

# Discretized Time and Conditional Duration Modelling for Stock Transaction Data\*

Kurt Brännäs and Ola Simonsen

Department of Economics & USBE, Umeå University

SE-901 87 Umeå, Sweden

## Abstract

The paper considers conditional duration models in which durations are in continuous time but measured in grouped or discretized form. This feature of recorded durations in combination with a frequently traded stock is expected to negatively influence the performance of conventional estimators. A few estimators that account for the discreteness are discussed and compared in a Monte Carlo experiment. An EM-algorithm accounting for the discrete data performs better than those which do not. Empirical results are reported for trading durations in Ericsson B at Stockholmsbörsen for a three-week period of July 2002. The incorporation of level variables for past trading is rejected in favour of change variables. This enables an interpretation in terms of news effects. No evidence of asymmetric responses to news about prices and spreads is found.

**Key Words:** Grouped data, Maximum likelihood, EM-algorithm, Estimation, Finance, News.

**JEL Classification:** C12, C22, C41, C51, G12, G14.

*Umeå Economic Studies 610, May 2003*

---

\* The financial support from the Wallander Foundation and Umeå School of Business and Economics is gratefully acknowledged. The comments of seminar/conference audiences at the universities of Porto and Umeå and at the 2003 Nordic Econometric Meeting and Xavier de Luna have improved the exposition.

## 1. INTRODUCTION

This paper considers the discrete nature of empirical duration data in frequently traded stocks. The output from a trading system records transactions to prevailing second and the resulting durations then contain a large fraction of zeros and very short durations. We take the continuous underlying duration density to be discretized and study the consequences for the econometric treatment of such data. A second objective of the paper is to give empirical evidence on the reaction to news in a specific stock.

Engle and Russell (1998) suggest conditional duration models for high frequency or intra-day time series data and emphasized the appealing properties of the quasi-maximum likelihood (QML) estimator based on the exponential duration model. There is an increasing amount of empirical research building on this QML estimator (e.g., Bauwens and Giot, 2001). Among the questions of interest in this field is the reaction to news as reflected, e.g., by indicators reflecting recent transactions (e.g., Easley and O'Hara, 1992; Engle and Russell, 1998).

When the data are available in only a discretized form the QML estimator loses some of its appeal. In essence, the consistency property requires a correct conditional mean specification and when data is discretized this is much harder to achieve as the conditional mean will then depend on the true underlying density. Kulldorff (1961, ch. 2) shows in a time invariant case the inconsistency of the maximum likelihood (ML) estimator arising from, e.g., using mid-interval values to represent the interval when data are discretized or grouped. He also demonstrates that discretization and the use of mid-interval values have more serious effects on the performance of ML estimators when the sample records durations to belonging to only a few and wide groups or intervals for continuous durations. For frequently traded stocks the groups are relatively few but the width is generally quite small. We may then expect the inconsistency to remain, but the actual performance of the ML or QML estimators may remain relatively advantageous. The inconsistency will remain when the model contains explanatory variables. In addition, if the true duration is viewed as continuous but only discrete time observations are available, any specification containing lagged durations will then be contaminated by a measurement error in a way to be made clear below.

In this paper we consider estimators that to some extent account for the outlined features of the data. Grouped data ML estimators and EM-algorithm versions are among

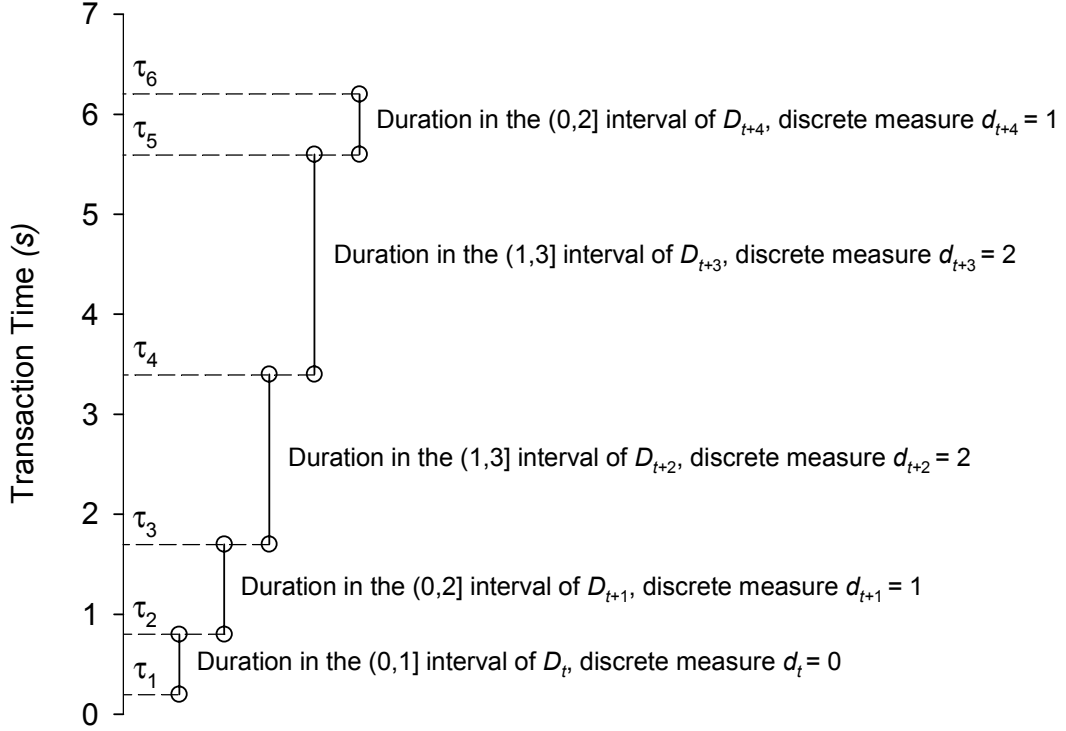


Figure 1: Examples of potential transaction times  $\tau_i$  and the registered duration variable  $d_t$  with the intervals of the underlying continuous duration variable  $D_t$ .

the studied estimators. We conduct a small Monte Carlo study focusing on the consequences of the various specification choices and the chosen estimators. Empirical results for a three-week period of transaction durations in Ericsson B at the order driven Stockholmsbörsen stock exchange in Stockholm are also to be reported. Of particular interest, beyond the focus on estimators, is here the reaction to news and whether the response to positive and negative news are different. We will use past price, spread and volume as indicators of new information to the market.

In Section 2 we discuss the model and discuss the ML and EM-algorithm estimators for discretized duration data. Section 3 reports the results from a set of Monte Carlo experiments conducted to study the consequences of the alternative ways of handling the discretized data. Section 4 reports the empirical results and the final section concludes.

## 2. MODEL AND ESTIMATORS

Let the  $t$ th continuous duration be denoted by  $D_t$ . The duration arises as a difference between two real transaction times,  $\tau$ , indexed by  $k$  and  $k - 1$ , i.e.  $D_t = \tau_k - \tau_{k-1}$ . When transactions are recorded at a second-level scale, the observed duration measure  $d_t = [\tau_k] - [\tau_{k-1}]$ , where  $[\cdot]$  signifies integer-value, is in seconds and hence integer-valued. For a frequently traded stock the durations are on average short and then  $d_t$  will take on one value from a set  $\{0, 1, 2, \dots, M_t\}$ , where  $\max\{M_t\}_{t=1}^T$  is a relatively small number and  $T$  is the length of the time series sequence of consecutive durations.

Figure 1 illustrates how the transaction times,  $\tau_k$ , the continuous durations,  $D_t$ , and the discretized durations,  $d_t$ , are related. Note that, except for  $d_t = 0$ , the  $d_t$ -values represent mid-interval values. For  $d_t = 0$  the mid-interval value is 0.5.

Given an assumption about the continuous and conditional distribution of  $D_t$  and given the information set  $\Delta_{t-1} = \{D_1, D_2, \dots, D_{t-1}\}$  it is straightforward to obtain the probability for  $d_t$  equal to some integer  $k$  given  $\Delta_{t-1}$ , as

$$\begin{aligned}
 \Pr(d_t = 0 | \Delta_{t-1}) &= \Pr(D_t \in (0, 1] | \Delta_{t-1}) \\
 &= \Pr(D_t \leq 1 | \Delta_{t-1}) \\
 &= q_{t1} \\
 \Pr(d_t = k | \Delta_{t-1}) &= \Pr(D_t \in (k-1, k+1] | \Delta_{t-1}) \\
 &= \Pr(D_t \leq k+1 | \Delta_{t-1}) - \Pr(D_t \leq k-1 | \Delta_{t-1}) \\
 &= q_{t,k+1} - q_{t,k-1}, \quad k \geq 1,
 \end{aligned} \tag{1}$$

with  $q_{t0} = \Pr(D_t \leq 0 | \Delta_{t-1}) = 0$ .

In the sequel of this and the next section we only consider the conditional exponential duration model, but any other reasonable duration distribution could have been considered instead. By focusing on the exponential model the technical aspects are kept simpler than for most other models, and the main ideas remain unaltered. For the exponential conditional duration model with conditional mean  $E(D_t | \Delta_{t-1}) = \theta_t > 0$  and conditional variance  $V(D_t | \Delta_{t-1}) = \theta_t^2$  we get explicit expressions for  $q_{tk}$  as

$$\begin{aligned}
 \Pr(d_t = 0 | \Delta_{t-1}) &= q_{t1} = 1 - e^{-1/\theta_t} \\
 \Pr(d_t = k | \Delta_{t-1}) &= e^{-(k-1)/\theta_t} - e^{-(k+1)/\theta_t}, \quad k = 1, 2, \dots, M_t.
 \end{aligned} \tag{2}$$

Since we can write  $D_t = \theta_t \varepsilon_t$ , with  $\varepsilon_t$  exponentially distributed with parameter one, we obtain the conditional expectation of  $D_t$  given  $d_t = k$  as

$$\begin{aligned} E(D_t|d_t = 0, \Delta_{t-1}) &= \theta_t \left[ 1 - \frac{1}{\theta_t(e^{1/\theta_t} - 1)} \right] \\ E(D_t|d_t = k, \Delta_{t-1}) &= \theta_t E \left[ \varepsilon_t | \varepsilon_t \in \left( \frac{k-1}{\theta_t}, \frac{k+1}{\theta_t} \right) \right] \\ &= \theta_t + \frac{e^{1/\theta_t}(k-1) - e^{-1/\theta_t}(k+1)}{e^{1/\theta_t} - e^{-1/\theta_t}}, \quad \text{for } k \geq 1. \end{aligned} \quad (3)$$

It may be shown that  $E(D_t|d_t = k, \Delta_{t-1}) \leq k$  for  $k \geq 1$  and that  $E(D_t|d_t = 0, \Delta_{t-1}) \leq \frac{1}{2}$ , i.e. both conditional expectations can be expected to be smaller than their corresponding mid-interval values. Equality arises only when  $\theta_t \rightarrow \infty$ .

