Analyzing the Time between Trades with a Gamma Compounded Hazard Model. An Application to LIFFE Bund Future Transactions

Nikolaus Hautsch

Center of Finance and Econometrics, University of Konstanz, Germany

e-mail: nikolaus.hautsch@uni-konstanz.de

Working Paper
January 1999

Abstract

This paper investigates the time between transactions on financial markets. It is assumed that the interval between transactions is a random variable and the relationship between the probability to observe a transaction at each instant of time and the type of the previous trade is investigated. To estimate these effects, a semiparametric proportional hazard model is used which is based on approaches proposed by Han and Hausman (1990) and Meyer (1990). Considering grouped durations the log-likelihood is formed by using differences in the survivor function. Hence, the model corresponds to an ordered response approach whereby the baseline hazard is estimated simultaneously with the coefficients of the covariates and is calculated by the thresholds. Clustering of the durations is taken into account by including lagged durations. A test is proposed to check for serial correlation in the errors based on the concept of generalized residuals along the lines of the work of Gourieroux, Monfort and Trognon (1987). Unobservable heterogeneity is implemented parametrically by a gamma distributed random variable entering the hazard function. It is shown that the resulting compounded model follows a BurrII form. In an empirical analysis high frequency intraday transaction data from the London International Financial Futures and Options Exchange (LIFFE) is investigated.

1 Introduction

The development of high frequency data bases makes it possible to analyze the workings of financial markets based on the lowest aggregation level, i.e. on the basis of transaction data. The availability of a maximum amount of information contained in the data provides the opportunity to gain insights into market microstructures and allows for testing the validity of implications originating from theoretical microstructure approaches.

An important feature of transaction data is that it occurs in irregular time intervals, i.e. one trader is allowed to transact at any point in time. On one trading day (see Figure 2 [1]) there are periods in which transactions are generally infrequent, but there are also hours which have very high rates of activity. This may be due to some observable events such as the arrival of a news release or to unobservable events which are to be thought of as stochastic process. In market microstructure theory the timing of trades plays an important role in the learning mechanism of market participants who draw inferences from the trading process. In this context durations can be regarded as means to aggregate information on market activities [2], but also as indicators for the speed of the price adjustment process caused by news events [3]. Hence, studying the waiting times between events is essential for understanding the economics of financial markets.

It is assumed that the waiting time between two successive transactions is a random variable, implying the occurrence of probabilities to observe a trade at each instant of time. Theoretical approaches, like Admati and Pfleiderer (1988, 1989) and Easley and O'Hara (1992), suggest relationships between the probability of the existence of information and the characteristics associated with each trade, i.e. the price, the bid-ask spread, the volume and the duration between transactions. Hence, the probability of a trade occurring at each instant of time varies according to the type of the previous transaction.

In this study the time between trades is examined based on Bund Future trading from the London International Financial Future and Options Exchange (LIFFE). The Bund Future

---

[1] See Section 'Descriptive Statistics'.

[2] E.g. Coppejans and Domowitz (1998) use durations to compare the relative importance of information sets in limit order book trading.

[3] Engle and Lunde (1998) determine the speed of price adjustment by using a bivariate point process to jointly investigate transaction and quote arrivals.

is one of the most actively traded future contracts in Europe, resulting in a highly liquid market. Based on this data this paper investigates relationships between the economic characteristics of a transaction and the time until the next trade, i.e. the impact of the bid-ask spread, the volume and the price of a transaction on the conditional expectation of the waiting time until the next event are analyzed.

A specific feature of high frequency transaction data is the occurrence of clustering, which is found in different empirical investigations. Clustering structures are found in several economic variables, such as prices (Gwilym, Clare and Thomas, 1998 or Aitken et. al., 1996), volatilities (Bollerslev and Domowitz, 1993 or Andersen and Bollerslev, 1997) or durations (see Engle and Russell 1995, 1997, 1998). Admati and Pfleiderer (1988) propose a model which explains the concentration of trading in particular time periods within a trading day. In this approach the clustering of waiting times arises from strategic behaviour of informed and liquidity traders. The liquidity traders are partitioned in "discretionary" liquidity traders who have some choice over the timing of trades and "nondiscretionary" liquidity traders whose timing is random. It is shown that "discretionary" liquidity trading and also informed trading is typically concentrated. In the framework of Easley and O'Hara (1992) waiting times between trades carry information about the actual state of the market. They assume the existence of an uninformed market maker who uses market-activities to infer the existence of information. In this context concentrated-trading patterns arise from an information event which increases the number of informed traders.

To take duration clustering into account, Engle and Russell (1995, 1998) propose an Autoregressive Conditional Duration (ACD) model for intertemporally correlated event arrival times. The ACD model is based on a parametric specification for the conditional mean of the duration which captures the past trading history. Dividing each duration by their conditional mean yields "standardized" durations which are to be assumed as i.i.d. Thus, the ACD model provides a baseline hazard based on standardized durations [4].

This study uses a semiparametric proportional hazard model which also allows for the consideration of autoregressive structures but provides a nonparametric baseline hazard rate

---

[4]Lunde (1996) proposes a model which provides an alternative to the ACD framework. He captures duration clustering by allowing an unobserved stochastic process to drive the expected duration and the hazard function.

based on "non-standardized" durations. In this framework the error terms are to be assumed as conditional i.i.d., given the covariates. I.e. the explanatory variables have to capture all serial dependence leading to conditionally independent error terms. Based on a specification with conditionally i.i.d. disturbances a nonparametric baseline hazard for the durations is obtained without requiring to specify standardized waiting times. The approach is based on concepts by Han and Hausman (1990) and Meyer (1990). The main idea is to consider grouped durations which can be seen as a necessary device to sort information in the data and making it possible to treat the model in terms of an ordered response specification. [5] Using grouped durations the sample likelihood depends on the baseline hazard only through a finite number of discrete points and leads to a reduction in dimensionality (see An, 1998). An advantage of this model is that it obtains a nonparametric baseline hazard rate which is estimated simultaneously with the parameters of the covariates and is calculated by the thresholds.

A major reason to use categorized durations is the high number of ties. At the LIFFE the minimum time between events which can be recorded is one second. The occurrence of a high number of low durations generates the existence of ties, i.e. more than 50% of the reported durations between events have the discrete values 1,2,...,10 seconds [6]. The presence of ties makes it possible to treat each value as one category implying no loss of information and allowing for a close aproximation of the baseline hazard rate.

To take data clustering into account lagged durations are included. Furthermore, an underlying autoregressive process may not only cause correlations in the durations but also in the errors, leading to invalid inference procedures. Thus, it is necessary to check for serial correlation in the error terms. In this context a test is proposed based on the concept of generalized residuals (Gourieroux, Monfort and Trognon, 1987) which makes it possible to find a specification that captures the autoregressive structures and leads to conditionally noncorrelated error terms, given the explanatory variables. In such a specification all serial dependence has to be captured by the covariates leading to the assumption of conditionally

---

[5]The model is based on the general relationship between conventional models for ordered response specifications and related approaches for grouped durations (Sueyoshi, 1995), whereby the likelihood of a particular observation on a grouped duration is the probability of observing a series of binary outcomes.

[6]See Figure 1.

i.i.d. error terms.

