

INSTITUTE OF ECONOMIC STUDIES

WORKING PAPER SERIES

W00:20

December 2000

Monitoring the trading intensity of a stock market under infrequent trading

Helgi Tomasson

Address: Department of Economics
University of Iceland
Oddi at Sturlugata
101 Reykjavik
Iceland

Email: helgito@hi.is

Abstract

The durations between trading on a stock market are monitored. The durations are assumed to be exponentially distributed with a time-varying parameter. A gamma prior is used for the parameter, and as data are observed, the Bayes rule is used for updating. Given data, the parameters of the gamma distribution can be estimated. The posterior predictive duration is Pareto-distributed. The method is applied to Icelandic stock market data. Results suggest that most stocks are rather illiquid.

1 Introduction

When a new stock market opens it is of interest to have an objective expression concerning the intensity of trading. Traditionally, when activity is reported, it is customary to give a historical review, e.g., number of transactions last year, last 6 months, last 3 months, etc. The procedure introduced in this paper is designed to complement traditional reporting. In this paper an on-line monitoring method of duration between transactions is suggested. The approach taken is a stochastic one, where duration between trades is assumed to follow a given probability distribution. Assuming that transactions occur randomly, the choice of the exponential distribution is, due to its memoryless property, a natural base line. By the memoryless property, it is understood that past durations do not affect future durations. The analysis of durations is essentially analysis of waiting times. It is, however, unrealistic to assume that the trading intensity is constant over time. In practice trading intensity is also unknown. Therefore, some approach for dealing with this uncertainty is necessary.

The approach taken in this paper is to assume that durations are locally exponentially distributed with time-varying intensity. The setup is similar to Kalman filtering (Harvey, 1989), a filtering process consisting of recursively updating the estimated intensity as observations on durations become available. The same concepts as in state-space modeling apply here, i.e., predicted state, filtered state, predicted observation and the corresponding variances. The Kalman filter for a linear, normal, state-space model is based on applying the Bayes rule, recursively updating the estimated state as observations arrive. To achieve analytical results when the Bayes rule is applied requires that the model and the prior come from a conjugate family; for details see, e.g., Bernardo and Smith (1994), Zellner (1971) or some statistical textbooks like Casella and Berger (1990). The analytical form of the Kalman-filter updating and filtering equations is due to the fact that for normal observations, the conjugate prior is the normal distribution. The conjugate prior for the exponential distribution is the gamma distribution. As the durations are assumed to be locally exponential, it is convenient to let the intensity be gamma-distributed. In Section 2 a version of this idea is reviewed; for details see, e.g., Harvey (1989) . Precisely analogous to the normal state-space model, the unknown intensity is assigned a prior distribution from the gamma family, the variance of the predicted

state is taken as the variance of the gamma prior up-scaled with a smoothing constant. This constant represents the uncertainty about the future. This amounts to scaling down the parameters of the prior distribution. Having obtained an observed duration, it is straightforward to apply the Bayes rule and obtain the posterior distribution of the intensity, which will result in a gamma posterior-distribution. It can be shown that the predictive distribution of a duration, conditional on the last duration, will be a Pareto distribution with parameters given by the posterior gamma distribution. The Pareto distribution is a heavy-tailed distribution with only a finite number of moments(Mood et al., 1974). The parameters of the posterior gamma decide how many moments there exist. The smoothing constant can be estimated using some criterion, e.g., maximum likelihood.

The existence of moments of the Pareto distribution can be interpreted. More moments means less extreme values. In other words if a standard deviation does not exist, very long waiting times are plausible. If the standard deviation exists, the probability of exceeding a given bound can be calculated, e.g., by use of Chebychev inequality(Mood et al., 1974). If the higher moments exist, e.g., the fourth moment, this indicates that the standard deviation can be consistently estimated by use of the usual sample estimator. However, it is straightforward to calculate quantiles of the Pareto distribution which is useful for getting an objective idea on extreme values.

In Section 3, the method of Section 2 is applied to data from the Iceland Stock Exchange. The results indicate that the distribution of durations between transactions in Icelandic stocks is in general quite heavy-tailed, i.e., Icelandic stocks rather illiquid.

2 Stochastic intensity

In this section an on-line monitoring procedure is described. The idea is based on state-space modeling, i.e., a pair of equations, a measurement equation and a state equation. The measurements are assumed to be observations on a latent state of the system, and the state equation describes the dynamics of the latent state. For the normal model this is frequently monitored with the Kalman-filter recursions. As an example of a normal model, the measurement

equation is given by:

$$y_t = \mu_t + e_t \quad e_t \text{ i.i.d } N(0, \sigma_e^2) \quad (1)$$

and the state equation

$$\mu_t = \mu_{t-1} + \varepsilon_t \quad \varepsilon_t \text{ i.i.d. } N(0, \sigma_\varepsilon^2) \quad (2)$$

