced independently by McKenzie (1986) and Al-Osh and Alzaid (1987). The INARMA model can be written

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

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 $E(\varphi \circ v | v) = \varphi v$ and $V(\varphi \circ v | v) = \varphi(1 - \varphi)v$. 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 $\alpha_1, \dots, \alpha_p$ and $\beta_1, \dots, \beta_q$ 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 $\{y_t, t = 1, 2, \dots, T\}$ is an ARFIMA $(0, d, 0)$ model if

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

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

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

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 Δ^d is needed for $\text{AR}(\infty)$ and the Δ^{-d} is needed for $\text{MA}(\infty)$ 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 $(0, d, 0)$ 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

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

where $d_i = \Gamma(i + d) / [\Gamma(i + 1)\Gamma(d)]$, $i \geq 1$ 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

$$(1 - L^\circ)^d y_t = u_t \quad (7)$$

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

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

In (8) $\alpha(L^\circ) = 1 - \alpha_1 \circ L - \alpha_2 \circ L^2 - \dots - \alpha_p \circ L^p$ and $\beta(L^\circ) = 1 + \beta_1 \circ L + \beta_2 \circ L^2 + \dots + \beta_q \circ L^q$, 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)

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

$$\begin{aligned} V(y_t | Y_{t-1}) &= \sum_{j=1}^p \alpha_j (1 - \alpha_j) y_{t-j} + \sigma^2 + \sum_{j=1}^q \beta_j (1 - \beta_j) u_{t-j} \\ &\quad + \sum_{j=0}^q \beta_j \sum_{i=1}^{\infty} d_i (1 - \beta_j d_i) u_{t-i-j}. \end{aligned} \quad (9b)$$

Note that the moments are conditioned on only the previous observations, $Y_{t-1} = (y_{t-1}, y_{t-2}, \dots)$. 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

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

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

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

$$\begin{aligned} V(y_t) = & \lambda \sum_{j=0}^q \beta_j \left(\sum_{i=1}^{\infty} d_i(1 - d_i \beta_j) + \sum_{j=0}^q (1 - \beta_j) \right) \\ & + \sigma^2 D^2 \sum_{j=0}^q \beta_j^2 \end{aligned} \quad (11b)$$

with $D = 1 + \sum_{i=1}^{\infty} d_i$, $D^2 = 1 + \sum_{i=1}^{\infty} d_i^2$ and $d_i = \Gamma(i+d)/[\Gamma(i+1)\Gamma(d)]$, $i \geq 1$. 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

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

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

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

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

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

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

with $d_i = \Gamma(i + d) / [\Gamma(i + 1)\Gamma(d)]$. 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

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

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

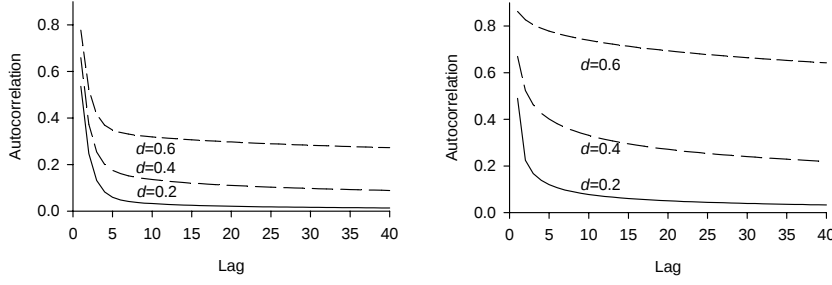


Figure 1: The autocorrelation functions for INARFIMA(1, d , 0) (left figure) with $\alpha = 0.4$, $\lambda = 2$ and $\sigma^2 = 50$ and INARFIMA(0, d , 1) (right figure) with $\beta = 0.4$ and the same values for the λ and σ^2 .