A further advantage of the model from Han and Hausman is the fact that it is easy to control for unobservable heterogeneity. Lancaster (1979) and Heckman and Singer (1984) show that ignoring unobserved heterogeneity can lead to biased estimates of the hazard function. In econometric duration literature two directions exist concerning the implementation of unobservable heterogeneity. E.g. Heckman and Singer (1984), Honore (1990) and Bearse, Canals and Rilstone (1994) use nonparametric specifications for heterogeneity effects, but also need parametric forms for the baseline hazard rate. Cox (1972), Kiefer (1988), Han and Hausman (1990) and Meyer (1990) allow non-parametric baseline hazards, but specify unobservable heterogeneity by parametric distributions. [7]

In this study heterogeneity effects are specified parametrically by a random variable which enters the hazard function multiplicatively. An important advantage of this approach is that the survivor function of the compounded model can be calculated in closed form and that it does not require numerical integration. By specifying unobservable heterogeneity using a gamma distribution, a gamma compounded hazard model is obtained. It will be shown that the survivor function of the resulting mixed proportional hazard model has a BurrII form, a generalization of the logistic distribution. Thus, an ordered response approach is obtained which parametrically nests the ordered logit model.

The outline of the paper is as follows: Section 2 describes the econometric approach. Section 3 discusses testable hypothesis originated by market microstructure models. Section 4 gives empirical findings based on the analysis of LIFFE data. Section 5 presents conclusions.

2   An econometric approach to analyzing durations

2.1   A gamma compounded hazard model for grouped durations

This section provides an econometric approach for duration analysis which is based on arbitrarily grouped duration data. It is assumed that the time between the ith and the $(i+1)$th transaction is a random variable T_i, $i = 1, \ldots, n$, which follows an unknown baseline distribution. The model from Han and Hausman (1990) is based on the proportional hazard

---

[7]An (1998) proposes a mixed proportional hazard model which allows both the baseline hazard and the distribution of the unobserved heterogeneity to be flexible.

specification of Cox (1972) which assumes a nonparametric baseline hazard λ_0:

$$\lambda(t_i|X_i) = \lambda_0(t_i)exp(-X_i'\beta), \quad i = 1,\ldots,n, \tag{1}$$

whereby t_i denotes the time between two succeding trades, X_i a $R \times 1$-vector of covariates and β a $R \times 1$-vector of coefficients.

In duration literature different approaches exist to estimate this specification. Cox (1975) treats the baseline function as a nuisance function and estimates the parameter vector by proposing a partial likelihood which does not depend of the baseline hazard. Kalbfleisch and Prentice (1980) extend this approach by considering a baseline function which can be estimated by discrete baseline hazard parameters. Moffitt (1985) estimates the parameter vector and the baseline hazard simultaneously, but his specification does not guarantee that the probabilities are bounded between 0 and 1. Meyer (1990) and Han and Hausman (1990) form the likelihood by using differences in the survivor function which makes it possible to estimate the baseline hazard jointly with the regression parameters whereby the probabilities of surviving each period are constrained to lie in the unit interval. An advantage of this framework is that it is relatively easy to implement unobservable heterogeneity. Such effects are included by specifying a random variable ω which acts multiplicatively with the hazard function, whereby it is assumed that ω is independent of X_i. The compounder can be regarded as an additional degree of freedom which captures a part of the variations of the endogenous variable which cannot be explained by the covariates [8]. Thus, the better the model is specified, the lower the variance of this variable should be. Like Lancaster (1979) and Han and Hausman (1990), it is assumed that this random variable has a gamma distribution with mean one and variance θ^{-1}. Given the covariates and unobservable heterogeneity effects, the conditional hazard is equal to

$$\lambda(t_i|X_i,\omega) = \lambda_0(t_i)exp(-X_i'\beta + \eta), \quad i = 1,\ldots,n, \tag{2}$$

where $\omega = exp(\eta)$. The survivor function of the compounded hazard model is obtained by calculating the conditional survivor function and integrating out ω:

$$S(t_i|X_i) = \int_0^{\infty} S(t_i|X_i,\omega)g(\omega)d\omega$$

---

[8]It has to taken into account that the compounder is time invariably and, thus, cannot capture the entire ommitted variations.

$$= \int_0^\infty exp\left[-exp(-X_i'\beta + \eta)\int_0^{t_i} \lambda_0(s)ds\right], (\theta)^{-1}\omega^{\theta-1}\theta^\theta exp(-\theta\omega)d\omega$$

$$= \left[1 + \theta^{-1}exp(-X_i'\beta)\Lambda_0(t_i)\right]^{-\theta}, \tag{3}$$

where $\Lambda_0(t_i)$ denotes the integrated baseline hazard rate and $g(\omega)$ the density of the gamma distribution.

Let $t_i^* \equiv ln[\Lambda_0(t_i)]$, then the hazard model can be written in the form

$$t_i^* = X_i'\beta + ln(\theta) - e_i, \quad i = 1, \ldots, n, \tag{4}$$

with

$$F(e_i) = \frac{1}{[1 + exp(-e_i)]^\theta}.$$

Hence, the specification can be written in the form of a linear regression model for the transformed duration t_i^* with an error term following a BurrII(θ) distribution. The BurrII(θ) distribution can be seen as a generalization of the logistic distribution, thus, the logistic distribution corresponds to the BurrII(1) distribution.

Using the fact that

$$\lim_{x \to \infty} (1 + \frac{x}{n})^n = exp(x),$$

it is shown that

$$\lim_{\theta \to \infty} \left[1 + \theta^{-1}exp(-X_i'\beta)\Lambda_0(t_i)\right]^{-\theta} = exp(-\Lambda_0(t_i)exp(-X_i'\beta)),$$

i.e. for $\theta^{-1} = Var(\omega) \to 0$ the BurrII distribution goes to an extreme value distribution. Hence, if no unobservable heterogeneity effects exist the proportional hazard model can be written as $t_i^* = X_i'\beta - e_i$, whereby e_i follows an extreme value distribution [9]

$$f(e_i) = exp(-e_i - exp(-e_i)).$$

For the error terms of equation (4) it is assumed conditional independence, given the covariates X_i. To take into account duration clustering lagged durations have to be included,

---

[9]See Johnson, Kotz and Balakrishnan (1994).

i.e. all serial dependence has to be captured by the X_i's leading to conditional i.i.d. error terms [10].

Considering grouped durations allows for the formation of the log likelihood by calculating the probabilities of one trade occurring in particular categories. In this framework the specification corresponds to an ordered response model and obtains a nonparametric baseline hazard which can be calculated by the estimated thresholds of the chosen categories. Define

$$\delta_k \equiv ln[\Lambda_0(t_k)], \tag{5}$$

where t_k, $k = 1, \ldots, K$, denote the bounds of the chosen categories. The conditional probability of failure, i. e. the occurence of a trade, in category k, conditioned on X_i is

$$Pr\left[t_{k-1} < t_i \leq t_k | X_i\right] = Pr\left[\delta_{k-1} < t_i^* \leq \delta_k | X_i\right] = \int_{X_i'\beta+ln(\theta)-\delta_{k-1}}^{X_i'\beta+ln(\theta)-\delta_k} f(s)ds, \tag{6}$$

where f(s) denotes the density function of the BurrII(θ) distribution, given by

$$f(s) = \frac{\theta exp(-s)}{(1 + exp(-s))^{\theta+1}}.$$

Hence, equation (4) can be seen as the latent model of the resulting ordered response approach. The log likelihood function takes the form

$$l(\beta, \delta) = \sum_{i=1}^{n} \sum_{k=K}^{k} y_{it} \, ln \int_{X_i'\beta+ln(\theta)-\delta_{k-1}}^{X_i'\beta+ln(\theta)-\delta_k} f(s)ds, \tag{7}$$

where the indicator variable y_{it} is defined by

$$y_{it} = \begin{cases} 1 & , & if \quad t_i \in [t_{k-1}, t_k] \\ 0 & , & else \end{cases} .$$

A nice feature of this approach is that the nonparametric baseline survivor function is obtained directly by the estimated thresholds. It can be calculated at the k discrete points by

$$S_0(t_k) = \frac{1}{[1 + exp(\delta_k - ln(\theta))]^\theta}, \quad k = 1, \ldots, K. \tag{8}$$

---

[10]In the ACD framework Engle and Russell (1998) define standardized durations by dividing each duration by their conditional expectation, given the past trading history. These standardized durations are to be assumed as i.i.d.