The model described by equations (1) and (2) is a noisy, normal, random-walk. The memoryless property of the random walk, i.e., only the last observation matters is an attractive feature in many applications, e.g., for the analysis of stock market data, where the efficient market hypothesis states that today's price is the best prediction of tomorrow's prices. As measurements arrive, information on μ_t is updated. The Kalman-filter recursions consist of the prediction equations and updating equations. The prediction equations are:

$$\mu_{t|t-1} = \mu_{t-1|t-1} \quad (3)$$

$$p_{t|t-1} = p_{t-1|t-1} + \sigma_\varepsilon^2 \quad (4)$$

$$y_{t|t-1} = \mu_{t|t-1} \quad (5)$$

$$f_{t|t-1} = p_{t|t-1} + \sigma_e^2 \quad (6)$$

Given information at time-point $t - 1$, $p_{t|t-1}$ and $p_{t-1|t-1}$ in equation (4) denote the variance of the unknown state at time t and $t - 1$, respectively. As new measurements of y_t arrive, one applies the updating, or filtering, equations:

$$\mu_{t|t} = \mu_{t|t-1} + p_{t|t-1}/f_{t|t-1}(y_t - \mu_{t|t-1}) \quad (7)$$

$$p_{t|t} = p_{t|t-1} - p_{t|t-1}^2/f_{t|t-1} \quad (8)$$

The updating equations are derived by use of the Bayes rule. The analytical form of equations (3) through (8) follows from the fact that both state and measurement are normal, i.e., because the conjugate prior for the normal distribution is the normal distribution. According to equation (3), the prediction for the next state is today's state, but the variance for the predicted state is larger than the one for the filtered (current) state. Models based on equations (1) through (8) are applied to analyze observed value dynamics in ISE in Tómasson (2000b) and Tómasson (2000a).

The purpose of this paper is to transfer these ideas for monitoring the intensity of trading. The duration between transactions is inherently positive, so a normal model is not appropriate. Due to its memoryless property the exponential distribution is a natural choice for the distribution of the duration between trades:

$$p(\delta_i | \lambda_i) = \lambda_i \exp(-\lambda_i \delta_i) \quad (9)$$

λ_i in equation (9) denotes the intensity of the trading process at transaction time number i . The intensity is assumed to evolve according to a state equation. An evolution scheme in the spirit of the Bayesian updating as in the Kalman filter is to assume

$$\lambda_i | \Delta_{i-1} \sim \Gamma(\alpha_{i|i-1}, \beta_{i|i-1}) \quad (10)$$

and

$$\lambda_i | \Delta_i \sim \Gamma(\alpha_{i|i}, \beta_{i|i}) \quad (11)$$

Equation (10) defines the distribution of intensity for transactions i conditional on past history, $\Gamma(\alpha_{i-1}, \beta_{i-1})$ denotes the gamma distribution and $\Delta_{i-1} = (\delta_{i-1}, \delta_{i-2}, \dots)$, denotes past values of duration between transactions. The numbers, $(\alpha_{i|i-1}, \beta_{i|i-1})$, and $(\alpha_{i|i}, \beta_{i|i})$ are the predicted and filtered estimates of the parameters, respectively. Similarly, equation (11) denotes the on-line distribution of the intensity. The gamma distribution is chosen here because it is the conjugate prior of the exponential distribution, and therefore recursive use of the Bayes rule will yield analytical results. The idea is that the expected future intensity should be the same as the current one, but the future variance conditional on current information should be greater than the current one. Standard results for the gamma distribution give:

$$E(\lambda_i | \Delta_i) = \frac{\alpha_{i|i}}{\beta_{i|i}} \quad (12)$$

and

$$V(\lambda_i | \Delta_i) = \frac{\alpha_{i|i}}{\beta_{i|i}^2} \quad (13)$$

We want $E(\lambda_i | \Delta_{i-1}) = E(\lambda_{i-1} | \Delta_{i-1})$, and due to the added uncertainty of prediction, $V(\lambda_i | \Delta_{i-1}) > V(\lambda_{i-1} | \Delta_{i-1})$, say by a constant factor ω , i.e.,

$$\omega V(\lambda_i | \Delta_{i-1}) = V(\lambda_{i-1} | \Delta_{i-1})$$

where $0 < \omega < 1$. By equation (13)

$$V(\lambda_{i-1} | \Delta_{i-1}) = \frac{\alpha_{i-1|i-1}}{\beta_{i-1|i-1}^2} = \omega V(\lambda_i | \Delta_{i-1}) = \frac{\alpha_{i|i-1}}{\beta_{i|i-1}^2} \quad (14)$$

and

$$E(\lambda_{i-1} | \Delta_{i-1}) = \frac{\alpha_{i-1|i-1}}{\beta_{i-1|i-1}} = E(\lambda_i | \Delta_{i-1}) = \frac{\alpha_{i|i-1}}{\beta_{i|i-1}} \quad (15)$$

Solving equations (14) and (15) for $\alpha_{i|i}$ and $\beta_{i|i}$ in terms of $\alpha_{i|i-1}$ and $\beta_{i|i-1}$ gives the prediction scheme:

$$\alpha_{i|i-1} = \omega \alpha_{i-1|i-1} \quad (16)$$

$$\beta_{i|i-1} = \omega \beta_{i-1|i-1} \quad (17)$$

The measurement equation is based on observing a duration, δ_i , assumed to be exponentially distributed with mean $\lambda_{i|i-1}$. Applying the Bayes rule to the measurement with the a priori distribution of $\lambda_{i|i-1}$ gives the posterior distribution:

$$\lambda_{i|i} \sim \Gamma(\alpha_{i|i}, \beta_{i|i}) \quad (18)$$

where

$$\alpha_{i|i} = \alpha_{i|i-1} + 1 = \omega \alpha_{i-1|i-1} + 1 \quad (19)$$

$$\beta_{i|i} = \beta_{i|i-1} + \delta_i = \omega \beta_{i-1|i-1} + \delta_i \quad (20)$$

The on-line estimate of the intensity is thus $\lambda_{i|i} = \alpha_{i|i} / \beta_{i|i}$. Integrating $\lambda_{i|i-1}$ out of the predictive density shows that the predictive density of δ_i conditional on Δ_{i-1} , i.e., as a function of $\alpha_{i|i-1}$ and $\beta_{i|i-1}$, is proportional to:

$$\frac{1}{(\beta_{i|i-1} + \delta_i)^{\alpha_{i|i-1} + 1}} \quad (21)$$

Equation (21) shows that the predictive distribution of δ_i conditional on Δ_{i-1} is a Pareto distribution with density:

$$f(\delta_i | \Delta_{i-1}) = \frac{\alpha_{i|i-1} \beta_{i|i-1}^{\alpha_{i|i-1}}}{(\beta_{i|i-1} + \delta_i)^{\alpha_{i|i-1} + 1}} \quad (22)$$