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

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

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

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

where $d_i = \Gamma(i + d) / [\Gamma(i + 1)\Gamma(d)]$, $i \geq 1$ and the property $\varphi_1 \circ (\varphi_2 \circ v) \stackrel{d}{=} (\varphi_1\varphi_2) \circ v$ 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, δ , 0). We say that $\{y_t, t = 1, 2, \dots, T\}$ is an INARFIMA (p, δ, q) model when

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

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

$$E(y_t | Y_{t-1}) = \sum_{j=1}^p \alpha_j y_{t-j} + \lambda \theta(\delta) + \theta(\delta) \sum_{j=1}^q \beta_j u_{t-j} \quad (19a)$$

$$+ \theta(\delta) \sum_{j=0}^q \beta_j \sum_{i=1}^{\infty} \delta_i u_{t-i-j}$$

$$V(y_t | Y_{t-1}) = \sum_{j=1}^p \alpha_j (1 - \alpha_j) y_{t-j} + \theta^2(\delta) \sigma^2 + \sum_{j=1}^q \theta(\delta) \beta_j (1 - \theta(\delta) \beta_j) u_{t-j} \\ + \theta(\delta) (1 - \theta(\delta)) \lambda \quad (19b)$$

$$+ \sum_{j=0}^q \beta_j \sum_{i=1}^{\infty} \theta(\delta) \delta_i (1 - \theta(\delta) \beta_j \delta_i) u_{t-i-j}.$$

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, δ , 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 $\text{INMA}(\infty)$ representation of the $\text{INARFIMA}(0, d, 0)$ or the corresponding part of the $\text{INARFIMA}(p, d, q)$ model in generating the data. The first 500 observations are discarded from 10500 observations in order to avoid start up effects.

The identification of $\text{INARFIMA}(p, d, q)$ models is not straightforward. The autocorrelation function for $\text{INARFIMA}(1, d, 0)$ and $\text{INARFIMA}(1, d, 1)$ look almost alike while the partial-autocorrelation functions are different. The autocorrelation and the partial-autocorrelation functions of the $\text{INARFIMA}(0, \delta, 0)$ are quite similar to those of an $\text{INARFIMA}(0, d, 1)$. In general, an $\text{INARFIMA}(p, d, q)$ model and an $\text{INARFIMA}(p, \delta, q)$ model have more slowly decaying autocorrelation functions than an $\text{INARMA}(p, q)$. 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

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

we can choose the lag length of the model. Here, $\hat{\sigma}^2$ is the variance estimate based on the residuals from the $\text{INARFIMA}(p, d, q)$ 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 $\text{INARFIMA}(p, \delta, q)$ 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