The baseline hazard rate can be approximated by

$$\lambda_0(t_k) \approx \frac{S_0(t_k) - S_0(t_{k+1})}{S_0(t_k)} \cdot \frac{1}{\Delta t}, \quad k = 1, \ldots, K. \tag{9}$$

2.2 Testing for serial correlation

If autocorrelations between the waiting times exist it is necessary to include lagged durations in the regression. Thus, serial dependence has to be captured by the X_i's leading to conditional independence of the error terms given X_i. To check for serial dependence of the errors a test for serial correlation is proposed. This test opens up the possibility of finding a suitable specification that captures autoregressive structures and leads to noncorrelated error terms that ensures valid inference processes. The test for serial dependence is based on the concept of generalized residuals proposed by Gourieroux, Monfort and Trognon (1987). The authors provide a direct relationship between the score of the observable and the latent model. By using generalized residuals the observable score can be calculated by the conditional expectation of the latent score. As shown in Section 2.1 the latent model can be written as

$$t_i^* = X_i'\beta + ln(\theta) - u_i, \tag{10}$$

with $u_i = \rho_j u_{i-j} + e_i$, where e_i is i.i.d. following a BurrII form and j denotes the number of the tested lag. The null hypothesis is $H_0 : \rho_j = 0$. Along the lines of the work of Gourieroux, Monfort and Trognon the observable score is equal to the conditional expectation of the latent score, given the observable endogenous variable, i. e.

$$\frac{\partial l(t; \psi)}{\partial \psi} = E\left[\frac{\partial l^*(t^*, \psi)}{\partial \psi}\bigg| t\right],$$

where ψ denotes the parameter vector and $l^*(.)$ the latent score. The log likelihood function of the latent model is

$$\begin{aligned}
l^*(t^*, \beta, \theta, \rho_j) &= \sum_{i=j+1}^{n} ln f(u_i - \rho_j u_{i-j}) \\
&= \sum_{i=j+1}^{n} \left[ln(\theta) + \rho_j u_{i-j} - u_i - (\theta + 1)ln\left[1 + exp(\rho_j u_{i-j} - u_i)\right]\right], \quad (11)
\end{aligned}$$

where $f(.)$ denotes the BurrII density function. Under the null, ξ is given by

$$\xi = \frac{\partial l(t, \beta, \theta, 0)}{\partial \rho} = \sum_{i=j+1}^{n} E_0[u_{i-j}|t_i]\left[1 - (\theta + 1)E_0\left[\frac{exp(-u_i)}{1 + exp(-u_i)}\bigg| t_i\right]\right]. \tag{12}$$

9

The conditional expectations of the disturbances under the null are called generalized residuals and are denoted by $\tilde{u}_i = E_0[u_i|t_i]$ and $\tilde{w}_i = E_0\left[\frac{exp(-u_i)}{1+exp(-u_i)}\middle| t_i\right]$. Based on these generalized residuals (see Appendix), the score statistic is given by

$$\hat{\xi} = \sum_{i=j+1}^{n} \hat{\tilde{u}}_{i-j}\left[1 - (\hat{\theta}+1)\hat{\tilde{w}}_i\right]. \tag{13}$$

In the Appendix it is shown that under the null the expectation of $\hat{\tilde{w}}_i$ is given by

$$E[\hat{\tilde{w}}_i] = \frac{1}{\theta+1}.$$

Thus, under the H_0 the expectation of the estimated score is $E[\hat{\xi}] = 0$. Using the asymptotic normality of the score, i.e.

$$\frac{1}{\sqrt{n}}\hat{\xi} \xrightarrow{d} N\left(0, plim\frac{1}{n}\sum_{i=j+1}^{n} \tilde{u}_{i-j}^2\left[1 - (\theta+1)\tilde{w}_i\right]^2\right),$$

the chi-square statistic is obtained by

$$S = \frac{\left[\sum_{i=j+1}^{n} \hat{\tilde{u}}_{i-j}\left[1 - (\hat{\theta}+1)\hat{\tilde{w}}_i\right]\right]^2}{\sum_{i=j+1}^{n} \hat{\tilde{u}}_{i-j}^2\left[1 - (\hat{\theta}+1)\hat{\tilde{w}}_i\right]^2} \sim \chi_1^2. \tag{14}$$

3 Market Microstructure

The main focus of market microstructure theory is the investigation of how market participants draw inferences from the trading process. In microstructure literature several approaches exist which model relationships between characteristics associated with transactions, such as bid ask spread, volume, price and waiting time between events. See O'Hara (1995) for a survey. In this literature traders are often partitioned into informed market participants, who only trade after obtaining private information, and liquidity traders (noise traders) whose trading is based on exogenous reasons, like liquidity or portfolio-balancing aspects. An important assumption of many microstructure approaches is the existence of an uniformed market maker learning from order flow characteristics.

In the trading models of Glosten and Milgrom (1985) and Kyle (1985) insiders and liquidity traders submit orders to a specialist who sets prices such that his expected profits are

zero given the information set. The disadvantage of these models is that the waiting times between events have no informational content.

Admati and Pfleiderer (1988) and Easley and O'Hara (1992) assume that trades take place sequentially with only one trader being allowed to transact at any point in time. In the approach of Easley and O'Hara (1992) the specialist updates his conditional expectations in response to order flow characteristics. His expectations are based on a priori probabilities for the arrival of orders executed by insiders and the composition of the traders. Thus, the specialist behaves according to a Bayesian learning process and uses no-trade-intervals to infer the existence of new information. Observing a buy, a sell or the absence of trading he determines a posteriori probabilities for the presence of news.

In this framework a positive correlation between the bid ask spread and the probability for the existence of information is proposed. This relationship arises because the higher the probability for the existence of information, the higher the probability for the presence of insiders and, thus, the higher the probability that the market maker executes transactions with informed market participants. Because these insiders have an informational advantage, the market maker transacts at a loss. He compensates his losses by setting a wider spread between bid and ask. Because a higher probability for the presence of news implies lower expected waiting times, a negative correlation between bid ask spread and the expected duration until the next trade is proposed. This implication is formulated in the testable hypothesis H1:

Hypothesis H1: *The bid ask spread carries information and has a negative impact on the time until the next transaction.*

Furthermore, Easley and O'Hara (1987) assume that informed traders prefer to trade larger amounts at any given price, thus, the strategy of the market maker must depend on trade sizes. Blume, Easley and O'Hara (1994) investigate the informational role of volume. In this framework volume captures the important information contained in the quality of traders' information signals. Easley and O'Hara (1992) consider the market makers' belief, given an observed trading history, and propose informational content of aggregate volumes. These implications are summarized in H2:

11

Hypothesis H2(a): *Aggregated volume has a negative impact on the expected durations.*

In this context it has to taken into account that aggregate volumes are influenced by past durations, i.e. the lower the time between trades within a certain time interval, the more events occur within this time interval and the higher the aggregated volume is. To investigate pure volume effects which are not influenced by the time between the events, average volumes are used:

Hypothesis H2(b): *Average volume has a negative impact on the expected durations.*

Admati and Pfleiderer (1988) assume two types of uninformed traders, "discretionary" liquidity traders, who have some choice over the time at which they transact, and "nondiscretionary" liquidity traders whose orders are assumed to arrive in a random fashion. In this framework it is optimal for liquidity traders and also for insiders to trade together leading to concentrations of trading in particular time periods and suggests clustering of transactions. In the Easley and O'Hara (1992) framework the interval between two arrival times carries information about the actual state of the market. Because the probability for the occurrence of further transactions increases with the probability for the existence of information, the durations are positive autocorrelated. Engle and Russell find evidence for this hypothesis by analyzing IBM transactions (Engle and Russell, 1995, 1998) and foreign exchange data (Engle and Russell, 1997). In this study this feature will also be investigated for Bund Future transactions and is formulated in H3:

Hypothesis H3 *The durations are positive autocorrelated.*

In this context the implementation of lagged endogenous variables is not only useful to check this hypothesis but is also necessary to ensure the conditional independence of the error terms given the explanatory variables (see Section 2.1).