In the conventional continuous duration framework advanced by Engle and Russell (1998) the  $\theta_t$  function is of the type:

$$\begin{aligned} \theta_t &= \alpha_0 + \alpha_1 D_{t-1} + \dots + \alpha_q D_{t-q} + \beta_1 \theta_{t-1} + \dots + \beta_p \theta_{t-p} + \mathbf{x}_t \boldsymbol{\pi} \\ &= \mathbf{z}_t \boldsymbol{\psi}, \end{aligned} \quad (4)$$

where  $\mathbf{x}_t$  is a vector of predetermined variables containing, e.g., past prices. Setting  $\xi_t = D_t - \theta_t$  in (4) enables us to rewrite the model on the alternative form

$$\begin{aligned} D_t &= \alpha_0 + (\alpha_1 + \beta_1) D_{t-1} + \dots + (\alpha_q + \beta_q) D_{t-q} + \beta_{q+1} D_{t-q-1} + \dots + \beta_p D_{t-p} \\ &\quad + \xi_t - \beta_1 \xi_{t-1} - \dots - \beta_p \xi_{t-p} + \mathbf{x}_t \boldsymbol{\pi}, \quad \text{for } p \geq q. \end{aligned} \quad (5)$$

This is an ARMAX model in the continuous exponential duration variable. Obviously, other specifications are also feasible (e.g., Bauwens and Giot, 2001, ch. 3).

## 2.1 Estimators

We first consider estimation that accounts for the discreteness in the conditional variable  $d_t$  to be explained. Later we extend the estimation setup by also considering the discreteness in the lagged durations that serve as explanatory variables in the  $\theta_t$  function.

The log-likelihood function for the discrete conditional variable of (2)-(4) takes the form

$$\ln L = \sum_{t=r}^T \ln(e^{-\eta_t(y_t-1)/\theta_t} - e^{-(y_t+1)/\theta_t}), \quad (6)$$

where  $r = \max(p, q) + 1$  and  $\eta_t = 0$ , for  $y_t = 0$ , and  $\eta_t = 1$ , for  $y_t \geq 1$ . The associated score vector can be expressed

$$\frac{\partial \ln L}{\partial \psi} = \sum_{t=r}^T \frac{\mathbf{z}_t'}{(\mathbf{z}_t \psi)^2} \left[ \frac{\eta_t(y_t - 1)e^{-\eta_t(y_t-1)(\mathbf{z}_t \psi)^{-1}} - (y_t + 1)e^{-(y_t+1)(\mathbf{z}_t \psi)^{-1}}}{e^{-\eta_t(y_t-1)(\mathbf{z}_t \psi)^{-1}} - e^{-(y_t+1)(\mathbf{z}_t \psi)^{-1}}} \right].$$

Obviously, other duration densities such as Weibull or log-logistic could also have been applied. In the absence of strong a priori arguments for a particular model one avenue would be to specify an even wider class of densities such as the generalized gamma. The Appendix gives expressions for the Weibull and Burr models, which also will be used in the empirical study below.

For (4)-(5) lagged continuous  $D_{t-i}, i = 1, \dots, q$ , variables are assumed observed. Obviously, if durations are measured discretely  $D_{t-i}$  are not observed but the  $d_{t-i}, i = 1, \dots, q$ , are. A consequence of using the discrete  $d_{t-i}, i = 1, \dots, q$ , is that measurement errors are introduced. These then imply that the ML estimator is inconsistent.

Consider as a simple example of this inconsistency an underlying exponentially distributed variable and observations falling into either of the two intervals  $[0, 1]$  and  $(1, \infty)$  with the indicator variable  $d_t = 1$  for the latter interval. Let  $\theta_t = \alpha D_{t-1}$  in the true case and  $\theta_t = \alpha d_{t-1}$  in the assumed case. The score for the assumed model is

$$\frac{\partial \ln L}{\partial \alpha} = \sum_{t=2}^T \frac{d_t - e^{-1/(\alpha d_{t-1})}}{\alpha^2 d_{t-1} [1 - e^{-1/(\alpha d_{t-1})}]},$$

with  $d_t = 0$  or  $d_t = 1$ . As  $E(d_t | \Delta_{t-1}) = \Pr(d_t = 1 | \Delta_{t-1}) = \exp(-1/(\alpha D_{t-1}))$ , a first order expansion of the score  $\partial \ln L / \partial \hat{\alpha}$  around the true parameter value and manipulation shows that the bias depends on  $D_{t-1} - d_{t-1}$ . As this difference can be expected to be larger than zero, the ML estimator  $\hat{\alpha}$  can be expected to be too large, or alternatively the estimated mean can be expected to be too small.

The ML estimator is consistent and asymptotically normal when  $\alpha_1 = \alpha_2 = \dots = \alpha_q = 0$  (no measurement errors as  $d_{t-i}, i = 1, \dots, q$ , are not included in  $\theta_t$ ) and the  $\beta_1, \dots, \beta_p$  parameters are such that the  $\{D_t\}$  sequence is stationary. An early proof of the asymptotic results for a scalar case is due to Kulldorff (1961), who also studied the loss in efficiency that results from discretizing the time scale. Engle and Russell (1998) consider the case of explanatory variables, continuous durations and the QML estimator. Given a correctly specified stationary model the QML estimator can be shown to be consistent and asymptotically normal.

The EM-algorithm (Dempster, Laird and Rubin, 1987) provides a general framework for dealing with aspects of the limited information in the  $\{d_t\}$  sequence that we have in this case. If we take the lagged  $d_{t-i}$  variables as is and focus only on the grouped  $d_t$  indicator, we can easily extend the constant parameter and grouped exponential model of Little and Rubin (1987, ch. 11). The M-step maximizes the conditional expectation of the log-likelihood function for  $D_t$  given  $d_t$  with respect to  $\psi$ , i.e.  $Q(\psi, \tilde{\psi}) = E_{\tilde{\psi}} [\ln L(D)|d, \Delta] = \sum_{t=r}^T \left[ -\ln \mathbf{z}_t \psi - E_{\tilde{\psi}}(D_t|d_t, \Delta_{t-1})/\mathbf{z}_t \psi \right]$ . The required conditional expectation is given in (3) and should be evaluated at  $\tilde{\psi}$  (the E-step). The E and M steps are iterated until convergence. Note that the M-step uses the continuous exponential variable log-likelihood function and should therefore be computationally straightforward. The score vector is  $\partial Q/\partial \psi = \sum_{t=r}^T \mathbf{z}_t' \left[ E_{\tilde{\psi}}(D_t|d_t, \Delta_{t-1}) - \mathbf{z}_t \psi \right] / (\mathbf{z}_t \psi)^2$  and the Hessian is  $\partial^2 Q/\partial \psi \partial \psi' = \sum_{t=r}^T \mathbf{z}_t' \mathbf{z}_t / (\mathbf{z}_t \psi)^2$ .

Consider next an EM-algorithm that attempts to account also for the presence of  $d_{t-i}$  when the true  $D_{t-i}$  would have been preferred in  $\theta_t$ . Little and Rubin (1987, ch. 8) consider a problem of missing observations in a Gaussian AR(1) model, but there appears to be no reported research on the type of problem we have in mind. Recall that the density is a conditional one so that conditioning on past  $d_{t-i}$  and not on future  $d_{t+j}$  appears reasonable. Then  $Q(\psi, \tilde{\psi}) = \sum_{t=r}^T \left[ -E_{\tilde{\psi}}(\ln \mathbf{z}_t \psi | \bar{d}_t, \Delta_{t-1}) - E_{\tilde{\psi}}(D_t/\mathbf{z}_t \psi | \bar{d}_t, \Delta_{t-1}) \right]$ , where  $\bar{d}_t = (d_1, \dots, d_t)$ . This expression is a difficult one to use as it involves taking expectations of past  $D_{t-i}$  with respect to more recent  $d_t$ , because it involves nonlinearities, and because different time periods are interwoven in the final expectation expression. Therefore, no attempt is made at obtaining an exact EM-algorithm in this paper.

An ad hoc EM-algorithm could use  $E_{\tilde{\psi}}(D_t|d_t, \Delta_{t-1})$  for all  $t$  and hence also for those lags that are included in the  $\mathbf{z}_t$  vector. The performance of this estimator is studied by Monte Carlo simulation in the next section and also used empirically.

The simulated maximum likelihood (SML) estimator (Gourieroux and Monfort, 1991) offers an interesting approach to coping with the discreteness of the data. Unfortunately, the results of Cappé et al. (2002) suggest that the SML estimator may be a very time demanding exercise as the number of replications should increase with the number of observations. For high frequency data the sample size is usually large. Note also that the current context differs from the dynamic limited dependent variable model considered by, e.g., Lee (1997).

An obvious way of attempting to avoid the bias arising from measurement errors in lagged variables would be to specify the joint distribution for  $\{\Delta_t\}$ . It would then be possible to avoid the measurement error problem by accounting for the discretization for all  $t$ . Unfortunately, such modelling would also be subject to even larger risks of distributional misspecification as multivariate distributions can come in many alternative shapes. Computationally it is potentially a difficult problem. Closely related to this idea and more directly focusing on the discrete data would be a direct specification of an inhomogeneous Markov chain with transition probabilities depending, e.g., on  $\mathbf{x}_t$ .