The parameters of the Pareto distribution in equation (22), $\alpha_{i|i-1}$ and $\beta_{i|i-1}$ may be used as measures of the liquidity of the market. The mean of the distribution given in equation (22) is:

$$\frac{\beta_{i|i-1}}{\alpha_{i|i-1} - 1} \quad (23)$$

which is finite if $\alpha_{i|i-1} > 1$. The variance of the distribution given by (22) is

$$\frac{\alpha_{i|i-1}\beta_{i|i-1}^2}{(\alpha_{i|i-1} - 1)^2(\alpha_{i|i-1} - 2)} \quad (24)$$

which is finite if $\alpha_{i|i-1} > 2$. In general m moment exist if $\alpha_{i|i-1} > m$.

The necessary parameters are the starting values, $\alpha_{0|0}, \beta_{0|0}$, of the gamma distribution, and ω , the speed of learning. Given these parameters and n observed durations, the values of these parameters, the system can be calculated recursively, and by equation (22) the log-likelihood can be calculated:

$$\log(L(\omega, \alpha_{0|0}, \beta_{0|0})) = \sum_{i=1}^n \log(f(\delta_i | \Delta_i)) \quad (25)$$

Maximizing (25) numerically yields a maximum-likelihood estimate.

Some algebra and a standard result from calculus

$$\lim_{x \rightarrow \infty} \left(\frac{x}{x+1}\right)^{\alpha x} = \exp(-\alpha)$$

shows that if

$$a_{i|i-1} \rightarrow \infty, \quad b_{i|i-1} \rightarrow \infty \text{ such that } \frac{a_{i|i-1}}{b_{i|i-1}} = \lambda \quad (26)$$

then the density in equation (22) approaches the exponential density with parameter λ . The condition $b_{i|i-1} \rightarrow \infty$ would also mean that the variance of the prior goes to zero, i.e., that the uncertainty about λ vanishes.

If $\omega = 1$ and $a_{0|0} = b_{0|0} = 0$, then

$$\frac{b_{n|n}}{a_{n|n}} = \frac{\sum_{i=1}^n 1}{\sum_{i=1}^n \delta_i} = \hat{\lambda}_{ML}$$

where λ_{ML} is the usual maximum-likelihood estimator for a pure exponential model. If $\omega < 1$, then $a_{\infty|\infty} = 1/(1 - \omega)$. The number $a_{\infty|\infty}$ can be interpreted because it says how many

moments of the Pareto distribution, equation (22), exist. Fewer moments means that more extreme durations can be expected. At each transaction time point equation, (22) can be used to calculate a 95% prediction interval for next duration, δ_i :

$$(\delta_l, \delta_u) = (\beta_{i|i-1}(1 - \frac{1}{(1 - 0.975)^{1/\alpha_{i|i-1}}}), \beta_{i|i-1}(1 - \frac{1}{(1 - 0.025)^{1/\alpha_{i|i-1}}})) \quad (27)$$

Equation (27) can be generalized to the case when the duration from transaction $i - 1$ is already δ , so it is known that $\delta_i > \delta$ by noting that:

$$F(\delta_i|\Delta_{i-1}, \delta_i > \delta) = 1 - \frac{(\beta_{i|i-1} + \delta)^{\alpha_{i|i-1}}}{(\beta_{i|i-1} + \delta_i)^{\alpha_{i|i-1}}} \quad (28)$$

At any point in time it is possible to monitor the trading intensity and give a probability statement about the duration to the next transaction. Direct interpretation of the condition of existing moments for the fat-tail nature of a distribution is perhaps not straightforward. To get an idea of the fat tail nature of the Pareto distribution one can apply standard order statistics (Mood et al., 1974) to calculate the median of maximum of the next 100 durations, assuming that they are independently Pareto distributed with $\alpha_{i|i-1}$ and $\beta_{i|i-1}$ given by equation (22).

The hazard rate of a waiting time distribution is given by:

$$h(\delta) = \frac{f(\delta)}{1 - F(\delta)}$$

where f and F denote the density function and the distribution functions, respectively. The Pareto density, equation (22), has its hazard function given by:

$$h(\delta|\Delta_{i-1}) = \frac{\alpha_{i|i-1}}{(\beta_{i|i-1} + \delta)} \quad (29)$$

i.e., the hazard is a decreasing function of δ . This means that at each point in time, the available information consists of past information Δ_{i-1} and the fact that no transaction has taken place after δ_{i-1} , i.e., δ_i will be greater than δ . This expresses a growing pessimism as time passes without transactions. The expected value of the next duration, δ_i , conditional on this information is

$$E(\delta_i|\Delta_{i-1}, \delta_i > \delta) = \frac{\beta_{i|i-1} + \delta}{\alpha_{i|i-1} - 1} \quad (30)$$

I.e., the expected remaining duration increases linearly with slope $1/(\alpha_{i|i-1} - 1)$. Both equation (29) and equation (30) illustrate that the risk of a transaction decreases as the time from the

last transaction increases. As $\alpha_{i|i-1} \rightarrow \infty$, however, the properties of exponential distribution, i.e., constant hazard rate and constant expected remaining duration become more apparent.