4   Empirical findings

4.1   The data

The hypotheses derived in section 3 will be tested by Bund-Future transaction data from the LIFFE. The Bund-Future is a notional German government bond of DEM 250,000 face

value which matures in 8.5 to 10.5 years at contract expiration. The months of March, June, September, and December are the four contract maturities per year.

The study uses intradaily data of the 22 trading days in August 1995 corresponding to 38977 transactions. The dataset obtains time stamped prices, bids, asks and volumes associated with the transactions.

The dependent variable is obtained from the time between two successive trades. Figure 1 [11] shows the distribution of the durations indicating a high number of ties especially for low values. The existence of these ties makes it possible to consider each low duration as one category which enables to calculate a close approximation to the baseline hazard. Thus, the regressions are based on the categories [0;1],(1;2],...,(49;50],(50,∞).

Explanatory variables are obtained from the characteristics of each transaction. To test Hypothesis H1 a dummy for certain tick sizes of the bid ask spread is used. Because almost 98% of the reported spreads are equal to one or two (see Figure 4), a dummy is generated which indicates tick sizes larger than at least two ticks.

For testing the volume hypotheses H2(a) and H2(b) aggregated and average volumes are included, while for checking Hypothesis H3 lags of the durations are used.

Furthermore, it is possible to generate further explanatory variables. To investigate the informational content of the volume associated with one single trade a dummy is used which indicates above average volumes of the past trade. Figure 3 shows the distribution of transaction volumes with an obvious peak at 41 units [12]. This feature arises from the fact that at the LIFFE transactions are registered by price reporters who record the volume of each trade by approximate values, for example, the price reporters indicate middle volumes with the value 41. Therefore, it seems reasonable to generate a dummy variable which indicates volumes larger than at least 41 units as an explanatory variable.

The impact of prices can be examined by using midquotes, which are defined by

$$mq_i = \frac{bid_i + ask_i}{2}.$$

This definition eliminates the problem of "bid-ask bounces" (see Roll, 1984).

Figure 5 shows the distribution of midquote changes between successive trades. To inves-

---

[11]See section 'Descriptive Statistics'.

[12]One unit corresponds to 250,000 DEM.

tigate the impact of midquote changes on the expected duration a dummy is used, which indicates midquote changes larger than at least 1.5 ticks. Additionally, a dummy is generated, which registers whether midquotes have not changed in the last 5 trades. This variable can be seen as a proxy for volatility. On the basis of this dummy it is possible to investigate the impact of past midquote movements on the probability for the occurence of further transactions.

On the basis of this data it is possible to identify buys and sells by assuming that a transaction is a buy (sell) if it occurs at a price equal or higher (lower) than the previously reported ask (bid) quote. Along the lines of microstructure theory, frequent changes between sells and buys (or vice versa) indicate trading by noise traders, who transact only for private reasons. This leads to the testable hypothesis that the hazard rate will increase if in the last few transactions only buys (sells) followed one another. This implication will be tested by a dummy which indicates changes between buys and sells in the last 5 transactions.

An often observed feature of transaction data is the existence of intraday and interday seasonalities (see Goodhart and O'Hara, 1997, Guillaume et. al., 1997). Figure 6 shows the average hourly durations of the LIFFE Bund Future market which is open from 8:30 a.m. to 5:15 p.m. A U-shaped pattern of daily market activities is evident with transactions occurring, on average, every 11 seconds at the opening, every 31 seconds at lunch time and every 12 seconds at the closing. Engle and Russell (1997, 1998) generate seasonal adjusted durations by dividing each waiting time by its seasonal component which is obtained by spline regressions. In this study seasonal patterns are taken into consideration by dummy variables which partition one trading day in five time-intervals. Furthermore, intradaily seasonalities within a trading week (see Figure 7) are investigated by dummies indicating the particular days of the week.

To avoid endogeneity all explanatory variables enter the equation in lagged form i.e. the impact of the characteristics of the previous trade on the expected waiting time until the next transaction is investigated.

14

4.2 Results

Regression 1 (Table 1) shows the results of a pooled regression of LIFFE Bund Future transactions for August 1995, corresponding to 22 single trading days. To take the clustering of the waiting times into account, lagged log-durations are included. Based on this specification the null hypothesis of no serial correlation has not to be rejected at the 5% level.

The negative coefficient of the spread-dummy illustrates a positive correlation between the size of the bid ask spread and the probability for the existence of information[13]. Hence, the probability for the occurence of a further transaction increases with the size of the spread, confirming H1.

To check Hypothesis H2(a) time aggregates of volumes are included. The strikingly high significance of the aggregated volumes within the first 5 minutes implies that this variable contains the main part of informational content. The sign of the coefficient of this variable indicates a negative impact on the expected duration, confirming H2(a). The parameters of the other volume variables provide no clear results. The coefficients of these covariates are very close to zero which impairs the significance of these variables substantielly. Hence, based on these explanatory variables no clear empirical evidence is found.

In this context it also has to taken into account that aggregated volumes are influenced by past durations (see Section 3) causing correlations between the volume variables [14]. To eliminate these correlations and to investigate Hypothesis H2(b), average volumes are used (see Regression 2, Table 2). In this context a significant negative coefficient of the average volumes within a time horizon of up to 5 minutes is obtained, while average volumes dating back longer have a positive impact on the expected waiting time. Thus, for the impact of average volumes of LIFFE Bund Future transactions on the expected time until the next trade no clear empirical evidence is found.

The inclusion of average volumes makes it possible to use higher lags of log-durations without causing correlations between these covariates and the volume variables. The lagged durations enter the regression in aggregated form. Regression 2 shows that waiting times

---

[13]Note that a negative coefficient of the covariate increases the hazard rate and, thus, reduces the time until the next transaction.

[14]To avoid correlations between the transformed lagged durations and the aggregated volumes only four lags are used whereby the volume variables are generated by trades that occurred before the fourth lag.

until lag 60 have a positive significant impact on the expected duration. Thus, positive autocorrelations between the durations are evidently, whereby H3 cannot be rejected.

Hence, based on this results it is found empirical evidence for the market microstructure hypotheses H1, H2(a) and H3, whereby for average volumes no clear result is obtained.

Furthermore, a significant negative coefficient for the dummy indicating above-average volumes of the previous trade is found. Thus, informational content of volumes associated whith single trades is evidently.

For the dummies indicating midquote movements and changes of buys to sells (or vice versa) significant negative values are presented. Thus, midquote changes indicate the existence of information and have a negative impact on the expected waiting time.

The significant value of the dummy registering buy-sell changes implies that the probability for the occurence of information increases if only buys (sells) follow one another. Frequent changes between buys and sells indicate that no important information exists and that market activities are dominated by noise trading.

For the intraday seasonality dummies significant coefficients are obtained. In particular the intraday variables indicate a strong impact of seasonal patterns on the time between trades within one trading day. Waiting times between trades which occur between 12:00 and 14:00 are significantly higher, while lower waiting times at the opening and closing are evident, confirming the hypothesis of U-shaped patterns of daily market activities (see Admati and Pfleiderer, 1988, Bollerslev and Domowitz, 1993, or Madhavan, Richardson and Roomans, 1997). For the coefficients of the dummies indicating trading at Wednesday and Thursday significantly [15] coefficients are obtained. Hence, weak evidence is found for inverse U-shaped seasonalities of the weekly trading structure with higher durations at Monday and Friday and lower durations at he middle of the week.