The geometric distribution is the discrete time equivalent of the exponential distribution and with a conditional interpretation the lagged  $d_{t-i}$  presents no problem. When data is genuinely continuous this can at most be regarded as an approximation. The likelihood function for a mean  $\mu_t$  and variance  $\mu_t(1+\mu_t)$  geometric distribution is (e.g., Cameron and Trivedi, 1998)  $L = \prod_{t=r}^T \mu_t^{d_t} (1 + \mu_t)^{-1-d_t}$ . For time invariant  $\mu$  the maximum likelihood estimator is of the explicit form;  $\hat{\mu} = \sum_{t=1}^T d_t / T$  and  $V(\hat{\mu}) = \mu(1+\mu)/T$ . It can be demonstrated that the estimator is biased and inconsistent when data is generated according to a continuous exponential model. To specify a dynamic model we set the conditional mean  $\mu_t = \theta_t = \mathbf{z}_t \boldsymbol{\psi}$  to obtain the score vector  $\partial \ln L / \partial \boldsymbol{\psi} = \sum_{t=r}^T \mathbf{z}_t' [d_t - \mathbf{z}_t \boldsymbol{\psi}] / [\mathbf{z}_t \boldsymbol{\psi} (1 + \mathbf{z}_t \boldsymbol{\psi})]$ . The covariance matrix of the estimator is estimated by the inverse of the information matrix  $\sum_{t=r}^T \mathbf{z}_t' \mathbf{z}_t / [\mathbf{z}_t \boldsymbol{\psi} (1 + \mathbf{z}_t \boldsymbol{\psi})]$  and evaluated at estimates.

### 3. MONTE CARLO STUDY

In this section we study the properties of the estimators for the various model versions when data are artificially generated according to conditional exponential and Weibull models.

We specify the conditional mean function that is used in generating the underlying  $D_t$  data as

$$\theta_t = \alpha_0 + \alpha_1 D_{t-1} + \beta_1 \theta_{t-1}.$$

Engle and Russell (1998) give the following moment results for the exponential model

$$\begin{aligned} E(D_t) &= \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \\ V(D_t) &= \frac{1 - \beta_1^2 - 2\alpha_1\beta_1}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 2\alpha_1^2} \times E^2(D_t). \end{aligned}$$



These results can be obtained by substituting  $\theta_{t-1}\varepsilon_{t-1}$  for  $D_{t-1}$  with  $\varepsilon_{t-1}$  exponentially distributed with parameter one. From the variance expression it follows that the parameters should satisfy  $2\alpha_1^2 + \beta_1^2 + 2\alpha_1\beta_1 < 1$  and from the mean expression they should also satisfy  $\alpha + \beta < 1$ . In the experiments with Weibull distributed durations we employ standardization to obtain the mean and variance of the exponential model, cf. the Appendix.

The study uses  $\alpha_1 = 0.2, 0.3$  and  $0.4$ ,  $\beta_1 = 0.15, 0.2$  and  $0.25$ , and  $\alpha_0 = 2.5, 5$  and  $10$ , to give mean durations in the range  $3.8 - 28.6$  seconds with variances in the range  $4.2 - 50.5$ . For the Weibull model we use  $\gamma = 0.8$ , which corresponds to negative duration dependence. The time series length is set at  $T = 5\,000$  and  $50\,000$ . The  $T = 5\,000$  case corresponds to a short time series length for frequently traded stocks, and  $T = 50\,000$  is used only for the shortest durations ( $\alpha_0 = 2.5$ ) and exponential data. In each design cell 1000 replications are generated starting from the same initial seed. In generating the series an initial part of 100 observations is dropped. Data are next discretized in accordance with the discussion in Section 2. The results are reported in terms of bias and mean square error (MSE) measures in Figures 2-3 and in Tables A1-A5 of the Appendix. We confine the presentation of results mainly to the  $\alpha_1$  and  $\beta_1$  parameters and the exponential model for  $T = 5\,000$ .

The following models and estimation algorithms are used: (i) the continuous time exponential model (indicated by C) serves as a base case and is estimated by ML and a scoring algorithm. All other data sets are based on discrete duration  $\{d_t\}$  sequences. (ii) The same ML algorithm as in (i) is used with discrete data (indicated by D). Note that for  $d_t = 0$  we use  $d_t = 0.5$  instead. This corresponds to the mid-interval value as for other  $d_t$ -values. (iii) the grouped data ML estimator with  $d_{t-1}$  in the  $\theta_t$  function is estimated by a BHHH algorithm (indicated by G); (iv) the EM-algorithm with  $\hat{D}_t$  replacing  $D_t$  for all  $t$  is estimated by alternating between a ML and an E-step (indicated by EM). Hence, even if data is generated as Weibull distributed the employed density underlying the estimators is throughout the exponential. Note that all estimators are conditional ones, as estimation is throughout conditional on initial observations. The number of iterations is limited to 100 and true parameter values are used to initialize iterations. All computations are performed using Fortran code on a 1.9 GHz Laptop.

The biases of the estimators of  $\alpha_1$  and  $\beta_1$  are displayed in Figure 2 for  $T = 5\,000$

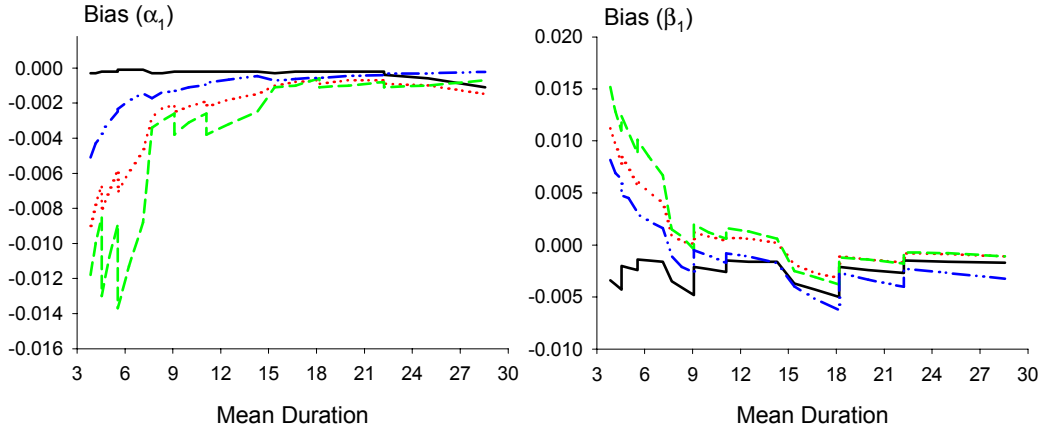


Figure 2: Biases for the estimators of  $\alpha_1$  and  $\beta_1$  ( $T = 5\,000$  and  $1000$  replications; Solid line indicates continuous duration ML estimator (C), dotted line discrete data with the continuous duration ML estimator (D), dashed line indicates the grouped data ML estimator (G) and the dot-dashed line the EM-algorithm (EM)).

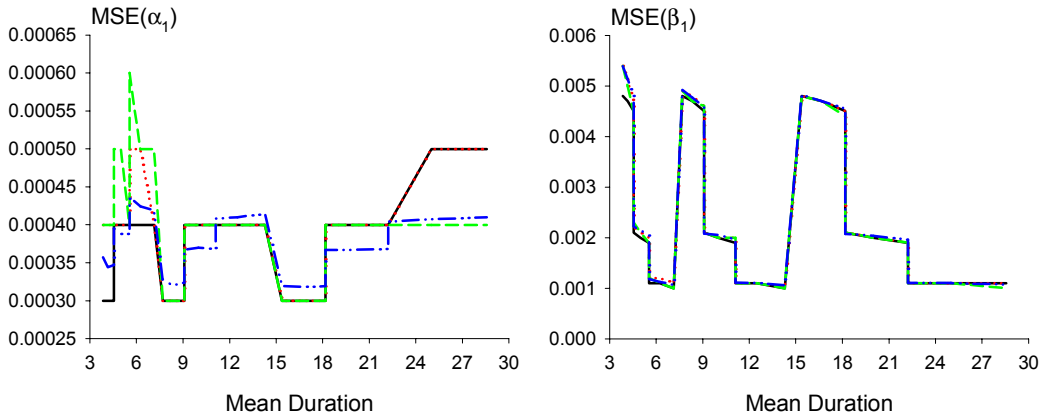


Figure 3: MSEs for the estimators of  $\alpha_1$  and  $\beta_1$  ( $T = 5\,000$  and  $1000$  replications; Solid line indicates continuous duration ML estimator (C), dotted line discrete data with the continuous duration ML estimator (D), dashed line indicates the grouped data ML estimator (G) and the dot-dashed line the EM-algorithm (EM)).

and exponential data. It is quite obvious from the patterns for both parameters that the ML estimator based on continuous data has small bias. All other estimators are based on discretized data and manifest some bias for short durations, while bias is much less of an issue for longer durations. The largest bias for  $\alpha_1 = .2$  and the shortest duration of 3.8 seconds is noted for the grouped data ML estimator and amounts to 6 percent. For the EM-algorithm the corresponding bias is less than 3 percent. For both parameters there is a clear-cut ranking of the estimators, in particular for the short mean durations. The biases of the EM-algorithm are smaller than the biases of the discretized data ML and grouped data ML estimators. It appears that the grouped data ML estimator has the weakest performance. As the EM-algorithm in this particular case is rather fast to calculate it is our tentative choice of a best estimator. Table A1 contains the detailed biases (all multiplied by 100) for the parameter estimators. For  $T = 50\,000$  corresponding results are given in Table A3 for short mean durations ( $\alpha_0 = 2.5$ ). The results reiterate the main conclusions derived from Figure 2. It is quite apparent that all estimators but the continuous data ML estimator (C) have a bias and that the EM-algorithm comes out as the least biased estimator for discretized data. For Weibull data, cf. Table A4, the internal ranking between estimators remain relatively unaltered.

The MSE results of the  $\alpha_1$  and  $\beta_1$  estimators are exhibited in Figure 3 for  $T = 5\,000$  and the exponential data. When it comes to the MSEs for  $\beta_1$  the most apparent feature is their striking similarity across mean durations. For this parameter the MSE is then completely dominated by the variance component. For  $\alpha_1$  there is some variation for short durations and for the long ones (see also Table A2). The MSE of the EM-algorithm is not much different from those of the continuous duration ML estimator based on discretized data and the grouped data ML estimator. Among the latter two, the grouped data ML estimator has the weaker performance for short mean durations. For  $T = 50\,000$  there is an expected drop in MSEs due to sample size, but the ranking between estimators remain unaltered. As expected the MSEs of the base case ML estimator for the exponential model (C) are the smallest in most cases, and also for the Weibull generated data, cf. Table A5.