3 Some empirical results

The methods of Section 2 were applied to data from the Iceland Stock Exchange(ISE). The ISE is an automated, computerized market, which opened in 1991. Active trading started in 1993 and has gradually increased, in both the number of companies registered and trading intensity. By the year 2000, over 70 companies were registered. The accessible data consist of every registered transaction since the opening of the ISE. The data used in this paper consist of durations between all transactions from 1 January 1998 up to May tenth 2000, a total of roughly 63000 transactions on 38 stocks. Since the opening of the ISE up to 10 May 2000, roughly 88000 transactions on 76 stocks were observed. The observation period is 860 days long, of which the market was open on 586 days. The period from the first trade of the day until the last trade of the day was defined as the duration per day. The average period between the first transaction and the last one was 5.7 hours. This means that if the first transaction on a particular day was in stock happening to be the last one traded the day before, then those transactions were defined as simultaneous. All simultaneous transactions were considered as a single transaction, i.e., only positive durations were considered. Time was measured in 24-hours days, so a typical day at the market lasts 5.7 hours ≈ 0.24 days. A measured duration of one day is therefore about four days in real life.

Tables 1 and 2 show numerical results for the procedure described in Section 2. Table 1 shows results for stocks already on the market 1 January 1998, and Table 2 shows results for stocks entering the market during the observation period. The number k denotes the stock code. As new companies are registered on the ISE, they are assigned a number in chronological order, i.e., $k = 1$ denotes the first company registered. The number n denotes the number of transactions during the observation period. All stocks with 30 or more transactions in the observation period were included. The numbers $\bar{\delta}$ and s_{δ} denote the sample mean and sample standard deviation, respectively. The parameter ω was estimated by maximizing the likelihood function. This was done both for the case with the restriction $\alpha_{0|0} = \beta_{0|0} \approx 0$,

k	n	δ	s_δ	$\hat{\omega}$	$\alpha_{n n}$	$\beta_{n n}$	$Q_{0.5}(100)$	$l_2 - l_1$
1	328	0.44	0.85	0.72	3.60	1.56	6.53	65.79
2	2754	0.05	0.09	0.70	3.36	0.04	0.18	377.04
3	2385	0.06	0.12	0.51	2.05	0.08	4.71	744.26
4	2041	0.07	0.16	0.58	2.38	0.08	1.63	761.92
5	39	3.58	3.56	1.00	39.00	68.02	17.82	0.00
6	172	0.80	1.31	0.71	3.47	1.81	8.39	1.07
7	61	2.28	2.70	0.91	10.82	13.92	8.33	2.16
8	31	4.49	7.19	0.83	5.77	13.71	20.84	1.14
9	803	0.19	0.37	0.58	2.37	0.54	11.28	288.62
10	1220	0.12	0.20	0.67	3.00	0.57	4.16	110.42
12	266	0.52	0.89	0.65	2.87	0.87	7.59	35.00
13	582	0.25	0.33	0.79	4.79	1.97	4.23	18.61
14	454	0.31	0.43	0.75	4.05	0.98	3.02	48.89
15	3890	0.04	0.12	0.56	2.28	0.22	5.79	1190.23
16	629	0.22	0.34	0.78	4.63	1.78	4.11	48.95
17	71	1.93	2.63	0.55	2.20	1.71	57.38	16.22
18	488	0.28	0.47	0.74	3.88	2.26	7.78	77.32
19	799	0.18	0.28	0.83	5.93	1.66	2.41	39.90
20	463	0.31	0.56	0.71	3.44	0.86	4.10	3150.28
21	718	0.20	0.35	0.71	3.46	0.32	1.48	108.81
22	662	0.21	0.35	0.72	3.58	1.18	5.02	127.27
23	675	0.21	0.33	0.89	8.91	1.52	1.18	39.46
24	899	0.15	0.25	0.75	4.01	0.65	2.04	95.86
25	527	0.27	0.61	0.73	3.65	0.54	2.19	322.92
26	36	3.89	5.40	0.86	7.01	15.61	17.24	1.27
27	66	2.10	3.01	0.81	5.22	4.61	8.38	5.61
28	55	2.47	3.86	0.34	1.52	5.83	26273.52	24.61
29	114	1.08	1.44	0.81	5.13	1.83	3.43	22.45
30	206	0.67	1.00	0.95	18.42	10.17	3.18	2.14
31	460	0.30	0.48	0.78	4.45	3.48	8.73	133.43
32	563	0.25	0.44	0.76	4.15	0.62	1.83	113.58
33	39	3.53	5.54	0.87	7.42	14.55	14.72	1.04
34	44	2.49	4.13	0.42	1.73	4.53	1740.31	12.92
35	237	0.59	0.89	0.62	2.62	1.25	15.89	26.14
36	1590	0.09	0.21	0.62	2.63	0.14	1.80	360.13
38	1745	0.09	0.19	0.56	2.26	0.22	6.34	564.80

Table 1: Results for stocks registered before 1 January 1998.