For the variance of the heterogeneity variable small values are obtained, implying the existence of only weak unobservable effects. Based on Regression 2 a value of $Var[\omega] = \theta^{-1} = 0.0671$ is obtained. Thus, the estimation is based on a BurrII(14.903) distribution which is skewed to the right and is very close to the extreme value specification (see Figure 8). Thus, the extreme value specification can be seen as a very close approximation to the

---

[15]At the 10% level.

BurrII(14.903) distribution. Furthermore, the heterogeneity parameter differs significantly from $\theta = 1$, corresponding to an ordered logit approach, implying that the ordered logit model is misspecified.

Running regressions by omitting covariates show that a part of the omission is captured by the variance of the heterogeneity parameter. The less of the variation of the dependent variable can be explained by the covariates, the higher the variance of the heterogeneity variable. By including only the first lag of the transformed duration as one explanatory variable, a value of $Var[\omega] = \theta^{-1} = 0.3809$ is obtained (see Regression 4, Table 3), which is significantly higher than in Regression 1 or 2. Thus, the heterogeneity parameter behaves like a residual variance. Addison and Portugal (1998) find a similar result by using proportional hazard models with a parametrically specified baseline hazard which follows a Burr distribution. Of course it has to be taken into account that the heterogeneity variable cannot entirely compensate for the omission of covariates because the parameter is time invariably.

To check the robustness of the estimated coefficients of the covariates different categorizations are used. Table 2 (Regression 3) shows the results of a regression based on only four categories [0;5],(5;10],(10;20],[20,∞). It is recognizable that the estimation of the coefficients is not influenced by the chosen categorization.

Figure 9 and 10 [16] show the baseline survivor function and hazard rate based on Regression 2. To demonstrate the effects caused by misspecifications of the heterogeneity variable Figure 9 and 10 show also the estimated baseline functions based on an ordered logit model. It is recognizable that the misspecification of the ordered logit approach is reflected in the baseline survivor function and hazard rate. Comparing the densities of the BurrII and the ordered logit specification (see Figure 8) it is recognizable that the ordered logit model has more (less) probability mass in the lower (upper) tail, leading to higher (lower) values of the baseline survivor function for higher (lower) durations. This misspecification of the ordered logit model is also reflected in the baseline hazard rate leading to a significantly different shape of the nonparametric function. This result indicates the importance of including a flexible heterogeneity parameter.

---

[16]For all estimated values of the estimated baseline survivor and hazard functions significance at the 1% level is obtained.

The baseline hazard based on the gamma compounded model has an increasing non-monotonic shape until it reaches 7 seconds, then it is followed by a slowly decreasing hazard. Hence, it is recognizable that the traders need about 7 seconds to react to the arrival of the last transaction. Because of this delay the probability for the occurrence of further trades increases within this time. After this period the impact of information on the trading behaviour of market participants seems to decrease. Bisière and Kamionka (1998) obtain very similar shapes of baseline hazards by analyzing high frequency data from the Paris Bourse with a competing risk model.

5   Conclusions and outlook

This paper deals with the analysis of the time between trades in financial markets. It uses an approach which is based on concepts by Han and Hausman (1990) and Meyer (1990) and provides a nonparametric baseline hazard. The main idea is to regard the waiting time between events as categorized durations making it possible to form the likelihood by using differences in the survivor function. In this framework a nonparametric baseline hazard is obtained which can be calculated by the thresholds and, thus, can be estimated simultaneously with the coefficients of the covariates.

To take into account clustering of transaction data and to ensure the conditionally independence of the error terms, lagged durations have to be included. A test for autocorrelation in the errors based on the concept of generalized residuals (Gourieroux, Monfort and Trognon, 1987) makes it possible to find a suitable specification which takes the autoregressive structure of the data into account and also leads to noncorrelated errors and, thus, valid inferences. Hence, in such a specification all serial dependence is captured by the covariates leading to conditionally independent errors, given the explanatory variables.

The implementation of a gamma distributed random variable which specifies unobservable heterogeneity and enters the hazard rate multiplicatively leads to a gamma compounded hazard model whereby the survivor function can be calculated in closed form. It is shown that the resulting ordered response approach is based on a latent model which has a BurrII form.

This paper investigates relationships between the economic characteristics of a transaction

and the time until the next trade. Three hypotheses originating from market microstructure models are tested on the basis of Bund Future transactions from the LIFFE:

1. The bid ask spread has a negative impact on the time until the next transaction.

2. Aggregated and average volumes of past transactions are negative correlated with the waiting time until the next event.

3. The durations are positive autocorrelated.

Based on the empirical analysis evidence is found for the market microstructure hypotheses 1 and 3, whereby for volumes no clear results are obtained. Especially for the informational content of average volumes no significant evidence can be shown.

Furthermore it is shown that midquote movements indicate the existence of information and have a significant impact on the time between trades. Using a dummy registering changes between buy and sell transaction shows that frequent changes indicate that no essential information exists and that market activities are dominated by noise traders.

The heterogeneity parameter behaves like a residual variance, i.e. the less of the variations of the endogenous variable can be explained by the covariates, the higher the heterogeneity variance.

Estimating the resulting baseline hazard provides a shape which increases until 7 seconds and decreases slowly for higher values. This result shows that traders react whith delay to the arrival of the last transaction.

Furthermore, it is recognizable that misspecifications of the heterogeneity variable lead to a biased baseline survivor and hazard function showing the importance of including a flexible heterogeneity parameter.

Future research will be concerned with the power of this approach to other financial markets, especially the screen based automated trading system of the Deutsche Terminbörse (DTB). At the DTB the Bund Future is traded in an almost identical fashion, therefore transaction data based on these two exchanges can be used to compare both trading systems.

Furthermore, it would be interesting to analyze price durations, i.e. waiting times between certain price changes [17]. Because it is easy to take into account censoring of the data it is possible to investigate price changes which occur over night. In this context calculating the

---

[17]See e.g. Engle and Russell (1997, 1998) or Darolles, Gourieroux and Le Fol (1998).

19

probability for price movements in certain time intervals can be seen as an instrument for risk management strategies.

6 Appendix

Calculation of the generalized residuals

If $t_i \in [t_{k-1}, t_k]$, the generalized residuals are calculated by

$$\tilde{u}_i = E_0[u_i|t_i] = E[\epsilon_i|t_i] = \frac{\int\limits_{a_i}^{b_i} s \frac{\theta exp(-s)}{[1+exp(-s)]^{\theta+1}} ds}{F_{E_i}(b_i) - F_{E_i}(a_i)}$$

and

$$\tilde{w}_i = E_0\left[\left.\frac{exp(-u_i)}{1+exp(-u_i)}\right| t_i\right] = E\left[\left.\frac{exp(-\epsilon_i)}{1+exp(-\epsilon_i)}\right| t_i\right] = \frac{\int\limits_{g(a_i)}^{g(b_i)} s f_{g(E)}(s) ds}{F_{g(E_i)}(g(b_i)) - F_{g(E_i)}(g(a_i))},$$

where

$$
\begin{aligned}
a_i &= X_i'\beta + ln(\theta) - \delta_{t-1}, \\
b_i &= X_i'\beta + ln(\theta) - \delta_t
\end{aligned}
$$

and $F_{E_i}(.)$ denotes the BurrII(θ) distribution function

$$F_{E_i}(e_i) = \frac{1}{[1+exp(-e_i)]^\theta}.$$

$f_{g(E_i)}(s)$ and $F_{g(E_i)}(s)$ denote the density respectively the distribution function of the transformed residuals

$$g(u_i) = \frac{exp(-u_i)}{1+exp(-u_i)}.$$

The density $f_{g(\epsilon)}(g(e))$ is obtained by transformation. Let $w = g(\epsilon)$ the transformation and $h(w) = y = g^{-1}(\epsilon)$ the inverse transformation, then

$$f_\epsilon(h(w)) = \theta y(1-y)^\theta).$$

Thus, the density of the transformed random variable is given by

$$f_{g(\epsilon)}(w) = \frac{f_\epsilon(h(w))}{|g'(h(w))|} = \theta(1-w)^{\theta-1} \mathbf{1}_{[0;1]}(w),$$

20

whereby the distribution function can be calculated by

$$F_{g(\epsilon)}(w) = 1 - (1 - w)^\theta \mathbf{1}_{[0;1]}(w).$$

The expectation of ω is obtained by $E[\omega] = \frac{1}{\theta+1}$.