In summary, among the estimators accounting for discretized data model the EM-algorithm is the preferred estimator in terms of bias. With respect to MSE it is not worse than the two competitors, though differences are quite small. No estimator manages to completely avoid bias for short mean durations.

## 4. EMPIRICAL RESULTS

### 4.1 Data and Descriptives

Empirical results are reported for 15 days of trading (July 2, 2002 – July 22, 2002) in Ericsson B at the order driven Stockholmsbörsen stock exchange in Stockholm. The data were downloaded from the Ecovison system and processed further by the authors. The number of observed durations or the time series length is, after some reduction due to day changes, 57 735. On average there are 3849 durations per day. Figure 4 gives a histogram of the durations. The estimated average of the integer-valued duration is 7.4 seconds with a standard deviation of 11.2 seconds. The average varies between 2.9 to 13.5 seconds over the 15 days. About 79 percent of the durations are 10 seconds or shorter and the longest duration is 403 seconds. To give an indication of the trading volume, the number of traded stocks during the first day of the sample period is 12 596 496 with a closing price of 14.90 SEK. The trading volume in the major summer vacation month of July is usually smaller than during other months.

Table 1 reports the one step transitions (in row percent) between successive durations. The matrix is asymmetric so that independently of the size of  $d_{t-1}$  the next duration  $d_t$  is most likely shorter. For instance, given  $d_{t-1} = 3$ , 52.7 percent of the durations at  $t$  can be expected to be shorter than or equal to  $d_t = 3$ . Figure 5 gives the autocorrelation function for the time series of successive durations. The autocorrelations are quite small but the function decreases only slowly. Note that all autocorrelations are positive. The partial autocorrelations decrease rather quickly and are approximately zero after 5-6 lags. The patterns of Table 1, Figure 5 and the partial autocorrelation function indicate that the model should be able to capture low order both autoregressive and moving average effects.

Figure 6 exhibits the seasonal pattern across the hours of the day. There appears to be a weakly increasing pattern so that trading is slightly less frequent (longer durations) towards the end of the trading day. There appears to be no strong reasons for deseasonalizing the series as done in some previous studies.

For a pure time series analytical approach (i.e.  $\boldsymbol{\pi} = \mathbf{0}$  in (4)) a reasonable starting point is to search for a model with  $p \leq 3$  and  $q \leq 3$ . In addition, in the final models we include as explanatory variables the price (mean 14.61, standard deviation 1.90), the

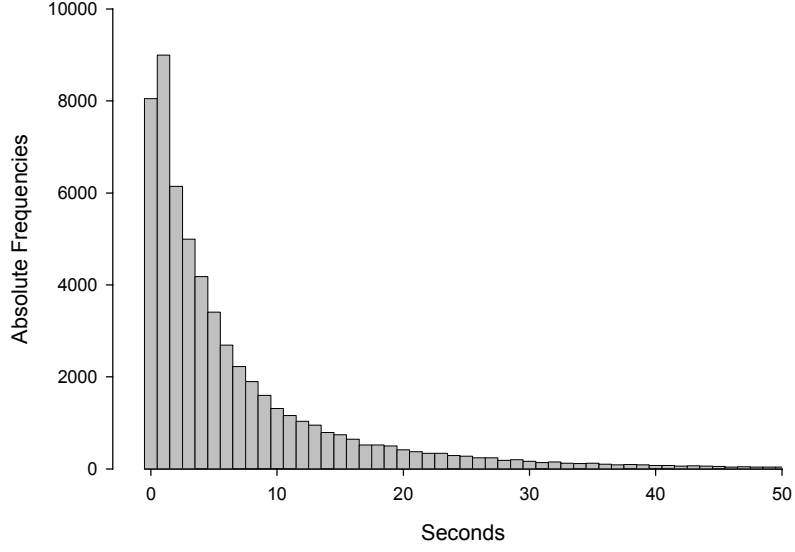


Figure 4: Histogram of discrete durations  $d_t$  ( $T = 57\,735$ , the one percent of durations exceeding 50 seconds is excluded).

Table 1: Transition matrix between successive durations (only durations and their lags shorter than 10 seconds are included in the table; in percent).

$d_{t-1}$	$d_t$										
	0	1	2	3	4	5	6	7	8	9	10
0	16.8	21.4	11.6	8.5	6.5	5.7	4.1	3.5	2.6	2.2	1.8
1	18.6	19.5	11.9	9.1	7.3	5.5	4.5	3.3	3.0	2.1	1.9
2	14.9	18.0	12.6	9.7	7.6	6.3	4.8	3.5	2.9	2.4	2.2
3	13.8	17.3	12.3	9.3	7.5	5.7	4.7	3.7	3.0	2.5	2.2
4	12.6	14.8	11.6	9.3	7.8	6.9	4.9	3.9	3.3	2.9	2.1
5	11.6	14.5	12.3	10.6	7.7	5.6	4.3	4.1	3.2	3.3	1.9
6	12.2	13.8	10.0	8.7	7.1	6.1	5.0	4.5	3.3	3.0	2.3
7	12.3	13.6	9.3	7.9	7.9	6.1	4.9	4.5	3.7	3.1	3.0
8	11.8	12.9	9.7	8.4	7.3	6.3	5.5	4.3	4.0	3.4	2.5
9	11.9	12.1	8.7	7.8	7.4	6.4	5.3	4.2	4.0	2.8	2.8
10	11.5	12.0	10.0	8.5	7.2	6.6	4.0	3.1	4.2	2.9	2.7

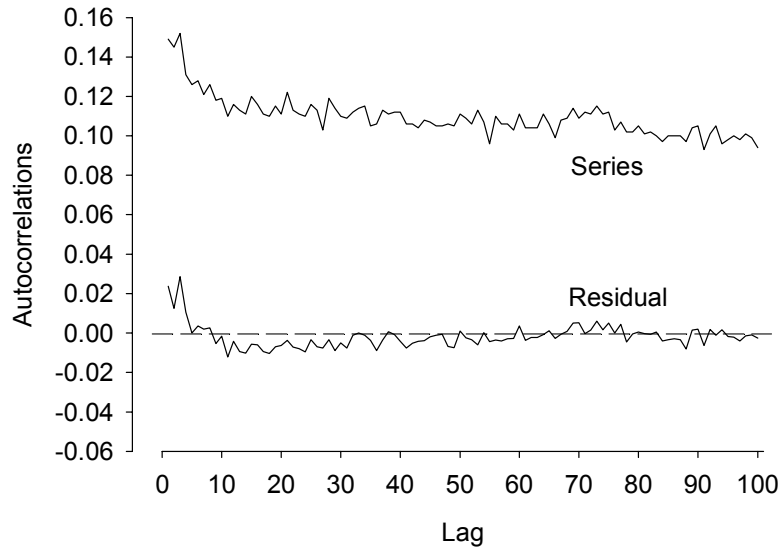


Figure 5: Autocorrelation function for durations  $d_t$  and residuals (calculated as  $(d_t - \hat{\theta}_t)/\hat{\theta}_t$ ) from the exponential model.

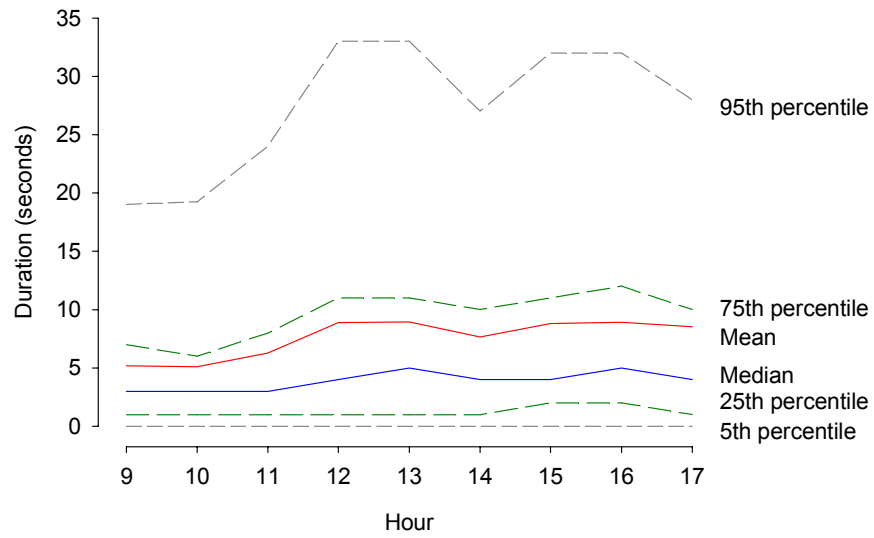


Figure 6: Descriptive statistics for durations per hour of the day (entire time series).

spread (0.10, 0.02) and the number of traded stocks ( $2134.5, 1.48 \cdot 10^6$ ) ending the previous duration. Including the first two variables as changes instead of as levels was rendered empirical support, see below. For example, the price part of the model,  $\pi_1 p_{t-1} + \pi_2 p_{t-2}$ , was used and empirically we found  $\hat{\pi}_1 \approx -\hat{\pi}_2$ . This suggests the use of a restricted  $\pi(p_{t-1} - p_{t-2}) = \pi \nabla p_{t-1}$  specification, i.e. in terms of a change.

#### 4.2 Estimation Results

To estimate the parameters we assume three parametric density specifications that have been used previously – the exponential, the Weibull and the Burr (see the Appendix for a brief account of the latter two distributions). The Weibull contains the exponential model as a special case. The Burr model is more flexible than Weibull in that it has more parameters and then a more flexible hazard function. The Burr model does not nest neither the exponential nor the Weibull models, so that straightforward use of, e.g., likelihood ratio tests for model selection is ruled out. We employ two versions of the EM-algorithm for the exponential model.

The continuous exponential model served as a tool for determining the model specification.<sup>1</sup> The best model has  $R^2 = 0.1$ . There is some remaining serial correlation in all models to be reported and this could not be eliminated, cf. Figure 5 for the autocorrelation function corresponding to column one of Table 2.<sup>2</sup> Note that no serial correlations are determined for the discretized models. No serial correlation remains in the squared residuals, except for in the Burr model and for the model of the final column of Table 2. Individual correlations are, however, quite small and the Ljung-Box statistic is obviously influenced by the large sample size.