k	n	$\bar{\delta}$	s_δ	$\hat{\omega}$	$\alpha_{n n}$	$\beta_{n n}$	$Q_{0.5}(100)$	$l_2 - l_1$
39	611	0.24	0.44	0.62	2.66	0.30	3.53	153.62
40	37	3.78	6.45	0.90	9.73	26.81	18.46	0.22
41	1552	0.09	0.26	0.60	2.48	0.14	2.28	817.86
43	304	0.47	1.02	0.69	3.20	1.73	10.25	110.63
44	141	0.99	1.95	0.52	2.10	0.90	43.05	66.59
45	70	1.95	2.64	0.89	9.14	15.00	11.26	0.51
47	1125	0.13	0.25	0.79	4.66	0.52	1.19	416.23
48	702	0.18	0.34	0.58	2.40	0.44	8.58	670.27
49	458	0.30	0.61	0.72	3.61	6.79	28.13	136.49
50	158	0.84	1.29	0.67	3.06	4.83	33.00	21.07
51	53	2.44	3.07	0.58	2.38	6.99	143.85	2.32
52	322	0.41	0.71	0.55	2.22	0.36	11.26	74.20
53	101	1.29	2.60	0.65	2.82	2.02	18.78	8.15
54	40	2.95	3.69	0.85	6.73	11.04	13.00	2.28
55	930	0.13	0.22	0.79	4.78	1.74	3.77	230.24
56	91	1.22	1.56	0.92	12.22	22.14	11.34	3.45
58	181	0.57	0.89	0.79	4.71	3.09	6.86	18.87
59	827	0.12	0.20	0.74	3.90	1.09	3.69	86.51
60	578	0.16	0.28	0.73	3.65	0.64	2.55	98.38
61	271	0.34	0.52	0.50	2.00	0.09	6.82	31.57
62	2177	0.05	0.19	0.52	2.10	0.10	5.03	940.19
63	3208	0.03	0.08	0.61	2.58	0.10	1.32	1009.76
64	90	1.01	1.38	0.83	5.99	9.22	13.14	8.75
65	1513	0.06	0.10	0.70	3.29	0.02	0.09	386.98
66	1524	0.06	0.14	0.71	3.48	0.22	1.02	423.86
67	87	0.95	1.21	0.77	4.26	7.17	19.75	9.14
68	171	0.47	1.01	0.59	2.41	0.70	13.45	43.81
70	1096	0.06	0.13	0.64	2.79	0.16	1.49	215.68
71	49	1.12	1.26	1.00	49.00	55.01	5.57	0.00
73	207	0.28	0.44	0.85	6.74	1.32	1.55	24.48
75	1419	0.04	0.18	0.60	2.49	0.02	0.25	731.24
76	1660	0.03	0.19	0.49	1.96	0.00	0.42	1014.78

Table 2: Results for stocks registered after 1 January 1998.

and for the case where $\alpha_{0|0}$ and $\beta_{0|0}$ were estimated as free parameters along with ω . Both cases are logical choices. If $\omega < 1$ then $\alpha_{i|i}$ quickly converges to $1/(1-\omega)$, so perhaps there is not much point in estimating $\alpha_{0|0}$. The variance in the prior distribution of the intensity, λ_i , is proportional to $1/\beta_i$. As it is natural to assume that in the beginning of the period little information is available on λ it is natural to choose the initial value such that the variance is big, i.e., choose $\beta_{0|0}$ close to zero, e.g., 10^{-20} . It is observed that the fixed parameter exponential model has $\omega = 1$ and $\lambda = \alpha_i/\beta_i$ for all i . This corresponds to setting $\beta_{0|0}$ to a large number. Therefore estimating the fixed parameter exponential model is the same as estimating a restricted version of the three-parameter model with parameter $(\omega, \alpha_{0|0}, \beta_{0|0})$. It is then possible to calculate the difference in the log-likelihood values for the two models to get an idea of the statistical significance of the deviation from the exponential model. This difference is denoted $l_2 - l_1$ in the tables. In most cases, the exponential model is clearly rejected. The maximum-likelihood estimate of ω , $\hat{\omega}$, is shown in the tables. In most cases, the difference of $\hat{\omega}$ between the case when the initial values are fixed close to zero and the case when they are estimated is marginal. In some cases the maximum-likelihood estimates of $\alpha_{0|0}$ and $\beta_{0|0}$ are very large, e.g., 10^{20} . In most of these cases, however, the likelihood is relatively flat, i.e., it changes by only a few units relative to large changes in $\alpha_{0|0}$ and $\beta_{0|0}$. The estimate $\hat{\omega}$ is shown in the tables. The situations where $l_2 - l_1$ is small, i.e., estimating ω is not a significant improvement over the exponential model ($\omega = 1$), is characterized by small sample sizes. In two cases, the maximum-likelihood estimate of ω is exactly one. This should not be surprising because ω was estimated with the restriction $\omega \leq 1$. In small samples, it is plausible that the unconstrained estimate of ω is greater than one, so the small sample distribution of the constrained maximum-likelihood estimator therefore has a point mass at one. The final estimates of α and β , $\alpha_{n|n}$ and $\beta_{n|n}$ are also shown. Judging from the the estimate of $\alpha_{n|n}$, which for many stocks is a low number, it is clear that liquidity is low on the ISE. Due to the low values of $\alpha_{n|n}$, only a few moments of the predictive distribution exist, i.e., predictive distributions are heavy-tailed. To get an objective measure of the extreme values, the median for the maximum of the 100 next durations was calculated, based on $\alpha_{n|n}$. This median is denoted by $Q_{0.5}(100)$ in the tables. For example, for stock number 3, the median

value for the maximum of the 100 next durations seems to be 10 times the expected value. All calculations were done using the statistical package R (Ihaka and Gentleman, 1996).