7   References

ADDISON, J.T., AND P. PORTUGAL (1998): "Some Specification Issues in Unemployment Duration Analysis," *Labour Economics,* 5, 53-66.

ADMATI, A.R., AND P. PFLEIDERER (1988): "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies,* 1, 3-40.

ADMATI, A.R., AND P. PFLEIDERER (1989): "Divide and Conquer: A Theory of Intraday and Day-of-the-Week Mean Effects," *Review of Financial Studies,* 2, 189-223.

AITKEN, M., P. BROWN, C. BUCKLAND, H.Y. IZAN, AND T. WALTER (1996): "Price Clustering on the Australian Stock Exchange," *Pacific-Basin Finance Journal,* 4, 297-314.

AN, M.Y. (1998): "Consistent Estimation of Regression Coefficients in Mixed Proportional Hazard Models," *Preliminary, Presented at the Conference on Panel Data and Structural Labour Market Models, Sandbjerg, DK, June 1998.*

ANDERSEN, T.G., AND T. BOLLERSLEV (1997): "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance,* 4, 115-158.

BEARSE, P., J. CANALS, AND P. RILSTONE (1994): "Semiparametric Maximum Likelihood Estimation of Duration Models with Unobserved Heterogeneity," Working Paper, Department of Economics, York University.

BISIÈRE, C., AND T. KAMIONKA (1998): "Timing of Orders, Orders Aggressiveness and the Order Book in the Paris Bourse," Working Paper, Presented at the "Symposium on Microstructure and High Frequency Data" in Paris, 1998.

BLUME, L., D. EASLEY, AND M. O'HARA (1994): "Market Statistics and Technical Analysis: The Role of Volume," *Journal of Finance,* 49, 153-181.

BOLLERSLEV, T., AND I. DOMOWITZ (1993): "Trading Patterns and Prices in the Interbank Foreign Exchange Market," *Journal of Finance,* 48, 1421-1443.

BOX, G., AND D. COX (1964): "An Analysis of Transformations," *Journal of the Royal Statistical Society, Series B,* 26, 211-243.

COPPEJANS, M., AND I. DOMOWITZ (1998): "Stock and Flow Information as Inputs to Limit Order Book Trading Activity," Working Paper, Presented at the "Symposium on Microstructure and High Frequency Data" in Paris, 1998.

COX, D.R. (1972): "Regression Models and Life-Tables," *Journal of the Royal Statistical Society, Series B,* 34, 187-200.

COX, D.R. (1975): "Partial Likelihood," *Biometrika,* 62, 269-276.

DAROLLES, S., C. GOURIEROUX, AND G. LE FOL (1998): "Intraday Transaction Price Dynamics," Working Paper, Presented at the "Symposium on Microstructure and High Frequency Data" in Paris, 1998.

EASLEY, D., AND M. O'HARA (1987): "Price, Trade Size and Information in Securities Markets," *Journal of Financial Economics,* 19, 69-90.

EASLEY, D., AND M. O'HARA (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance,* 47, 577-607.

ENGLE, R.F., AND A. LUNDE (1998): "Trades and Quotes: A Bivariate Point Process," Working Paper, Presented at the "Symposium on Microstructure and High Frequency Data" in Paris, 1998.

ENGLE, R.F., AND J.R. RUSSELL (1995): "Forecasting Transaction Rates: The Autoregressive Conditional Duration," *UCSD Discussion Paper,*

ENGLE R.F., AND J.R. RUSSELL (1997): "Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model," *Journal of Empirical Finance,* 4, 187-212.

ENGLE, R.F., AND J.R. RUSSELL (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica,* 66, 1127-1162.

GUILLAUME, D.M., M.M. DACOROGNA, R.R. DAVE, U.A. MÜELLER, R. OLSEN, AND O.V. PICTET (1997): "From the Bird's Eye to the Microscope: A Survey of New Stylized Facts of the Intra-Daily Foreign Exchange Markets," *Finance and Stochastics,* 1, 95-129.

GLOSTEN, L.R., AND P. MILGROM (1985): "Bid, Ask and Transaction Prices in Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics,* 14, 71-100.

GOODHART, C.A.E., AND M. O'HARA (1997): "High Frequency Data in Financial Markets: Issues and applications," *Journal of Empirical Finance,* 4, 73-114.

GOURIEROUX, C., A. MONFORT, AND A. TROGNON (1987): "A General Approach to Serial Correlation," *Econometric Theory,* 1, 315-340.

GWILYM, O., A. CLARE, AND S. THOMAS (1998): "Extreme Price Clustering in the London Equity Index Futures and Options Markets," *Journal of Banking and Finance,* 22, 1193-1206.

HAN, A., AND A.J. HAUSMAN (1990): "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics,* 5, 1-28.

HECKMAN, J., AND B. SINGER (1984): "Econometric Duration Analyses," *Journal of Econometrics,* 24, 63-132.

HONORE, B.E. (1990): "Simple Estimation of a Duration Model with Unobserved Heterogeneity," *Econometrica,* 58, 453-473.

JOHNSON, N.L., S. KOTZ, AND N. BALAKRISHNAN (1994): *Continuous Univariate Distributions. Volume 1.* Wiley, New York.

KALBFLEISCH, J.D., AND R. PRENTICE (1980): *The Statistical Analsysis of Failure Time Data.* Wiley, New York.

KIEFER, N. M. (1988): "Economic Duration Data and Hazard Functions," *Journal of Eco-*

nomic Literature, 26, 646-679.

KYLE, A.S. (1985): "Continous Auctions and Insider Trading," *Econometrica,* 53, 1315-1335.

LANCASTER, T. (1979): "Econometric Methods for the Duration of Unemployment," *Econometrica,* 47, 939-956.

LUNDE, A. (1996): "A Conjugate Gamma Model for Duration in Transactions Data," Working Paper, Centre for Nonlinear Modelling in Economics, University of Aarhus.

MADHAVAN, A., M. RICHARDSON, AND M. ROOMANS (1997): "Why do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks," *Review of Financial Studies,*10, 1035-1064.

MEYER, B. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica,* 58, 757-782.

MOFFIT, R. (1990): "Unemployment Insurance and the Distribution of Unemployment Spells," *Journal of Economentrics,* 28, 85-101.

O'HARA, M. (1995): *Market Microstructure Theory.* Blackwell, Oxford.

ROLL, R. (1984): "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *Journal of Finance,* 39, 1127-1140.