The estimation results are presented in Table 2-4. The parameter estimates are throughout almost exclusively of the same sign, roughly of similar sizes and when significant this happens across models and estimators. Note that there are more lags in these models than in most previous models.

Table 2 reports results based on an assumed, continuous variable exponential model. It is found that the estimated models of this table (and other tables) satisfy the station-

---

<sup>1</sup>In this case and whenever continuous variable methods are employed 0 is replaced by 0.5 seconds to reflect mid-interval-value in the same way as for longer durations.

<sup>2</sup>The residual is defined as  $r_t = (d_t - E(D_t|\Delta_{t-1}))/V^{1/2}(D_t|\Delta_{t-1})$ , where  $E(\cdot)$  and  $V(\cdot)$  are different for the different distributions. The squared residual is  $r_t^2$ .

Table 2: Maximum likelihood estimates for alternative specifications of the continuous exponential model.

Variable	Coeff	s.e.	Coeff	s.e.	Coeff	s.e.	Coeff	s.e.
$d_{t-1}$	0.0378	0.0029	0.0381	0.0029	0.0408	0.0030	0.0356	0.0030
$\theta_{t-1}$	0.8962	0.0603	0.9162	0.0756	0.8712	0.0524	0.9905	0.0149
$\theta_{t-2}$	-0.2634	0.0803	-0.3072	0.0966	-0.3350	0.0697	-0.3120	0.0185
$\theta_{t-3}$	0.3282	0.0564	0.3506	0.0645	0.4818	0.0041	0.2834	0.0275
Price change	2.8274	0.3564	—	—	3.1143	0.3419	1.6284	0.2899
$p_{t-1}$	—	—	0.4985	0.0090	—	—	—	—
$p_{t-2}$	—	—	-0.4954	1.3161	—	—	—	—
Spread change	-1.7938	1.1494	-0.8952	0.8908	—	—	1.0475	1.0689
$s_{t-1}$	—	—	—	—	-1.0680	0.8481	—	—
$s_{t-2}$	—	—	—	—	1.2123	1.3161	—	—
Volume	0.6266	2.6438	-0.7224	1.1746	0.8041	1.1797	—	—
$v_{t-1}$	—	—	—	—	—	—	0.6973	0.0048
$v_{t-2}$	—	—	—	—	—	—	-0.6945	1.3161
Constant	0.0103	0.0354	-0.0250	0.1116	-0.0038	0.0935	0.0066	0.0531
LB <sub>100</sub>	235.4	—	233.1	—	234.9	—	382.5	—
LB <sub>100</sub> <sup>2</sup>	6.7	—	6.2	—	6.7	—	828.6	—
ln $L$	-165860	—	-165860	—	-165859	—	-165743	—

Notes: Volume pertains to the previous transaction, while  $v_t$  is the accumulated (within the day) trading volume. Both are throughout divided by 10 000 000. LB<sub>100</sub> is the Ljung-Box statistic of the standardized residual over 100 lags. LB<sub>100</sub><sup>2</sup> is the same statistic for squared residuals.



Table 3: Parameter estimates for the exponential model.

Variable	ML-Continuous		ML-Grouped		EM-Grouped		EM-full	
	Coeff	s.e.	Coeff	s.e.	Coeff	s.e.	Coeff	s.e.
$d_{t-1}$	0.0378	0.0029	0.0380	0.0029	0.0379	0.0029	0.0403	0.0029
$\theta_{t-1}$	0.8962	0.0603	0.9172	0.0548	0.8800	0.0532	0.8841	0.0530
$\theta_{t-2}$	-0.2634	0.0803	-0.3036	0.0742	-0.2681	0.0712	-0.2693	0.0704
$\theta_{t-3}$	0.3282	0.0564	0.3473	0.0522	0.3492	0.0507	0.3337	0.0496
Price change	2.8274	0.3564	2.9936	0.3350	3.1907	0.3330	3.1266	0.3303
Spread change	-1.7938	1.1494	-3.0104	1.0554	-2.9918	1.0574	-1.4293	1.0879
Volume	0.6266	2.6438	0.8230	2.5724	0.9936	2.5879	0.7678	2.5562
Constant	0.0103	0.0354	0.0092	0.0317	0.0078	0.0311	0.0096	0.0314
LB <sub>100</sub>	235.4		—		—		—	
LB <sub>100</sub> <sup>2</sup>	6.7		—		—		—	
$\ln L$	-165860		-131050		-165096		-165105	

Notes: All estimation results are obtained by Fortran coded programs. Repeated use is made of a simplex algorithm and the outer product gradient is used for the covariance matrix. EM-full is the EM-algorithm used in the Monte Carlo experiment. EM-Grouped is the grouped data EM-algorithm estimator.

arity condition on the  $\alpha$  and  $\beta$  parameters, albeit with a rather narrow margin. Initially alternative lag structures (different  $p$  and  $q$  values) were tried. Table 2 also reports on how explanatory variables should be included. There is strong support throughout for utilizing change variables for the price and spread. If, e.g., the price follows a random walk the change corresponds to the innovation or the unpredicted new information over the previous duration. A positive price change leads to a longer duration. The effect of the spread change is negative but not significant. A higher trading volume prolongs the next duration but not significantly so. The final column suggests that separate inclusion of  $v_{t-1}$  and  $v_{t-2}$  is preferable judging by the log-likelihood values. However, the serial correlation properties speak against this specification. The change variables will be retained in all further model estimations.

In Table 3 a comparison within the exponential model of using continuous or grouped data is reported. There are no substantial differences between the ML estimators based on the two data types. The two versions of the EM-algorithm are quite similar, too.

Table 4: Maximum likelihood estimates for alternative model specifications.

Variable	Weibull				Burr			
	Continuous		Grouped		Continuous		Grouped	
	Coeff	s.e.	Coeff	s.e.	Coeff	s.e.	Coeff	s.e.
$d_{t-1}$	0.0393	0.0425	0.0437	0.0042	0.0386	0.0033	0.0418	0.0042
$\theta_{t-1}$	0.8875	0.0678	0.7792	0.0698	0.9994	0.0629	0.9990	0.0771
$\theta_{t-2}$	-0.2967	0.0979	-0.2748	0.0860	-0.5001	0.0848	-0.5159	0.1045
$\theta_{t-3}$	0.3663	0.0629	0.4455	0.0631	0.4434	0.0589	0.4679	0.0709
Price change	3.0200	0.4241	3.1342	0.4640	2.9754	0.4057	2.5722	0.4725
Spread change	-1.2306	1.3314	-1.1535	1.4182	-0.7020	1.0991	-0.5374	1.3905
Volume	0.6712	3.0271	0.2148	3.3013	-0.9633	2.2659	-0.0063	3.1934
Constant	0.0132	0.0425	0.0045	0.0470	0.0286	0.0405	0.0044	0.0426
$\gamma$	0.9130	0.0023	0.8054	0.0005	1.1263	0.0019	0.8286	0.0044
$\lambda$	—		—		0.3600	0.0083	0.0350	0.0032
LB <sub>100</sub>	229.7		—		201.3		—	
LB <sub>100</sub> <sup>2</sup>	7.1		—		240.9		—	
$\ln L$	-165409		-129178		-164497		-129152	

Note: See Table 3 for explanations.

Given this result, arguments supportive of the conventional QML estimator even if data are discretized are strengthened.

Table 4 studies this issue further; if the QML is to be useful we would expect no large changes in parameter estimates even if the exponential model is not the 'true' one. The qualitative conclusions correspond to those of the exponential model, though sizes of estimates are slightly different. The exponential model is nested within the Weibull model and the exponential can be rejected against the Weibull model ( $\hat{\gamma}$  is significantly smaller than one). The Weibull model is not nested within the more general Burr model, though the  $\gamma$  and  $\lambda$  estimates of the latter model may indicate that the shape of the Weibull hazard is not supported by data. The Weibull duration dependence parameter  $\gamma$  is significantly smaller than one, which implies a decreasing hazard function and that the exponential model can be rejected. In a similar way the form of the Burr hazard function is an indication against the exponential model.

Figure 7 shows the Burr and Weibull hazard functions, when  $\theta$  is replaced by the

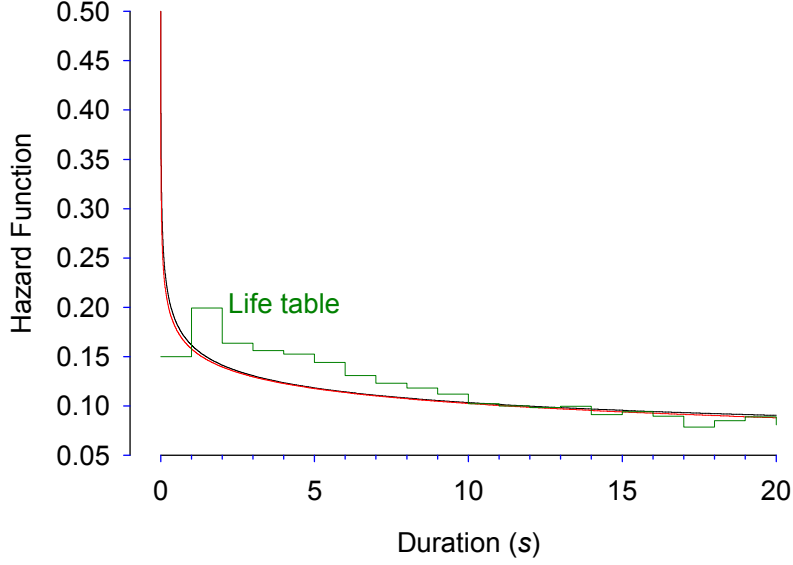


Figure 7: Hazard functions based on grouped ML estimates of Table 4 and evaluated at the sample mean of the duration variable, and a life table estimated hazard function.

sample mean and estimates from Table 4 are used for  $\gamma$  and  $\lambda$ . The hazard functions are hardly distinguishable and decrease rapidly within the first second, but are roughly constant thereafter. Hence, these hazards differ the most from the life table estimate in the  $(0,1)$  interval and discrimination between the two parametric models would obviously be much strengthened if short and continuous duration data in the  $(0,1)$  interval were available.

