The Number of Shareholders —
Time Series Modelling and Some Empirical Results

Kurt Brännäs
Department of Economics
Umeå School of Business and Economics, Umeå University
e-mail: kurt.brannas@econ.umu.se

Umeå Economic Studies 855, January 2013

Abstract

The paper discusses some model related issues for time series of the number of shareholders in a stock. The point of departure is an integer-valued autoregressive model of order one. Empirical results are presented for some frequently traded stocks on the Finnish and Swedish stock markets. In these stock markets public records of the number of owners are reported monthly (Finland) and quarterly (Sweden, and initially at biannual) intervals. The aggregate records are useful for, e.g., indirectly estimating average holding times, which are found to vary but to mostly exceed one year.

Key Words: Integer-valued, autoregression, nonlinear, holding time, stock, share

JEL: C22, C25, G10, G11


1 Introduction

This short paper models time series of the number of shareholders (owners) in large Finnish and Swedish stocks. The numbers of owners may beyond being interesting in their own right be viewed as integral parts of the time series of number of traded shares, trading volume, etc. While market data allows for intra-day studies, the number of owners is typically provided publicly by securities depositories and some companies at much lower frequencies.

The integer-valued autoregressive model of order one (INAR(1)) with constant parameters (McKenzie, 1985; Al-Osh and Alzaid, 1987) is taken as the basic modelling approach. This model catches the main aspects of the integer-valued, number of owners variable in a relatively simple way. In contrast to conventional time series modelling the adopted approach offers interesting interpretations in terms of the model’s parameters. We may obtain measures such as the mean holding time from the parameters, even though sampling individual transactions in the records of securities depositories may be a better alternative (cf. Bøhrens et al., 2006, for a study based on individual transactions). The available low sampling frequency or sparse but equidistant reporting dates brings along the problem of short holdings falling between measurement days. An approach to handling this problem is discussed and used empirically.

Empirically, we focus on the stocks of eight large stock market companies, the four largest (by the end of 2005) in each of Finland and Sweden, in terms of their numbers of shareholders. Some of the companies are traded not only in their home markets, but also elsewhere. The numbers of recorded owners in the Finnish and Swedish depositories also contain registered (but not all) foreign owners. The consequences of these and a few other extensions of the INAR(1) are also analyzed.

2 The Basic Model and Extensions

The number of owners/shareholders of a stock at some time point \( t \) is a count data or integer-valued variable \( y_t \) taking values \( 0, 1, 2, \ldots \). The \( y_t \) is composed of the number of new shareholders, \( \epsilon_t \), that buy into the stock in the period \( (t-1, t] \) and of those owners at \( t-1 \) that remain owners at time \( t \); these are denoted by \( \alpha \circ y_{t-1} \). We write the model as

\[
y_t = \alpha \circ y_{t-1} + \epsilon_t.
\]

Here, the symbol \( \circ \) indicates a binomial thinning operation and it replaces conventional multiplication. The thinning is defined through \( \alpha \circ y = \sum_{i=1}^{y} u_i \), where \( \{u_i\} \) is an independent sequence of random \( 0 - 1 \) variables with \( \Pr(u_i = 1) = \alpha \). Importantly, thinning gives integers in the \([0, y]\) range. The \( \alpha \) is the probability of remaining an owner to the
next time period, i.e. $a$ is a survival probability. Assuming that $y$ and $\{u_i\}$ are independent gives that $E(a \circ y) = E_y [E(a \circ y)|y] = aE(y)$ and $V(a \circ y) = a^2V(y) + a(1-a)E(y)$. For the $\{\varepsilon_i\}$ sequence we assume independence and that $E(\varepsilon_i) = \lambda$ and $V(\varepsilon_i) = \sigma^2$. The $\lambda$ corresponds to the average number of new owners over a time period. For the important special case of the Poisson distribution we have that $\sigma^2 = \lambda$. The model was first introduced by McKenzie (1985) and independently by Al-Osh and Alzaid (1987). It has later been discussed and generalized in a number of studies, cf. the survey of McKenzie (2003).