SUEYOSHI, G.T. (1995): "A Class of Binary Response Models for Grouped Duration Data," *Journal of Applied Econometrics,* 10, 411-431.
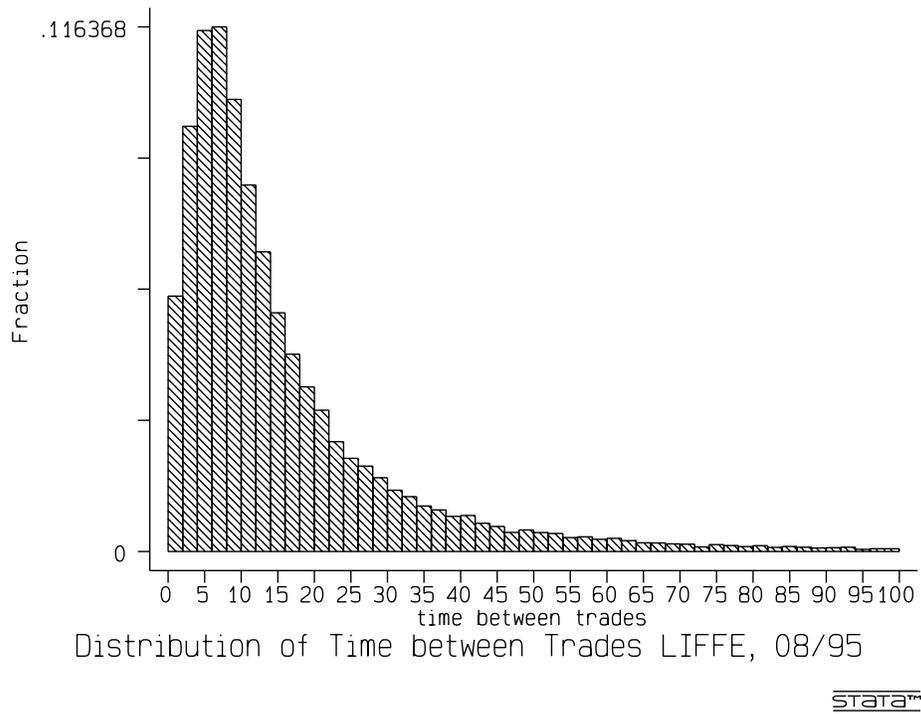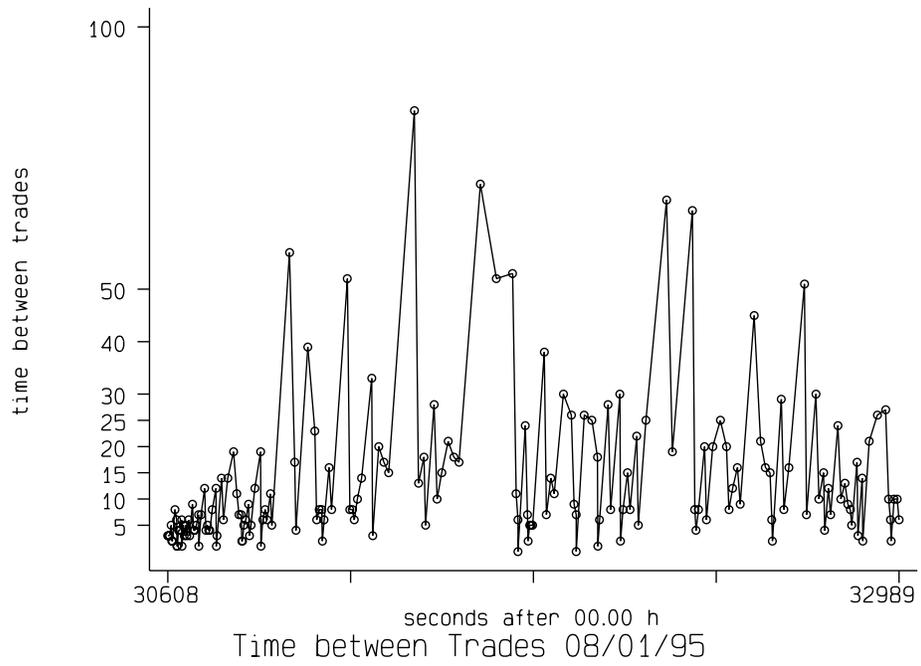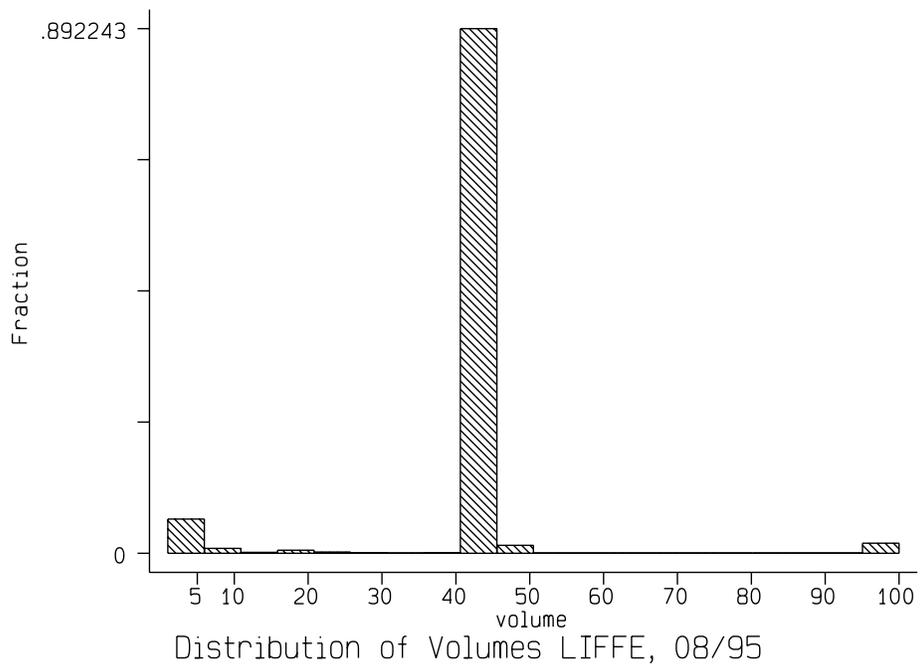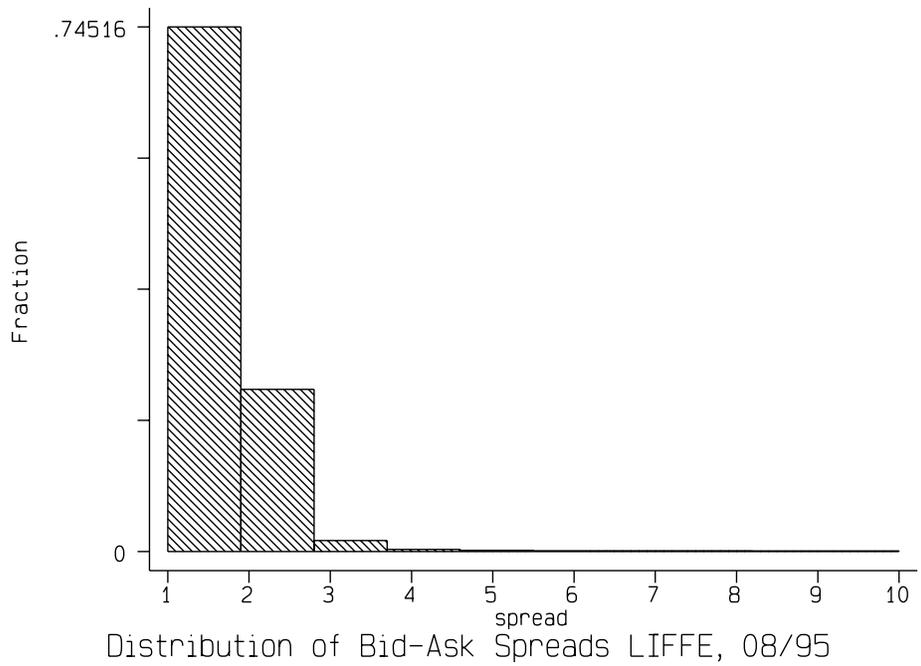
8 Descriptive Statistics