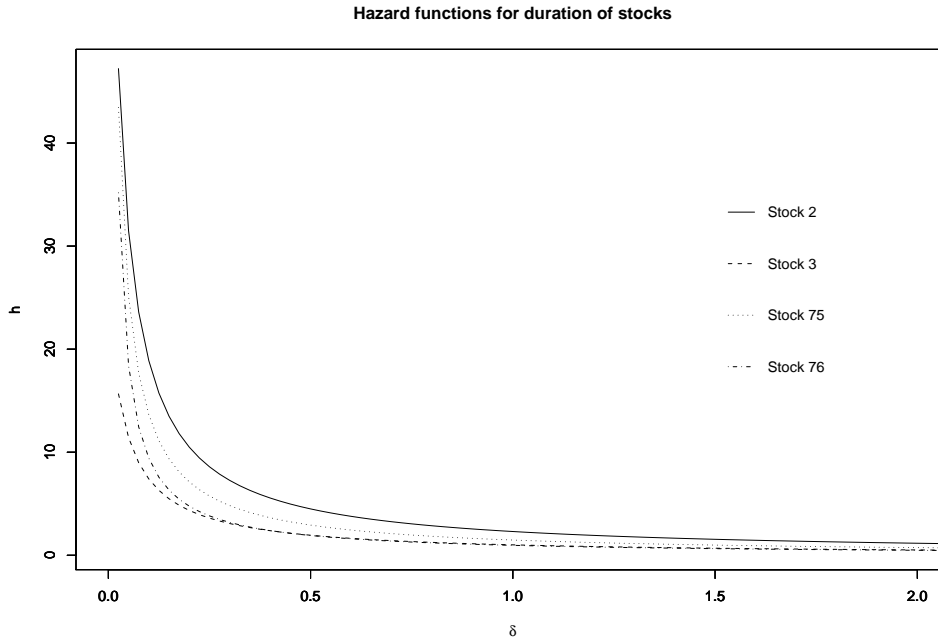


Figure 1: Hazard function for some stocks conditional on Δ_n .

In Figure 1, the hazard function conditional on the latest information for stocks 2, 3, 75 and 76 is plotted. Stocks 2 and 3 are well established on the market, whereas stocks 75 and 76 are registered at the end of the year 1999. The initial hazard for transaction is highest for stock 2 because it has the highest $\alpha_{n+1|n}$, but as time without a transaction increases, they all converge to zero. The time scale shown in Figure 1 goes from zero to two 24-hour days, which in real life is approximately one week. The common conclusion for all four stocks is that after one week of non-trading predictions for future transactions would be pessimistic.

Figure 2 shows the expected remaining duration as a function of δ . It is clear from looking at the slopes that stock number 2 is least affected by a period of non-trading. Stock 76 had $\hat{\omega} < 0.5$, this suggests that the predictive distribution does not have a finite expectation. The expected remaining duration is therefore a useless concept in such cases. It is, however, possible to analyse quantiles of the predictive distribution directly. Figures 3 and 4 are based on the estimates in Tables 1 and 2 and the formulas of Section 2. They show the median of the predictive distribution for a few stocks as a function of δ in relation to their 95% prediction