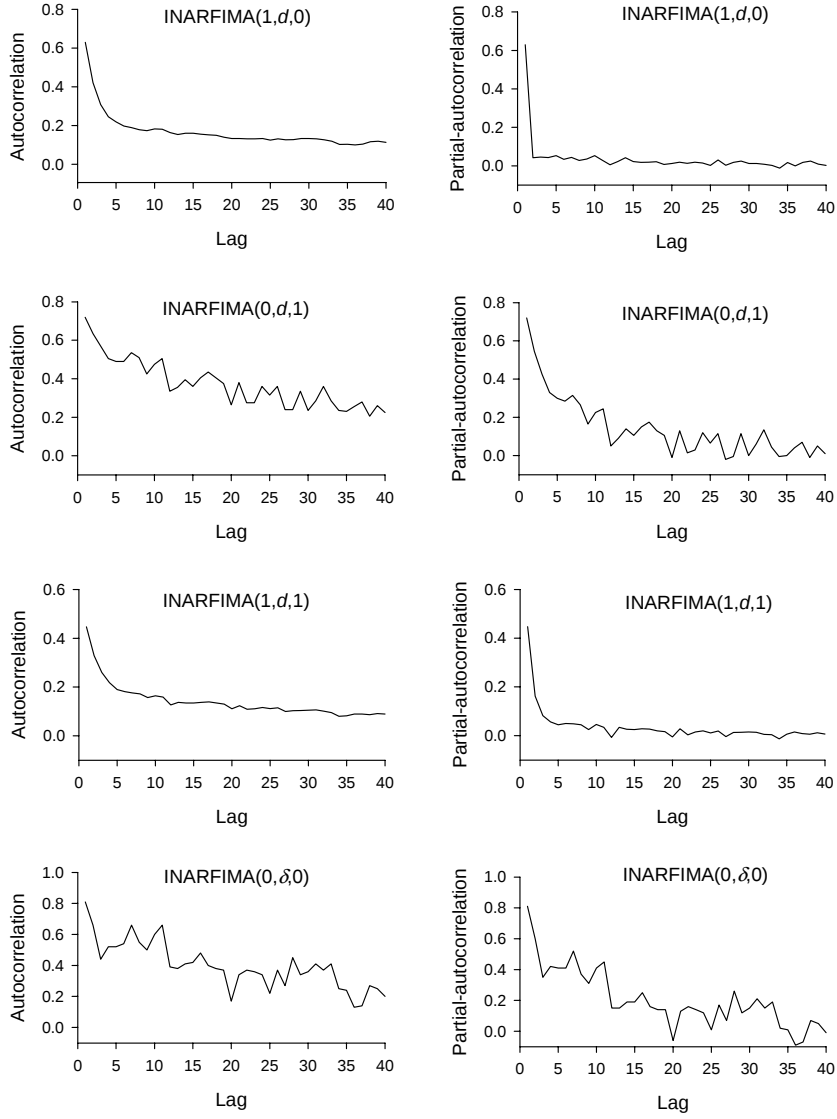


Figure 2: The autocorrelation and partial-autocorrelation functions for the different INARFIMA models with $\alpha = 0.5$, $d = 0.4$, $\theta(\delta) = 0.5$ and $\beta = 0.5$ when applicable. The autocorrelations and partial-autocorrelations for $\text{INARFIMA}(0,d,1)$ and $\text{INARFIMA}(0,\delta,0)$ 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

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

and the criterion function $S_1 = \sum_{t=m+1}^T e_{1t}^2$ is minimized with respect to the unknown parameters, i.e. $\psi = (\lambda, \alpha', \beta', \theta(\delta) \text{ and } \delta)$. Using a finite maximum lag m in (20) instead of infinite lags may have biasing effects. Due to the omitted variables, i.e. $u_{t-m-1}, \dots, u_{t-\infty}$ 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)

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

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

$$\sigma^2 = \theta^{-2}(\hat{\delta}) (T - m)^{-1} \sum_{t=m+1}^T \left[\hat{e}_{1t}^2 - \sum_{j=0}^p \hat{\alpha}_j (1 - \hat{\alpha}_j) y_{t-j} - \theta(\hat{\delta}) (1 - \theta(\hat{\delta})) \hat{\lambda} \right. \\ \left. - \sum_{j=1}^q \theta(\hat{\delta}) \hat{\beta}_j (1 - \theta(\hat{\delta}) \hat{\beta}_j) u_{t-j} - \sum_{j=0}^q \hat{\beta}_j \sum_{i=1}^{\infty} \theta(\hat{\delta}) \hat{\delta}_i (1 - \theta(\hat{\delta}) \hat{\beta}_j \hat{\delta}_i) u_{t-i-j} \right]$$

Finally, the FGLS estimator minimizes

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

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

$$\text{Cov}(\hat{\psi}_{CLS}) = \left(\sum_{t=m+1}^T \frac{\partial e_{1t}}{\partial \psi} \frac{\partial e_{1t}}{\partial \psi'} \right)^{-1}$$

$$\text{Cov}(\hat{\psi}_{FGLS}) = \left(\sum_{t=m+1}^T \hat{V}^{-1}(y_t | Y_{t-1}) \frac{\partial e_{1t}}{\partial \psi} \frac{\partial e_{1t}}{\partial \psi'} \right)^{-1}.$$