The basic model has the moment properties

\[
E(y_t|F_{t-1}) = ay_{t-1} + \lambda \\
E(y_t) = \lambda/(1-a) \\
V(y_t|F_{t-1}) = a(1-a)y_{t-1} + \sigma^2 \\
V(y_t) = \left[a(1-a)E(y_{t-1}) + \sigma^2\right]/(1-a^2),
\]

where $F_{t-1}$ is the available information up through time $t-1$. The expected number of owners is $\lambda/(1-a)$. The probability of remaining an owner over $k$ periods is $a^k$ and the expected duration of ownership (mean holding time at the $t$-scale) is $MH = 1/(1-a)$. The median duration falls into the interval where the survival function is 0.5. The model also has a conditional heteroskedasticity property as is evident from $V(y_t|F_{t-1})$.

Given that we write the model as $y_t = a \circ y_{t-1} + \gamma \circ \nu_t$, where $\nu_t$ may be viewed as potential entrants, it appears reasonable to consider, i.a., dependence between entry and exit mechanisms. This corresponds to dependence between the thinning operations and can be demonstrated to influence only the second order moment properties in most cases (Brännäs and Hellström, 2001). Hence, considering only first order moment properties is a way of assuring robustness to incorrect assumptions related to the thinning operations. On the other hand, if such assumptions could be justified they would pave the way for even more efficient estimation.

An alternative but more complex model for the number of owners is an integer-valued moving average (INMA) model with a special pattern of dependence between thinning operations, cf. Brännäs and Hall (2001). The INAR model can be viewed as an approximative and simpler dual representation of the INMA model. Obviously, the more parsimoniously parameterized or short lagged INAR(1) model has an empirical advantage with respect to the commonly short time series at hand.

Next we consider model consequences of a number of issues that may arise in this and related areas.
2.1 Low Sampling Frequency

Given the widely spread algorithmic trading at an almost continuous time scale and the availability of data at a monthly or even sparser time scale the question of appropriate time scale for the model in (2) is an important but difficult one. We have chosen to view the model in (2) as generating data on a daily basis. The available data is monthly or even quarterly. To account for a low sampling frequency of \( s = 21 \) days (for a trading month), \( 63 \) days (trading quarter) or even biannual data we consider the following modelling strategy based on temporal aggregation.

By successive substitution over \( s \) days from \( t \) to \( t + s \) we get

\[
y_{t+s} = \alpha^s \circ y_t + \sum_{i=1}^{s} \alpha^{s-i} \circ \varepsilon_{t+i},
\]

where equality is in distribution. In this context there are only observations at times that are separated by \( s \).

With \( \alpha < 0.9 \) the first term in (3) can in practice be disregarded, while for an \( \alpha \) closer to one the term remains important, at least, in the monthly or \( s = 21 \) trading days case.

The conditional expectation of \( y_{t+s} \) in (3) conditional on information up through the previous observation time, i.e. \( t \), takes the form

\[
E(y_{t+s} | \mathcal{F}_t) = \alpha^s y_t + \lambda \sum_{i=1}^{s} \alpha^{s-i} = \alpha^s y_t + \frac{\lambda (1 - \alpha^s)}{1 - \alpha},
\]

which indicates that the model is still of the INAR(1) type (cf. Brewer, 1973). Note that the parameters in the conditional expectation (4) differs from those of (2). In particular, the survival probability is, not surprisingly, smaller at the sparser time scale, and the mean entry is larger. The unconditional mean \( E(y_{t+s}) \) remains unchanged, while the conditional variance is less dependent on past \( y \) than the one in (2), since \( \alpha^s (1 - \alpha^s) < \alpha (1 - \alpha) \).