We also studied whether the response to news is symmetric in the sense that positive and negative news affect subsequent durations in the same way. The potentially asymmetric response to news (the variables are constructed as  $\nabla x_t^+ = \max(0, \nabla x_t)$  and  $\nabla x_t^- = \min(0, \nabla x_t)$ ) is studied in terms of the price and spread changes within the framework of the grouped data Weibull and Burr models. By likelihood ratio tests we find no evidence of asymmetric response to price changes and the two estimates for positive and negative changes are quite similar. There are different responses to spread changes depending on their signs, but not significantly so. Individually neither of the spread change effects appear to have a significant effect.

## 5. CONCLUSIONS

The paper has discussed the discrete nature of duration measures between transactions in stocks and studied the consequences of this discretization of a continuous time scale.

Grouped maximum likelihood and EM-algorithm estimators were discussed. In the small Monte Carlo study the EM-algorithm that accounts for the discrete nature of the data both in the outcome and the lagged explanatory variables comes out as the best estimator of the compared ones. In the empirical study the differences between estimators are generally quite small, and the EM-algorithm and ML estimators based on discrete data are not too different from ML based on grouped data and Weibull and Burr models.

When it comes to the effects of explanatory variables the study provided support for using changes rather than levels to reflect news. There is throughout a significant and positive effect of news about prices and a negative effect of a change in the spread. The spread effect is not significant, however. A higher volume has an insignificant but prolonging effect in most cases. We could not find statistically significant support for an asymmetric response to news about spreads nor about prices. The log-likelihood function value of the Burr is larger than for other models but the models are not nested. In addition, the serial correlation properties of the exponential and Weibull models speak in favor of these two models. A generalized gamma was also employed and provided a better fit to the data than both the exponential and Weibull models. A reason for not reporting generalized gamma results is the numerical problems we faced in obtaining standard errors.

## APPENDIX

### Weibull

Using the specification of Bauwens and Giot (2001, pp. 98-99) the Weibull model has hazard function  $\lambda(D) = \gamma D^{\gamma-1}/\theta^\gamma$ , which gives the integrated hazard function  $\Lambda(D) = (D/\theta)^\gamma$ . From this follows the distribution function  $F(D) = 1 - \exp(-\Lambda(D))$ , expected value  $E(D) = \theta \Gamma(1 + \gamma^{-1})$  and variance  $V(D) = \theta^2 [\Gamma(1 + 2\gamma^{-1}) - \Gamma(1 + \gamma^{-1})]$ . Random durations can be generated according to  $D = \theta [-\ln(1 - u)]^{\gamma^{-1}}$ , where  $u$  is a uniform  $[0, 1]$  random deviate. A standardization of  $D$  to get the moments of the exponential model is obtained by  $D_* = a^{-1/2}(D - \theta b)$ , where  $a = [\Gamma(1 + 2\gamma^{-1}) - \Gamma(1 + \gamma^{-1})]$  and  $b = \Gamma(1 + \gamma^{-1}) - a^{1/2}$ .

Corresponding to the log-likelihood function in (6) we have

$$\begin{aligned} \ln L &= \sum_{t=r}^T l_t = \sum_{t=r}^T \ln [\exp(-\eta_t [(d_t - 1)/\theta_t]^\gamma) - \exp(-[(d_t + 1)/\theta_t]^\gamma)] \\ &= \sum_{t=r}^T \ln [g_{1t} - g_{2t}], \end{aligned} \quad (\text{A.1})$$

where the final step is notational. The derivatives for  $l_t$  are

$$\begin{aligned} \frac{\partial l_t}{\partial \gamma} &= \frac{1}{\theta_t^\gamma} \frac{[-\eta_t (d_t - 1)^\gamma \ln\left(\frac{d_t - 1}{\theta_t}\right)] g_{1t} + [(d_t + 1)^\gamma \ln\left(\frac{d_t + 1}{\theta_t}\right)] g_{2t}}{g_{1t} - g_{2t}} \\ \frac{\partial l_t}{\partial \theta_t} &= \frac{\gamma}{\theta_t} \frac{\eta_t \left(\frac{d_t - 1}{\theta_t}\right)^\gamma g_{1t} - \left(\frac{d_t + 1}{\theta_t}\right)^\gamma g_{2t}}{g_{1t} - g_{2t}} \\ \frac{\partial l_t}{\partial \psi} &= \frac{\partial l_t}{\partial \theta_t} \cdot \frac{\partial \theta_t}{\partial \psi} = \mathbf{z}'_t \frac{\partial l_t}{\partial \theta_t}. \end{aligned} \quad (\text{A.2})$$

The conditional expectations corresponding to those in (3) and required for EM-algorithms are of the form

$$\begin{aligned} E(D_t | d_t = 0, \Delta_{t-1}) &= c \frac{P(1 + \gamma^{-1}, 1/\theta_t^\gamma)}{1 - e^{-1/\theta_t}} \\ E(D_t | d_t = k, \Delta_{t-1}) &= c \frac{P\left[1 + \gamma^{-1}, \left(\frac{k+1}{\theta_t}\right)^\gamma\right] - P\left[1 + \gamma^{-1}, \left(\frac{k-1}{\theta_t}\right)^\gamma\right]}{e^{-(k-1)^\gamma/\theta_t} - e^{-(k+1)^\gamma/\theta_t}}, \end{aligned} \quad (\text{A.3})$$

where  $c = \theta_t \Gamma(1 + \gamma^{-1})$  and  $P(.,.)$  is the incomplete gamma function (e.g., Press et al., 1992, p. 209).

Bauwens and Giot (2001, pp. 101-104) give the Burr density function:

$$f(D) = \frac{\gamma}{\theta} \left( \frac{D}{\theta} \right)^{\gamma-1} \left[ 1 + \lambda \left( \frac{D}{\theta} \right)^{\gamma} \right]^{-1-1/\lambda} \quad (\text{A.4})$$

with mean and variance

$$\begin{aligned} E(D) &= \theta \frac{\Gamma(1+\gamma^{-1})\Gamma(\lambda^{-1})}{\lambda^{1+\gamma^{-1}}\Gamma(1+\lambda^{-1})}, \quad \text{for } \gamma/\lambda > 1 \\ V(D) &= \theta^2 \frac{1}{\lambda^{1+2\gamma^{-1}}\Gamma(1+\lambda^{-1})} \\ &\times \left[ \Gamma(1+2\gamma^{-1})\Gamma(\lambda^{-1}+2\gamma^{-1}) - \frac{\Gamma(1+\gamma^{-1})\Gamma(\lambda^{-1})}{\lambda\Gamma(1+\lambda^{-1})} \right], \quad \text{for } \gamma/\lambda > 2. \end{aligned} \quad (\text{A.5})$$

The survival and hazard functions are

$$\begin{aligned} \bar{F}(D) &= \left[ 1 + \lambda \left( \frac{D}{\theta} \right)^{\gamma} \right]^{-1/\lambda} = c^{-\lambda^{-1}} \\ h(D) &= \frac{\gamma}{\theta} \left( \frac{D}{\theta} \right)^{\gamma-1} \left[ 1 + \lambda \left( \frac{D}{\theta} \right)^{\gamma} \right]^{-1}. \end{aligned} \quad (\text{A.6})$$

Using (A.4)-(A.6) it is then possible to obtain the log-likelihood function corresponding to (6) and then to obtain ML estimates. The derivatives of  $\bar{F}(D)$  with respect to  $\gamma$ ,  $\theta$  (and  $\boldsymbol{\psi}$ ) and  $\lambda$  make up the score vector and are given by

$$\begin{aligned} \frac{\partial \bar{F}(D)}{\partial \gamma} &= - \left( \frac{D}{\theta} \right)^{\gamma} \ln \left( \frac{D}{\theta} \right) c^{-\lambda^{-1}-1} \\ \frac{\partial \bar{F}(D)}{\partial \theta_t} &= \frac{\gamma}{\theta} \left( \frac{D}{\theta} \right)^{\gamma} c^{-\lambda^{-1}-1} \\ \frac{\partial \bar{F}(D)}{\partial \boldsymbol{\psi}} &= \frac{\partial \bar{F}(D)}{\partial \theta_t} \cdot \frac{\partial \theta_t}{\partial \boldsymbol{\psi}} = \mathbf{z}_t' \frac{\partial \bar{F}(D)}{\partial \theta_t} \\ \frac{\partial \bar{F}(D)}{\partial \lambda} &= \frac{1}{\lambda} \left[ \lambda^{-1} \ln(c) - c^{-1} \left( \frac{D}{\theta} \right)^{\gamma} \right] c^{-\lambda^{-1}}. \end{aligned} \quad (\text{A.7})$$

Table A1: Bias measures (times 100) of estimators for data generated by an exponential model in Monte Carlo experiment,  $T = 5000$ .