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

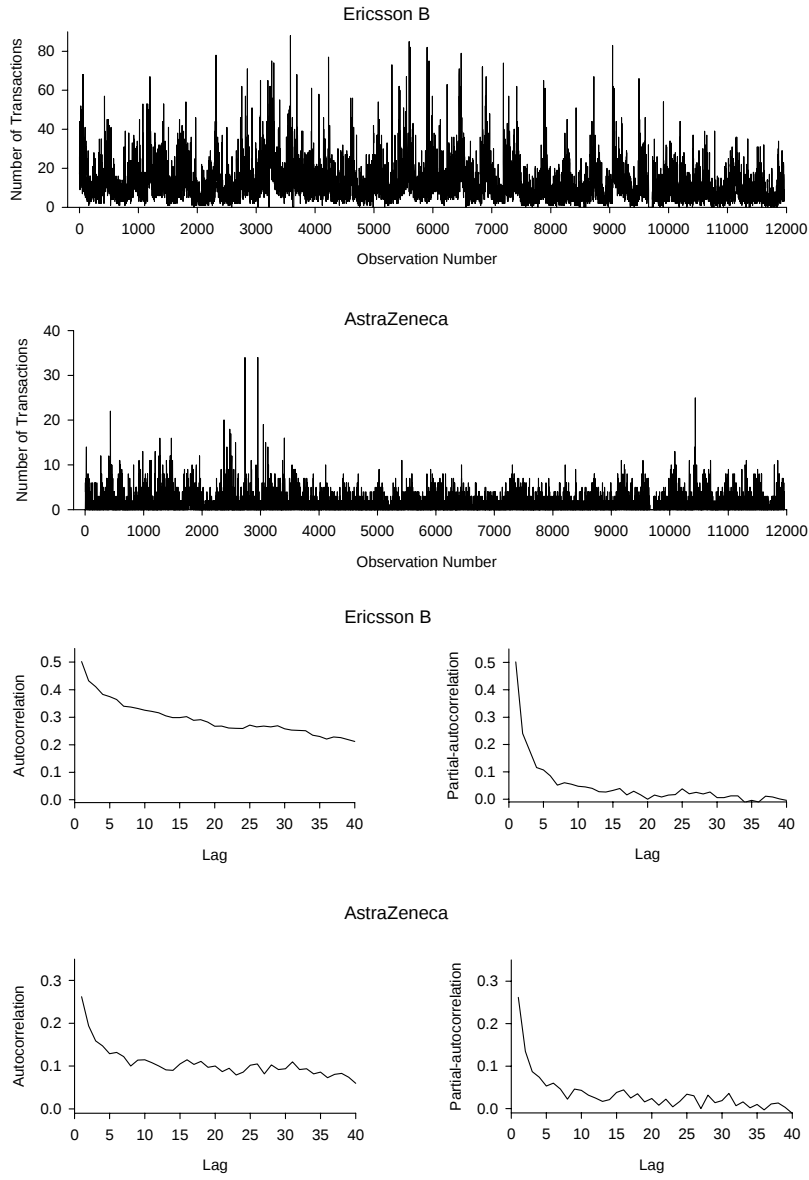


Figure 3: Time series plots for the Ericsson B (mean 11.73 variance 84.86, maximum 88) and AstraZeneca series (mean 1.33 variance 3.75, maximum 34) and their autocorrelation and partial-autocorrelation functions.

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, δ , 0) becomes the second best for both series and the estimated parameters are positive. The INARFIMA(0, δ , 0) is the most parsimonious model in terms of number of parameters.¹

The empirical results for INARFIMA(0, δ , 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

¹We 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

²As 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 $\theta(\delta) \sum_{i=0}^k \delta_i/w \geq 0.5$.

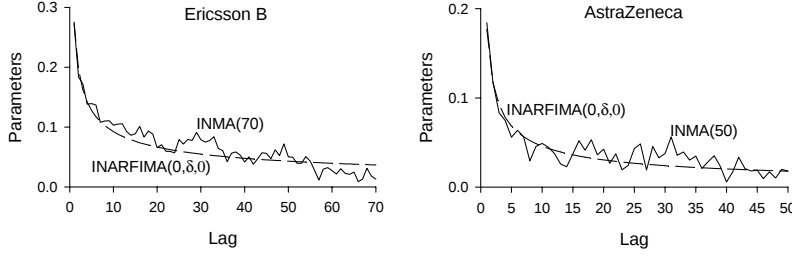


Figure 4: The parameters estimated with INARFIMA($0, \delta, 0$) (dashed line) and truncated INMA(∞) models for Ericsson B and AstraZeneca.

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 Qureshi, 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.