2.2 Two Owner Types

Consider a case of two types of owners with numbers \( y_{1t} \) and \( y_{2t} \) that are unobservable, while their sum is observed. Both follow INAR(1) models but with different parameters \( \alpha_i, \lambda_i, i = 1,2 \). We may consider \( \alpha_2 \) to be smaller than \( \alpha_1 \) and then to correspond to owners having short holding times. Since \( y_{t+s} = y_{1t+s} + y_{2t+s} \) we have that \( E(y_{t+s} | \mathcal{F}_t) \) arises as the sum of expressions of the type in (4), so that we can get, say, the conditional expectation

\[
E(y_{t+s} | \mathcal{F}_t) = \alpha_1^s y_t + (\alpha_2^s - \alpha_1^s) y_{2t} + \frac{\lambda_1 (1 - \alpha_1^s)}{1 - \alpha_1} + \frac{\lambda_2 (1 - \alpha_2^s)}{1 - \alpha_2}.
\]
When \( \alpha_2 \) is small \( \alpha_2^* \approx 0 \) leaving us with a second term \(-\alpha_1^* y_{2t}\) in (5). As we cannot observe \( y_{2t} \) we may replace it with its expected value \( \lambda_2/(1 - \alpha_2) \) to get an approximative expression

\[
E(y_{t+s}|\mathcal{F}_t) \approx \alpha_1^* y_t + \frac{\lambda_1(1 - \alpha_1^*)}{1 - \alpha_1} + \frac{\lambda_2(1 - \alpha_1^*)}{1 - \alpha_2}.
\] (6)

Even if \( \alpha_1 \) can be uniquely estimated using (6) the other parameters cannot be separately estimated without further information. Importantly, (6) then suggests that it may be empirically difficult to catch the survival probability \( \alpha_2 \) by this modelling approach, at least, with low sampling frequencies for the time series.

### 2.3 Two Depositories

Shareholders in a stock may be registered in either of two depositories, domestically \((D)\) or abroad \((A)\). Over time registered ownership may move between the two depositories according to

\[
y_{Dt} = \alpha_D \circ y_{D,t-1} + \beta_A \circ y_{A,t-1} + \epsilon_{Dt} \quad (7)
y_{At} = \alpha_A \circ y_{A,t-1} + \beta_D \circ y_{D,t-1} + \epsilon_{At}, \quad (8)
\]

where we expect the \( \beta_A \) and \( \beta_D \) migration probabilities to be small, and much smaller than the \( \alpha_D \) and \( \alpha_A \) probabilities.

If we only have access to the domestic series \( \{y_{Dt}\} \) we may substitute from (8) to get rid of the \( y_{A,t-1} \) part in (7). We use \( y_{it} = E(y_{it}|\mathcal{F}_{i-1}) + \zeta_{it}, i = A, D, \) where \( E(\zeta_{it}|\mathcal{F}_{i-1}) = 0 \) for both \( i \) to get

\[
y_{At} = \frac{\beta_D}{1 - \alpha_A L} y_{D,t-1} + \frac{\lambda_A}{1 - \alpha_A} + \frac{\zeta_{At}}{1 - \alpha_A L}.
\]

Inserting this into the conditional representation of (7) then gives

\[
E(y_{Dt}|\mathcal{F}_{t-1}) = \alpha_D y_{D,t-1} + \beta_A \beta_D y_{D,t-2} + \alpha_A \beta_D y_{D,t-3} + \alpha_2^2 \beta_A \beta_D y_{D,t-4} \ldots + \lambda_D + \frac{\beta_A \lambda_A}{1 - \alpha_A}. \quad (9)
\]

With a small \( \beta_A \beta_D \) the autoregression can in practice be expected to be of order one. When this is the case only \( \alpha_D \) can be separately estimated by, e.g., a conditional least squares estimator.

Systematically observing only every \( s \) observation of a daily series leads to an INAR with the same maximum lag as in (9) (Brewer, 1973). Given the complexity of the general expression it may appear reasonable to use (9) directly and then to interpret the parameters in terms of the \( s \) time scale.
2.4 Time Dependence

In practice both \( \lambda \) and \( \alpha \) can be expected to vary over time. Liquidity and the share price are examples of variables that can be expected to influence both parameters. Over longer time horizons parameters may also change due to the common use of share expansions in take-overs or due to repurchases of shares.