True			$\alpha_0$				$\alpha_1$				$\beta_1$			
$\alpha_0$	$\alpha_1$	$\beta_1$	C	D	G	EM	C	D	G	EM	C	D	G	EM
2.5	0.2	0.15	1.50	3.17	-3.02	-2.43	-0.03	-0.90	-1.18	-0.51	-0.34	1.12	1.52	0.82
2.5	0.2	0.20	1.82	2.65	-2.70	-2.15	-0.03	-0.77	-1.00	-0.43	-0.38	0.97	1.28	0.69
2.5	0.2	0.25	2.20	2.42	-2.41	-1.97	-0.02	-0.67	-0.85	-0.40	-0.43	0.78	1.10	0.63
2.5	0.3	0.15	1.10	2.83	-1.54	-1.52	-0.02	-0.82	-1.30	-0.36	-0.20	0.85	1.24	0.48
2.5	0.3	0.20	1.31	2.33	-1.54	-1.49	-0.02	-0.70	-1.09	-0.31	-0.22	0.73	1.08	0.45
2.5	0.3	0.25	1.54	2.05	-1.32	-1.00	-0.02	-0.58	-0.89	-0.25	-0.24	0.57	0.88	0.31
2.5	0.4	0.15	0.97	2.54	-0.12	-1.04	-0.01	-0.71	-1.37	-0.23	-0.14	0.61	1.01	0.29
2.5	0.4	0.20	1.13	2.08	-0.15	-0.84	-0.01	-0.59	-1.13	-0.18	-0.15	0.51	0.85	0.23
2.5	0.4	0.25	1.30	1.71	-0.14	-0.51	-0.01	-0.47	-0.88	-0.15	-0.16	0.40	0.67	0.16
5	0.2	0.15	3.15	3.64	0.70	1.64	-0.03	-0.28	-0.34	-0.17	-0.35	0.09	0.15	-0.11
5	0.2	0.20	3.88	3.64	1.18	2.50	-0.03	-0.23	-0.30	-0.14	-0.41	0.03	0.08	-0.21
5	0.2	0.25	4.76	3.49	2.24	3.36	-0.02	-0.21	-0.26	-0.13	-0.48	0.01	-0.04	-0.27
5	0.3	0.15	2.28	2.90	0.80	1.29	-0.02	-0.25	-0.38	-0.13	-0.21	0.12	0.20	-0.05
5	0.3	0.20	2.73	2.77	1.20	1.90	-0.02	-0.22	-0.31	-0.11	-0.23	0.08	0.12	-0.10
5	0.3	0.25	3.27	2.82	1.62	2.81	-0.02	-0.19	-0.26	-0.10	-0.26	0.04	0.06	-0.17
5	0.4	0.15	2.01	2.75	1.42	1.62	-0.02	-0.22	-0.38	-0.08	-0.15	0.07	0.16	-0.08
5	0.4	0.20	2.35	2.45	1.54	2.07	-0.02	-0.18	-0.32	-0.07	-0.16	0.06	0.13	-0.11
5	0.4	0.25	2.76	2.56	1.96	3.07	-0.02	-0.15	-0.25	-0.05	-0.16	0.02	0.06	-0.17
10	0.2	0.15	6.48	5.99	5.39	7.33	-0.03	-0.10	-0.11	-0.07	-0.37	-0.19	-0.25	-0.40
10	0.2	0.20	8.04	7.19	6.83	9.84	-0.02	-0.08	-0.10	-0.06	-0.43	-0.27	-0.31	-0.52
10	0.2	0.25	9.95	8.31	9.84	12.62	-0.02	-0.07	-0.06	-0.06	-0.50	-0.32	-0.38	-0.63
10	0.3	0.15	4.66	4.78	4.06	6.05	-0.02	-0.09	-0.11	-0.06	-0.21	-0.11	-0.12	-0.27
10	0.3	0.20	5.64	5.37	4.75	7.80	-0.02	-0.07	-0.10	-0.05	-0.24	-0.14	-0.14	-0.33
10	0.3	0.25	6.81	6.14	5.87	10.11	-0.02	-0.07	-0.08	-0.04	-0.27	-0.17	-0.18	-0.40
10	0.4	0.15	4.27	4.50	3.73	5.98	-0.04	-0.09	-0.11	-0.03	-0.15	-0.08	-0.07	-0.23
10	0.4	0.20	5.36	5.25	4.34	7.70	-0.06	-0.10	-0.10	-0.03	-0.16	-0.09	-0.08	-0.27
10	0.4	0.25	7.18	6.73	5.09	10.03	-0.11	-0.15	-0.07	-0.02	-0.17	-0.11	-0.11	-0.33

Table A2: MSE measures (times 100) of estimators in Monte Carlo experiment,  $T = 5000$ .  
Data are generated as exponentially distributed.

True			$\alpha_0$				$\alpha_1$				$\beta_1$			
$\alpha_0$	$\alpha_1$	$\beta_1$	C	D	G	EM	C	D	G	EM	C	D	G	EM
2.5	0.2	0.15	6.38	7.26	6.89	7.01	0.03	0.04	0.04	0.04	0.48	0.54	0.54	0.54
2.5	0.2	0.20	7.05	7.79	7.33	7.57	0.03	0.04	0.04	0.03	0.47	0.51	0.50	0.51
2.5	0.2	0.25	7.77	8.29	7.57	8.15	0.03	0.04	0.04	0.03	0.45	0.47	0.45	0.48
2.5	0.3	0.15	3.49	3.83	3.62	3.70	0.04	0.04	0.05	0.04	0.21	0.22	0.23	0.22
2.5	0.3	0.20	3.87	4.09	3.84	4.05	0.04	0.04	0.05	0.04	0.20	0.21	0.21	0.21
2.5	0.3	0.25	4.28	4.47	4.07	4.44	0.04	0.04	0.04	0.04	0.19	0.20	0.19	0.20
2.5	0.4	0.15	2.43	2.60	2.48	2.57	0.04	0.05	0.06	0.04	0.11	0.12	0.12	0.12
2.5	0.4	0.20	2.72	2.86	2.67	2.79	0.04	0.05	0.05	0.04	0.11	0.12	0.11	0.11
2.5	0.4	0.25	3.04	3.14	2.87	3.09	0.04	0.04	0.05	0.04	0.10	0.11	0.10	0.11
5	0.2	0.15	25.51	26.38	25.92	26.03	0.03	0.03	0.03	0.03	0.48	0.49	0.49	0.49
5	0.2	0.20	28.19	29.05	28.49	28.76	0.03	0.03	0.03	0.03	0.47	0.48	0.47	0.48
5	0.2	0.25	31.05	31.84	31.46	31.70	0.03	0.03	0.03	0.03	0.45	0.46	0.46	0.46
5	0.3	0.15	13.95	14.27	14.17	14.17	0.04	0.04	0.04	0.04	0.21	0.21	0.21	0.21
5	0.3	0.20	15.48	15.79	15.55	15.76	0.04	0.04	0.04	0.04	0.20	0.20	0.20	0.21
5	0.3	0.25	17.12	17.32	17.24	17.42	0.04	0.04	0.04	0.04	0.19	0.20	0.20	0.20
5	0.4	0.15	9.73	9.91	9.83	9.89	0.04	0.04	0.04	0.04	0.11	0.11	0.11	0.11
5	0.4	0.20	10.91	11.03	11.00	11.08	0.04	0.04	0.04	0.04	0.11	0.11	0.11	0.11
5	0.4	0.25	12.23	12.29	12.15	12.44	0.04	0.04	0.04	0.04	0.10	0.10	0.10	0.11
10	0.2	0.15	101.95	103.00	102.42	103.02	0.03	0.03	0.03	0.03	0.48	0.48	0.48	0.48
10	0.2	0.20	112.68	113.32	113.34	113.87	0.03	0.03	0.03	0.03	0.47	0.47	0.47	0.47
10	0.2	0.25	124.16	125.31	113.13	126.74	0.03	0.03	0.03	0.03	0.45	0.45	0.44	0.45
10	0.3	0.15	55.77	56.23	55.78	56.68	0.04	0.04	0.04	0.04	0.21	0.21	0.21	0.21
10	0.3	0.20	61.90	62.16	61.79	63.08	0.04	0.04	0.04	0.04	0.20	0.20	0.20	0.20
10	0.3	0.25	68.54	68.50	68.38	70.20	0.04	0.04	0.04	0.04	0.19	0.19	0.19	0.20
10	0.4	0.15	39.38	39.53	38.95	39.69	0.04	0.04	0.04	0.04	0.11	0.11	0.11	0.11
10	0.4	0.20	44.84	45.05	43.67	45.04	0.05	0.05	0.04	0.04	0.11	0.11	0.11	0.11
10	0.4	0.25	52.01	52.18	48.62	51.23	0.05	0.05	0.04	0.04	0.11	0.11	0.10	0.11



Table A3: Bias and MSE (times 100) sof estimators for the exponential model with  $\alpha_0 = 2.5, T = 50000$ .

True		$\alpha_0$				$\alpha_1$				$\beta_1$			
$\alpha_1$	$\beta_1$	C	D	G	EM	C	D	G	EM	C	D	G	EM
Bias													
0.2	0.15	-0.15	1.36	-3.73	-3.99	0.02	-0.84	-0.57	-0.44	0.02	1.50	1.39	1.13
0.2	0.20	-0.13	0.56	-3.79	-4.14	0.02	-0.72	-0.48	-0.38	0.02	1.39	1.26	1.10
0.2	0.25	-0.10	-0.15	-3.88	-4.14	0.02	-0.61	-0.41	-0.32	0.01	1.27	1.14	1.02
0.3	0.15	-0.14	1.60	-3.47	-2.86	0.02	-0.77	-0.47	-0.30	0.02	1.05	0.99	0.70
0.3	0.20	-0.12	0.86	-3.48	-2.94	0.02	-0.65	0.39	-0.26	0.01	0.95	0.88	0.67
0.3	0.25	-0.09	0.31	-3.36	-2.78	0.02	-0.53	-0.31	-0.21	0.00	0.83	0.74	0.58
0.4	0.15	-0.12	1.41	-3.46	-2.20	0.02	-0.67	-0.37	-0.19	0.01	0.76	0.72	0.45
0.4	0.20	-0.10	0.79	-3.37	-2.17	0.02	-0.55	-0.29	-0.16	0.00	0.66	0.61	0.40
0.4	0.25	-0.07	0.32	-3.24	-2.06	0.02	-0.43	-0.22	-0.12	0.00	0.55	0.50	0.34
MSE													
0.2	0.15	0.64	0.73	0.83	0.85	0.00	0.01	0.01	0.01	0.05	0.08	0.07	0.07
0.2	0.20	0.71	0.79	0.89	0.93	0.00	0.01	0.01	0.00	0.05	0.07	0.07	0.06
0.2	0.25	0.79	0.86	0.93	1.00	0.00	0.01	0.01	0.00	0.05	0.07	0.06	0.06
0.3	0.15	0.35	0.40	0.48	0.45	0.00	0.01	0.01	0.00	0.02	0.03	0.03	0.03
0.3	0.20	0.39	0.42	0.51	0.50	0.00	0.01	0.01	0.00	0.02	0.03	0.03	0.03
0.3	0.25	0.43	0.46	0.53	0.53	0.00	0.01	0.00	0.00	0.02	0.03	0.03	0.02
0.4	0.15	0.24	0.28	0.37	0.30	0.00	0.01	0.01	0.00	0.01	0.02	0.02	0.01
0.4	0.20	0.27	0.29	0.38	0.33	0.00	0.01	0.01	0.00	0.01	0.02	0.01	0.01
0.4	0.25	0.31	0.32	0.40	0.36	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.01

Table A4: Bias measures (times 100) of estimators for data generated by a Weibull model in Monte Carlo experiment,  $T = 5000$ .