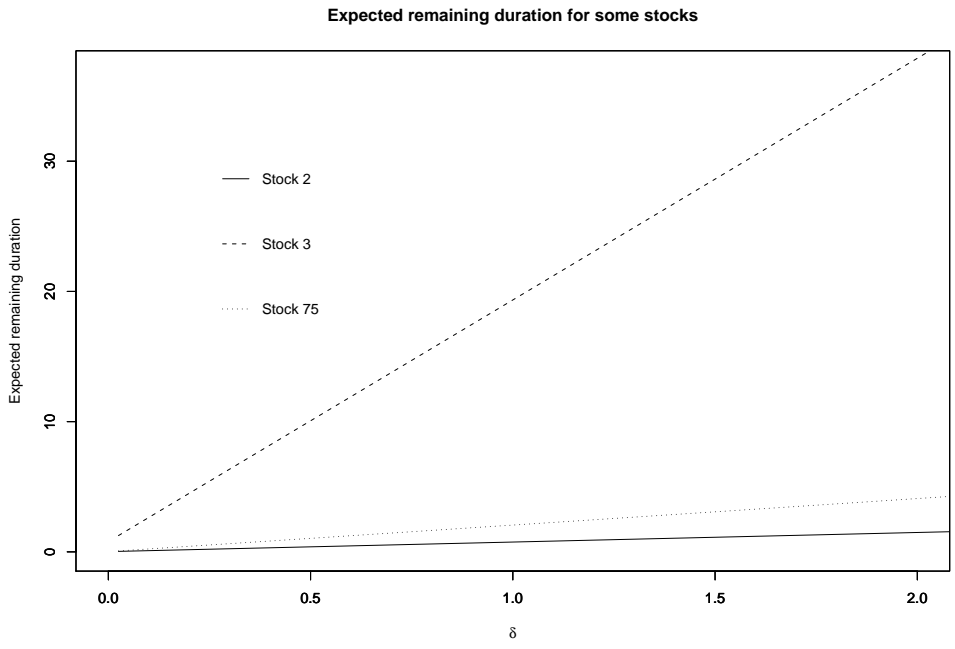


Figure 2: Expected remaining duration for three stocks

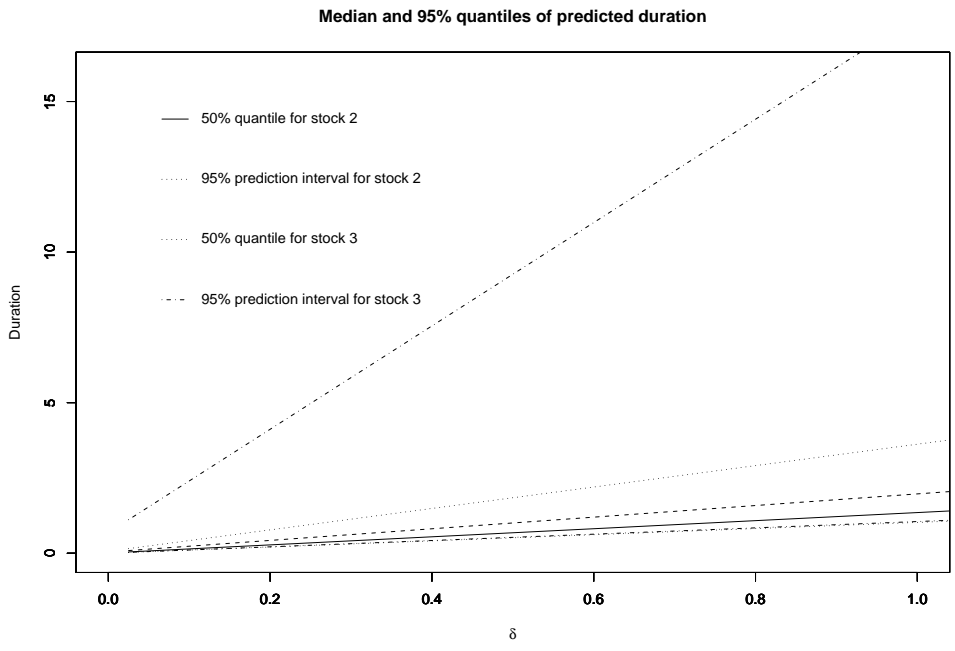


Figure 3: Median and 95% prediction intervals for stocks 2 and 3.

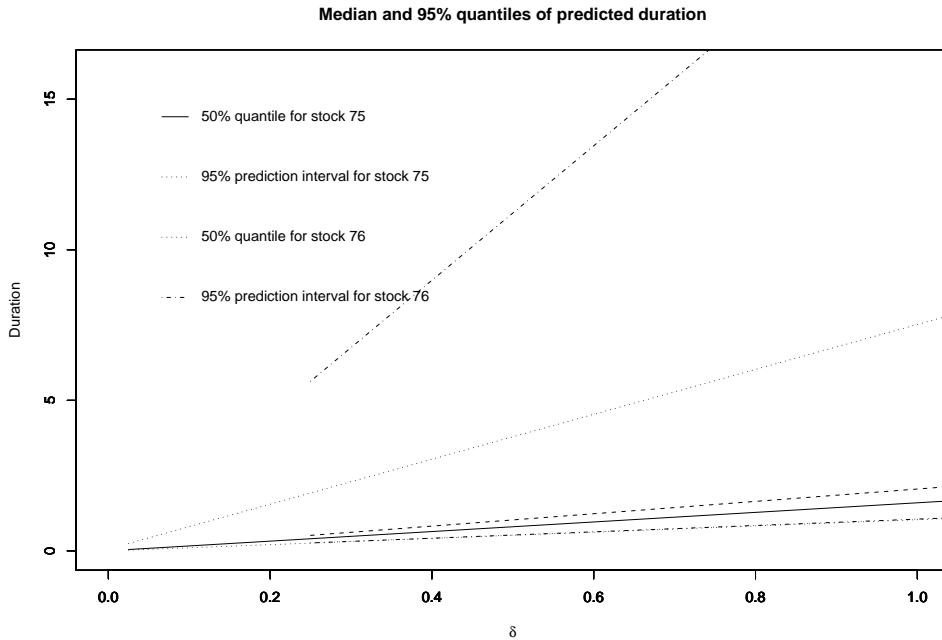


Figure 4: Median and 95% prediction intervals for stocks 75 and 76.

intervals. The heavy-tailed nature of the Pareto distribution is apparent from the asymmetric intervals. The stocks illustrated in these figures, stocks, 2,3,75 and 76, were chosen because they are among the more frequently traded ones in the observation period. Stocks 2 and 3 are among the oldest in the market, whereas stocks 75 and 76 are among the youngest. The stocks with very few transactions, say less than 100, were shown in Tables 1 and 2 to contrast them with the more frequently traded ones. It is clear that it is not possible to make very precise statements on stocks that are on average traded every five days.

4 Conclusion

In this paper, an objective method of monitoring durations in a financial market is proposed. The basis is a simple exponential model, modified by relaxing the assumption of a constant intensity. The computational process has a Bayesian flavor, i.e., the conjugate prior distribution is combined recursively with the observations. The interpretation need not, however, be Bayesian. The updating is controlled with a smoothing parameter ω , reminiscent of exponential smoothing shown in textbooks, see, e.g., Newbold (1995). The resulting predictive

distribution is easily interpretable. Expected duration can be calculated along with quantiles of the predictive distributions, order statistics, etc. This helps analysts and agents to get an objective view of liquidity on the market. It is also of potential value to market authorities, e.g., they can react if too many or too few transactions take place in a given period. The tool described in this paper is useful for comparing expected durations with observed ones as well as making predictions.

Acknowledgements

This research was supported by the Research Fund of Icelandic Banks and the Icelandic Science Fund. The author thanks the Iceland Stock Exchange for supplying data and the assistants at the Institute of Economic Studies at the University of Iceland for their help.

References

- Bernardo, J. M. and Smith, A. F. (1994). *Bayesian Theory*. John Wiley and Sons.
- Casella, G. and Berger, R. L. (1990). *Statistical Inference*. Wadworth and Brooks.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299–314.
- Mood, A. M., Graybill, F. A., and Boes, D. C. (1974). *Introduction to the Theory of Statistics*. McGraw-Hil, third edition.
- Newbold, P. (1995). *Statistics for Business and Economics*. Prentice-Hall, fourth edition.
- Tómasson, H. (2000a). Estimation of correlations in financial markets when trading is infrequent. Working Paper W00:18, The Institute of Economic Studies at the University of Iceland.

Tómasson, H. (2000b). Signal-noise decomposition in financial markets: an empirical stochastic process analysis for infrequent trading. Working paper W00:11, The Institute of Economic Studies at the University of Iceland.

Zellner, A. (1971). *An Introduction to Bayesian Econometrics*. John Wiley and Sons.