It is straightforward to introduce time dependence in the parameters of (4) as long as they remain constant within the month for \( s = 21 \), etc. We may write

\[
E(y_{t+s}|\mathcal{F}_t) = \alpha_t y_t + \frac{\lambda_t(1 - \alpha_t^s)}{1 - \alpha_t},
\]

(10)

Brännäs (1995) suggested that a logistic distribution function \( \alpha_t = 1/[1 + \exp(x_t \beta)] \) and an exponential function \( \lambda_t = \exp(z_t \gamma) \) provide convenient parameterizations where use is made of exogenously determined variable vectors \( x_t \) and \( z_t \) and the unknown parameter vectors \( \beta \) and \( \gamma \). Note that the unconditional mean in (2) suggests that incorporating the same variable in \( x_t \) and \( z_t \) will likely lead to strong negative correlation between the associated parameters.

2.5 Unit Root

Given the count data interpretation of the \( \{y_t\} \) sequence and the assumption \( \varepsilon_t \geq 0 \), a unit root or \( \alpha \equiv 1 \) gives \( y_t - y_{t-1} = \varepsilon_t \geq 0 \). Hence, a unit root implies that an owner series can remain constant or be growing, but it cannot decrease. Importantly, this corresponds to an infinite holding time. Therefore, a unit root hypothesis may be rejected on logical grounds in most cases. Note that there is some empirical evidence evidence of overestimating \( \alpha \) in AR(1) models when it is de facto time-varying rather than time invariant.

3 Estimation

From Al-Osh and Alzaid (1987) and followers several estimators for the INAR(1) model have been studied and compared. Among the simpler ones for the current context of very large numbers of shareholders the Yule-Walker and conditional least squares (CLS) estimators are both simple to use and have been found to perform well.

The constant parameter model in (4) has for monthly data, \( t = 1, \ldots, T \), the one-step-ahead predictor \( E(y_t|\mathcal{F}_{t-1}) = \alpha^s y_{t-1} + \lambda(1 - \alpha^s)/(1 - \alpha) \), and the prediction error is \( e_t = y_t - E(y_t|\mathcal{F}_{t-1}) \). For the other model specifications the appropriate conditional expectation should be used. The prediction errors have zero means, but conditional variances vary with, e.g., the dependence between and within the binomial thinning operations (cf. Brännäs and Hellström, 2001). For this reason estimation will here
not explicitly build on a conditional variance specification as would be required for a conditional weighed least squares estimator. Additionally, we refrain from making distributional assumptions, so that maximum likelihood estimation is precluded.

A linear CLS estimator is based on only a first conditional moment assumption and gives estimates \( a \) of \( \alpha \) and \( b \) of \( \lambda(1 - \alpha)/(1 - \alpha) \). The underlying \( \hat{\alpha} \) and \( \hat{\lambda} \) can be obtained from these expressions as \( \hat{\alpha} = a^{1/s} \) and \( \hat{\lambda} = b(1 - a^{1/s})/(1 - a) \). Standard errors can be obtained using the Delta method, e.g., for \( \hat{\alpha} \) the standard error is \( a^{(1/s)-1}V^{1/2}(a)/s \). The CLS estimator should have an accompanying robust covariance matrix for the \( a \) and \( b \) estimators to account for conditional heteroskedasticity of unspecified form. A nonlinear CLS estimator can also be used to obtain the \( \hat{\alpha} \) and \( \hat{\lambda} \) estimates directly. Introducing exogenous variables through the \( \alpha \) and \( \lambda \) parameters creates no substantially new and difficult obstacles for the nonlinear CLS estimator.

For the Swedish stocks the available time series is initially measured on a biannual scale and later on a quarterly one. The LS criterion function \( S \) can be decomposed accordingly, i.e. as \( S = S_1 + S_2 \), where \( S_1 \) is for the biannual data. The prediction errors are then based on \( s_1 = 126 \) and later \( s_2 = 63 \) days, respectively.

A Yule-Walker estimator of \( \alpha \) and \( \lambda \) is a moment estimator based on \( \hat{\gamma} = \lambda/(1 - \alpha) \) and \( r_1 = \alpha^{s} \), where \( \hat{\gamma} \) is the sample average and \( r_1 \) is the estimator of the lag one autocorrelation. Hence, this estimator is based on unconditional moments while CLS is based on conditional ones.