True			$\alpha_0$				$\alpha_1$				$\beta_1$			
$\alpha_0$	$\alpha_1$	$\beta_1$	C	D	G	EM	C	D	G	EM	C	D	G	EM
2.5	0.2	0.15	1.67	4.48	-1.52	-1.18	-0.04	-0.89	-1.23	-0.52	-0.37	0.90	1.35	0.65
2.5	0.2	0.20	1.97	3.65	-1.57	-1.15	-0.04	-0.80	-1.08	-0.48	-0.41	0.84	1.22	0.60
2.5	0.2	0.25	2.33	3.24	-1.39	-1.04	-0.04	-0.69	-0.91	-0.44	-0.45	0.68	1.00	0.53
2.5	0.3	0.15	1.22	3.46	-0.31	-1.01	-0.04	-0.96	-1.46	-0.48	-0.21	0.89	1.19	0.54
2.5	0.3	0.20	1.42	2.91	-0.38	-1.01	-0.04	-0.81	-1.23	-0.47	-0.23	0.74	1.02	0.53
2.5	0.3	0.25	1.64	2.38	-0.43	-0.85	-0.03	-0.72	-1.02	-0.39	-0.25	0.66	0.84	0.42
2.5	0.4	0.15	1.07	2.87	1.11	-0.83	-0.03	-0.94	-1.63	-0.47	-0.15	0.76	0.99	0.48
2.5	0.4	0.20	1.22	2.21	0.76	-0.72	-0.03	-0.81	-1.34	-0.40	-0.16	0.67	0.85	0.40
2.5	0.4	0.25	1.37	1.80	0.56	-0.47	-0.02	-0.68	-1.09	-0.35	-0.16	0.55	0.70	0.32
5	0.2	0.15	3.44	4.36	0.93	1.59	-0.04	-0.33	-0.38	-0.24	-0.38	0.01	0.09	-0.08
5	0.2	0.20	4.11	4.31	1.45	2.20	-0.04	-0.29	-0.31	-0.22	-0.43	-0.04	-0.01	-0.14
5	0.2	0.25	4.92	4.78	1.90	3.27	-0.04	-0.26	-0.26	-0.20	-0.48	-0.13	-0.08	-0.25
5	0.3	0.15	2.49	3.30	1.04	1.25	-0.04	-0.37	-0.42	-0.25	-0.22	0.13	0.13	0.01
5	0.3	0.20	2.92	3.24	1.09	1.74	-0.04	-0.31	-0.36	-0.21	-0.24	0.08	0.10	-0.04
5	0.3	0.25	3.43	3.41	1.36	2.60	-0.03	-0.28	-0.28	-0.19	-0.26	0.03	0.03	-0.12
5	0.4	0.15	2.17	2.88	1.51	1.58	-0.03	-0.37	-0.46	-0.24	-0.16	0.15	0.13	0.01
5	0.4	0.20	2.49	2.88	1.41	2.08	-0.03	-0.31	-0.36	-0.20	-0.17	0.10	0.09	-0.03
5	0.4	0.25	2.84	2.92	1.48	2.87	-0.02	-0.26	-0.29	-0.18	-0.17	0.06	0.05	-0.08
10	0.2	0.15	6.99	6.93	4.74	6.96	-0.04	-0.12	-0.10	-0.11	-0.39	-0.29	-0.28	-0.40
10	0.2	0.20	8.44	8.14	5.89	8.97	-0.04	-0.11	-0.08	-0.10	-0.44	-0.35	-0.34	-0.49
10	0.2	0.25	10.17	9.70	7.21	11.56	-0.04	-0.10	-0.07	-0.08	-0.50	-0.41	-0.38	-0.59
10	0.3	0.15	5.05	5.28	3.51	5.54	-0.04	-0.13	-0.10	-0.11	-0.22	-0.15	-0.16	-0.23
10	0.3	0.20	5.96	5.79	4.21	7.30	-0.04	-0.12	-0.08	-0.09	-0.25	-0.16	-0.18	-0.30
10	0.3	0.25	7.02	6.73	4.59	9.15	-0.03	-0.10	-0.06	-0.08	-0.27	-0.20	-0.19	-0.36
10	0.4	0.15	4.39	4.66	2.75	5.79	-0.03	-0.13	-0.08	-0.11	-0.16	-0.08	-0.09	-0.19
10	0.4	0.20	5.06	5.18	3.19	7.36	-0.03	-0.11	-0.02	-0.09	-0.17	-0.10	-0.12	-0.24
10	0.4	0.25	5.79	5.81	3.44	9.21	-0.02	-0.09	0.01	-0.07	-0.18	-0.12	-0.12	-0.28

Table A5: MSE measures (times 100) of estimators for data generated by a Weibull model in Monte Carlo experiment,  $T = 5000$ .

True			$\alpha_0$				$\alpha_1$				$\beta_1$			
$\alpha_0$	$\alpha_1$	$\beta_1$	C	D	G	EM	C	D	G	EM	C	D	G	EM
2.5	0.2	0.15	6.93	8.02	7.54	7.50	0.04	0.05	0.05	0.04	0.52	0.58	0.59	0.58
2.5	0.2	0.20	7.73	8.63	8.06	8.20	0.04	0.05	0.05	0.04	0.52	0.57	0.55	0.56
2.5	0.2	0.25	8.58	9.29	8.33	8.99	0.04	0.05	0.05	0.04	0.50	0.53	0.49	0.53
2.5	0.3	0.15	3.88	4.32	4.05	4.13	0.05	0.06	0.07	0.05	0.23	0.26	0.26	0.25
2.5	0.3	0.20	4.35	4.69	4.36	4.52	0.05	0.06	0.06	0.05	0.23	0.25	0.24	0.25
2.5	0.3	0.25	4.85	5.15	4.60	4.98	0.05	0.05	0.06	0.05	0.23	0.24	0.22	0.24
2.5	0.4	0.15	2.78	3.03	2.83	2.86	0.06	0.07	0.08	0.06	0.13	0.14	0.14	0.14
2.5	0.4	0.20	3.13	3.30	3.11	3.24	0.06	0.06	0.07	0.06	0.13	0.14	0.13	0.14
2.5	0.4	0.25	3.52	3.64	3.33	3.62	0.06	0.06	0.07	0.06	0.13	0.13	0.12	0.14
5	0.2	0.15	27.68	28.76	28.04	28.53	0.04	0.04	0.04	0.04	0.52	0.54	0.53	0.54
5	0.2	0.20	30.89	31.90	30.24	31.74	0.04	0.04	0.04	0.04	0.52	0.53	0.51	0.53
5	0.2	0.25	34.30	34.61	32.32	34.98	0.04	0.04	0.04	0.04	0.50	0.50	0.47	0.51
5	0.3	0.15	15.53	16.00	15.49	15.87	0.05	0.05	0.05	0.05	0.23	0.24	0.23	0.24
5	0.3	0.20	17.41	17.77	16.80	17.74	0.05	0.05	0.05	0.05	0.23	0.24	0.22	0.24
5	0.3	0.25	19.41	19.62	17.90	19.69	0.05	0.05	0.05	0.05	0.23	0.23	0.21	0.23
5	0.4	0.15	11.10	11.33	10.98	11.29	0.06	0.06	0.06	0.06	0.13	0.13	0.13	0.13
5	0.4	0.20	12.53	12.69	12.02	12.83	0.06	0.06	0.06	0.06	0.13	0.13	0.12	0.13
5	0.4	0.25	14.09	14.21	13.07	14.46	0.06	0.06	0.06	0.06	0.13	0.13	0.12	0.13
10	0.2	0.15	110.67	111.45	109.69	111.87	0.04	0.04	0.04	0.04	0.52	0.53	0.52	0.53
10	0.2	0.20	123.51	123.78	118.99	124.81	0.04	0.04	0.04	0.04	0.52	0.52	0.50	0.52
10	0.2	0.25	137.17	137.78	128.08	139.06	0.04	0.04	0.04	0.04	0.50	0.51	0.47	0.51
10	0.3	0.15	62.10	62.36	60.59	62.79	0.05	0.05	0.05	0.05	0.23	0.23	0.23	0.24
10	0.3	0.20	69.63	69.73	66.31	70.74	0.05	0.05	0.05	0.05	0.23	0.23	0.22	0.24
10	0.3	0.25	77.65	77.85	71.46	79.24	0.05	0.05	0.05	0.05	0.23	0.23	0.21	0.23
10	0.4	0.15	44.40	44.83	43.28	45.25	0.06	0.06	0.06	0.06	0.13	0.13	0.13	0.13
10	0.4	0.20	50.12	50.30	47.97	51.60	0.06	0.06	0.06	0.06	0.13	0.13	0.12	0.13
10	0.4	0.25	56.36	56.54	52.61	58.76	0.06	0.06	0.06	0.06	0.13	0.13	0.12	0.13

## REFERENCES

- Bauwens, L. and Giot, P. (2001). *Econometric Modelling of Stock Market Intraday Activity*. Kluwer, Dordrecht.
- Cameron, A.C. and Trivedi, P.K. (1998). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge.
- Cappé, O., Douc, R., Moulines, E. and Robert, C. (2002). On the Convergence of the Monte Carlo Maximum Likelihood Method for Latent Variable Models. *Scandinavian Journal of Statistics* **29**, 615-635.
- Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm (with discussion). *Journal of the Royal Statistical Society* **B39**, 1-38.
- Easley, D. and O'Hara, M. (1992). Time and the Process of Security Price Adjustment. *Journal of Finance* **47**, 576-605.
- 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 Monfort, A. (1991). Simulation Based Econometrics in Models with Heterogeneity. *Annales d'Economie et de Statistique* **20/1**, 69-107.
- Kulldorff, G. (1961). *Contributions to the Theory of Estimation from Grouped and Partially Grouped Samples*. PhD thesis. Almqvist & Wicksell, Stockholm.
- Lee, L-F. (1997). Simulated Maximum Likelihood Estimation of Dynamic Discrete Choice Statistical Models: Some Monte Carlo Results. *Journal of Econometrics* **82**, 1-35.
- Little, R.J.A. and Rubin, D.B. (1987). *Statistical Analysis with Missing Data*. Wiley, New York.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). *Numerical Recipes*. Cambridge University Press, Cambridge.