INSTITUTE OF ECONOMIC STUDIES WORKING PAPERS 1987-2000

Formerly Iceland Economic Papers Series

Editor Tryggvi Thor Herbertsson

A complete list of IoES Working Papers and Reprints can be accessed on our World Wide Web site at <http://www.ioes.hi.is>

W98:07 Thorarinn G. Petursson: The Representative Household's Demand for Money in a Cointegrated VAR Model

W98:08 Tryggvi Thor Herbertsson and Gylfi Zoega: Trade Surpluses and Life-cycle Saving Behaviour

W98:09 Thorvaldur Gylfason: Privatization, Efficiency, and Economic Growth

W98:10 Tor Einarsson and Milton Marquis: Uncertain Labor Quality and Asset Pricing

W98:11 J. Michael Orzag: Actuarial Discretion vs. Rules in Regulation of Pension Funds

W98:12 J. Michael Orzag and Dennis Snower: From Unemployment Benefits to Unemployment Support Accounts

W98:13 Sveinn Agnarsson, Axel Hall, Tryggvi Thor Herbertsson, Sigurdur Ingolfsson, Gylfi Magnusson, and Gylfi Zoega: EMU and the Icelandic Labor Market

W98:14 Thorvaldur Gylfason: Exports, Inflation, and Growth

W99:01 Alison L. Booth, Yu-Fu Chen, and Gylfi Zoega: Hiring and Firing: A Tale of Two Thresholds

W99:02 Martin Paldam and Gert Tinggaard Svendsen: Is Social Capital an Effective Smoke Condenser?

W99:03 J. Michael Orzag and Dennis J. Snower: Anatomy of Policy Complementarities

W99:04 Thorvaldur Gylfason and Tryggvi Thor Herbertsson: Does Inflation Matter for Growth?

W99:05 Marco Biamchi, Bjorn R. Gudmundsson, and Gylfi Zoega: Iceland's Natural Experiment in Supply-side Economics

W99:06 Haukur C. Benediktsson, Tryggvi Thor Herbertsson, Gylfi Magnússon, and Marta G. Skúladóttir: Generational Accounts for Iceland

W99:07 Axel Hall and Jon Thor Sturluson: Testing a CGE Model: The Tax Free Year in Iceland as a Natural Experiment

W99:08 Edmund S. Phelps: Equilibrium and Disequilibrium in 20th Century 'Macro': With Attention to the Share Price Boom of the 1990s

W99:09 Kenneth F. Wallis: Macroeconometric Modelling

W99:10 Tryggvi Thor Herbertsson, Marta G. Skúladóttir, and Gylfi Zoega: Three Symptoms and a Cure: A Contribution to the Economics of the Dutch Disease

W99:11 Tryggvi Thor Herbertsson and J. Michael Orzag: Issues in European Pension Reforms: Supplementary Pensions

W99:12 Tryggvi Thor Herbertsson, J. Michael Orszag and Peter R. Orszag: Population Dynamics and Convergence in Fertility Rates

W99:13 Ragnar Arnason: Costs of Fisheries Management: Theoretical and Practical Implications

W99:14 Ragnar Arnason: Economic Instruments to Achieve Ecosystem Objectives in Fisheries Management

W99:15 Ragnar Arnason: Property Rights as a Means of Economic Organization

W00:01 Jerry Coakley, Ana-Maria Fuertes and Gylfi Zoega: Testing the Persistence and Structuralist Theories of Unemployment

W00:02 Thrainn Eggertsson: Norms in Economics – With Special Reference to Economic Development

W00:03 Thorvaldur Gylfason: Growing Apart

W00:04 Jon Danielsson: The Emperor has no Clothes: Limits to Risk Modelling

W00:05 Thorolfur Matthiasson: The Icelandic Debate on the Case for a Fishing Fee: A Non-Technical Introduction

W00:06 Willem H. Buiter: Is Iceland an Optimal Currency Area?

W00:07 Alison L. Booth and Gylfi Zoega: Why do Firms Invest in General Training? ‘Good’ Firms and ‘Bad’ Firms as a Source of Monopsony Power

W00:08 Eduard Hochreiter: “Exchange rate regimes and capital mobility: Issues and some lessons from central and eastern European applicant countries”

W00:09 Thorvaldur Gylfason: Fix or Flex? Alternative Exchange Rate Regimes in an Era of Global Capital Mobility

W00:10 Thorvaldur Gylfason: Natural Resources, Education and Economic Development

W00:11 Helgi Tomasson: Signal-noise Decomposition in Financial Markets: An Empirical Stochastic Process Analysis for Infrequent Trading

W00:12 Thorolfur Matthiasson: Changing Rules for Regulation of Icelandic Fisheries

W00:13 E. Tumusiime-Mutebile: Economic Reforms and their Impact in Uganda

W00:14 Sveinn Agnarsson: Productivity in Icelandic Fish Processing Industry 1985 – 1995: A Comparison of Methods

W00:15 Sveinn Agnarsson: Development of Efficiency in Icelandic Fish Processing Firms: A DEA Approach

W00:16 Jon Danielsson, Bjorn N. Jorgensen and Casper G. de Vries: Risk Management and Regulation in Incomplete Markets

W00:17 Ragnar Arnason, Gylfi Magnusson and Sveinn Agnarsson: The Norwegian Spring Spawning Herring Fishery: A Stylised Game Model

W00:18 Helgi Tomasson: Estimation of Correlations in Financial Markets when Trading is Infrequent

W00:19 Helgi Tomasson: Computations of Bayesian Estimators in ARMA Models

W00:20 Helgi Tomasson: Monitoring the trading intensity of a stock market under infrequent trading