4 Empirical Findings

We present empirical results for the four largest and registered companies in terms of their numbers of owners (December 2005) of the Helsinki (Finland) and Stockholm (Sweden) stock markets. For Finland there are monthly time series 2000:12–2012:10 (\( T = 143 \)) and for Sweden (expressed in months) biannual observations 1999:12–2005:12 followed by quarterly observations 2006:3–2012:9 (\( T = 40 \)).\(^1\) Note that in each case the number of owners corresponds to owners registered domestically. Ownership of American Depository shares may therefore influence the interpretation of \( \hat{\lambda} \) (see, e.g., the annual reports 2008 of Nokia and Ericsson).

The Finnish time series are exhibited in Figure 1. The Elisa series contains a few jumps that correspond to takeovers when the Elisa stock was involved. Nokia has an expansion phase up to about the 40th month (2004:3) followed by a long recession before recovering towards the latter part of the series. Note that Nokia to some 90 percent is foreign-owned. The UPM series shows some fluctuation, while for Sampo the predominant impression is one of growth in the number of shareholders.

\(^1\)TeliaSonera has the first observation at 2000:6.
Figure 1: The Finnish shareholder time series vs observation month, 2000:12–2012:10.

Figure 2: The Swedish shareholder time series vs observation number, 1999:12-2005:12 (biannually), 2006:3-2012:9 (quarterly).
Table 1: Yule-Walker (Finland) and CLS (Sweden) estimates for constant parameter specifications. The mean holding time (MH) is in months.

<table>
<thead>
<tr>
<th>Stock</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\lambda}$</th>
<th>$T$</th>
<th>MH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Finland</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elisa</td>
<td>0.9994</td>
<td>3028.1</td>
<td>143</td>
<td>83</td>
</tr>
<tr>
<td>Nokia</td>
<td>0.9985</td>
<td>4960.3</td>
<td>143</td>
<td>31</td>
</tr>
<tr>
<td>Sampo</td>
<td>0.9990</td>
<td>1274.4</td>
<td>143</td>
<td>50</td>
</tr>
<tr>
<td>UPM</td>
<td>0.9991</td>
<td>1390.1</td>
<td>143</td>
<td>53</td>
</tr>
<tr>
<td><strong>Sweden</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ericsson</td>
<td>0.9982</td>
<td>1559.7</td>
<td>40</td>
<td>26</td>
</tr>
<tr>
<td>SEB</td>
<td>0.9992</td>
<td>219.6</td>
<td>40</td>
<td>59</td>
</tr>
<tr>
<td>SwedBank</td>
<td>0.9993</td>
<td>208.2</td>
<td>40</td>
<td>68</td>
</tr>
<tr>
<td>TeliaSonera</td>
<td>0.9881</td>
<td>8855.4</td>
<td>39</td>
<td>4</td>
</tr>
</tbody>
</table>

Note: The Yule-Walker estimator is based on the mean and lag one autocorrelation.

For each Finnish stock we consider two specifications, one in which the parameters are time invariant and one in which we have a time dependent $\alpha_t$ parameter of the logistic type and $\lambda_t$ as an exponential function. The time-varying specifications may indicate the role of time variation on, e.g., the MH. Here, $\alpha_t$ is set to depend on the price change in the previous month, i.e. $\Delta p_{t-1} = p_{t-1} - p_{t-2}$ and on the number of owners at the end of previous month $(y_{t-1})$ (divided by 10000). The $\lambda_t$ is a function of $\Delta y_{t-1}$. For Elisa we also include a dummy variable ($d_t$) taking value one for the months of larger changes in the stock in the time-varying $\lambda_t$.

For Nokia we find that the Yule Walker estimator gives $\hat{\alpha} = 0.9985$ and an implied holding time of about 31 trading months. The CLS estimator gives a longer MH. The fit is about $R^2 = 0.97$ but there is remaining serial correlation of the AR(1) type. Therefore, in an INAR(2) model the lag two parameter not surprisingly comes out with a negative sign. This also speaks against, e.g., a two depository and an INAR(2) interpretation. After some trial and error we give an estimated time dependent parameter model as $\hat{y}_t = y_{t-1}/(1 + \exp(-5.0 + 0.022\Delta p_{t-1} - 0.078y_{t-1}/10000)) + \exp(5.82 + 0.418\Delta y_{t-1})$, where all parameter estimates are significant. A positive price change is estimated to have a reducing effect on $\alpha_t$, while an increase in the number of owners has an enhancing effect. The implied holding times vary between 1.2 and 4.2 trading years, with an average of about 2.0 years. The longest holding times are noted for the final part of the series, where the share price is lowest. Hence, introducing explanatory variables seems to
reduce the mean holding time. Moreover, there is no remaining serial correlation (with the exception of the UPM model) thanks to the $\Delta y_{t-1}$ included in the $\lambda_t$-part.

Table 1 gives the Yule-Walker estimates for all Finnish and CLS estimates for all Swedish stock series. The mean holding times are reported in trading months. In general, the models fit the Finnish series very well, but there are serial correlation problems. The $\hat{\alpha}$ estimates are very close to one, and even closer for the CLS estimator.

Table 2 contains estimates for the time varying $\alpha_t$ and a $\lambda_t$ model specifications for the Finnish stocks. The price effect appears positive and the lagged $y_{t-1}$ effect comes out as significantly negative. The Elisa specification produces a holding time estimate that appears on the long side, while the other holding times appear more realistic. Average entry is affected positively by an increase in the lagged change in ownership. Only for UPM is there significant remaining serial correlation.

The Swedish stocks are displayed in Figure 2. In contrast to the Finnish stocks there is a negative trend, at least, in the latter parts of the series. Ericsson peaks at about the same time as Nokia. The Swedish series are very short and when estimated individually there is less room to enlarge by using time dependent parameters.

The constant parameter estimation results are given in Table 1. The implied mean holding times range between 0.3 and 5.7 years. The MH estimate of TeliaSonera is surprisingly small and its asymptotic 95 percent confidence interval is 0 - 8.2 months. The downward trend of the series is overall steeper than for any of the other series, which may be the reason for the small $\hat{\alpha}$ estimate. The confidence interval for Ericsson

<table>
<thead>
<tr>
<th>Stock</th>
<th>$\alpha_t$</th>
<th>$\lambda_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Const $\Delta p_{t-1} y_{t-1}^*$</td>
<td>Const $\Delta y_{t-1} d_t$</td>
</tr>
<tr>
<td>Elisa</td>
<td>-12.67 (0.01)</td>
<td>0.047 (0.009)</td>
</tr>
<tr>
<td>Nokia</td>
<td>-5.00 (0.36)</td>
<td>0.022 (0.011)</td>
</tr>
<tr>
<td>Sampo</td>
<td>-7.73 (0.05)</td>
<td>-0.009 (0.000)</td>
</tr>
<tr>
<td>UPM</td>
<td>-5.45 (0.01)</td>
<td>0.056 (0.024)</td>
</tr>
</tbody>
</table>

Note: $y_{t-1}^* = y_{t-1}/10000.$
is 19.6 - 32.4 trading months.

5 Concluding Remarks

The only other study from a Nordic country that we are aware of is Bøhrens et al. (2006), whose main interest was in corporate governance issues. They studied Norwegian registered firms and found, using micro-data 1989-1999, that median HM times were in the region 1-2 years across firms, excluding financial ones. Their sample included both small and large firms beyond being based on a different time period. Their results, e.g., indicate that foreign owners have shorter mean holding times and we find shorter ones for Nokia, Ericsson as well as TeliaSonera that have internationally spread ownership. It is obviously possible to use the current modeling approach to studying corporate governance related issues across both time and different stocks.

The empirical results indicate that mean holding can be expected to vary across time. The suggested specifications are, admittedly, quite ad hoc and some more serious thinking is required. Alternatively, some flexible time polynomial could be adopted to catch main features across time. Even so, the simple INAR(1) model provides a very good fit to the data, and there is not much remaining to explain.
References