Figure 1

Figure 2

Figure 3

26

Distribution of Bid-Ask Spreads LIFFE, 08/95

STATA™

Figure 4

Distribution of Midquote Changes LIFFE, 08/95
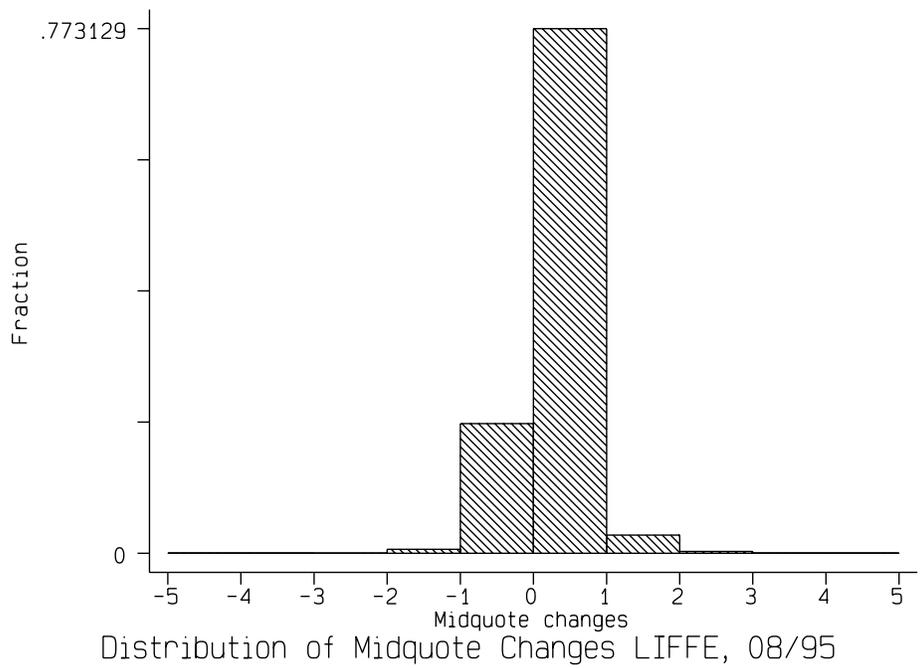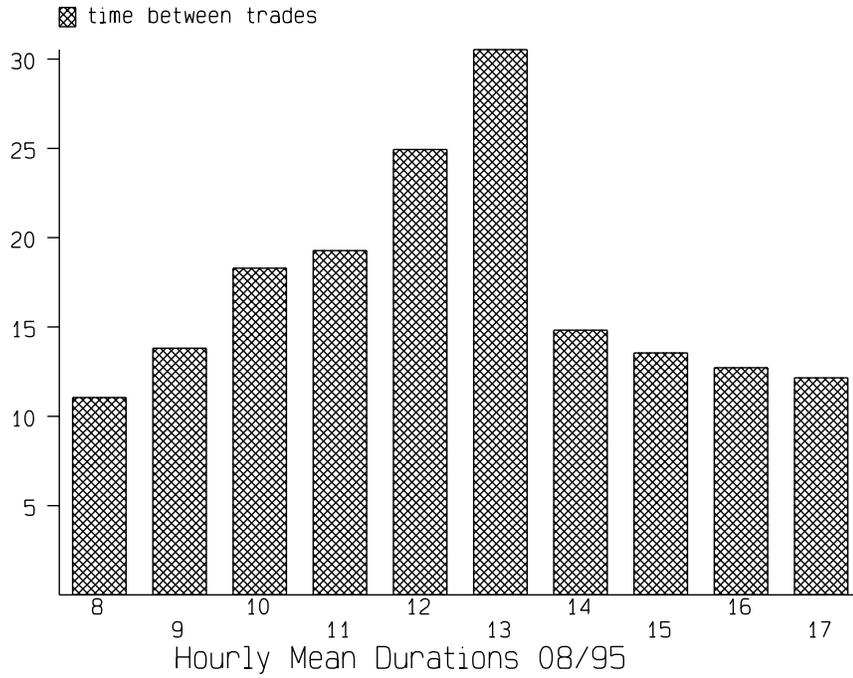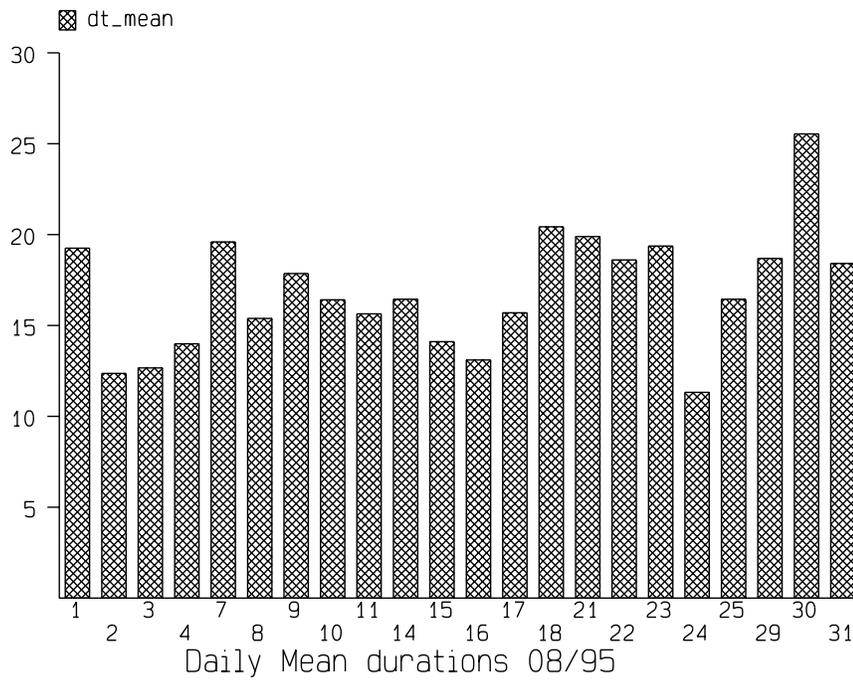
STATA™

Figure 5

Figure 6

Figure 7

9 Estimation and Diagnostic Results

Regression 1		
variables	coeff.	t-val.
Hypothesis H1		
Bid ask spread ≥ 2	-0.0367	-2.760
Hypothesis H2(a)		
Aggregated volumes 0-5 min.	-0.7572	-34.224
Aggregated volumes 5-10 min.	0.0475	2.934
Aggregated volumes 10-15 min.	-0.0130	-0.801
Aggregated volumes 15-20 min.	-0.0605	-3.649
Aggregated volumes 20-25 min.	-0.0014	-0.088
Aggregated volumes 25-30 min.	-0.0327	-2.375
Hypothesis H3		
Lag 1 of log-duration	0.0968	15.506
Lag 2 of log-duration	0.0406	6.936
Lag 3 of log-duration	0.0313	5.292
Lag 4 of log-duration	0.0348	5.773
Further explanatory variables		
Volume ≥ 41	-0.0360	-1.664
No midq. movements. (5 trades)	0.0921	5.431
Midquote changes ≥ 1.5	-0.2266	-4.200
No changes buys-sells (5 trades)	-0.0712	-3.517
Seasonal patterns		
Trades between 10:00 and 12:00	0.0802	4.327
Trades between 12:00 and 14:00	0.2406	10.997
Trades between 14:00 and 16:00	0.0784	4.479
Trades between 16:00 and 17:15	-0.0004	0.022
Trades at Tuesday	-0.0060	-0.302
Trades at Wednesday	-0.0312	-1.581
Trades at Thursday	-0.0345	-1.739
Trades at Friday	0.0154	0.729
Heterogeneity variance θ^{-1}	0.0534	2.478
$\chi^2(23)$		9253.33
Mc Kelvey-Zavoina-R^2		0.1537
Teststat. Partial Ser. Corr. Lag1		1.6977
Teststat. Partial Ser. Corr. Lag2		2.4741
Teststat. Partial Ser. Corr. Lag3		0.6132
Teststat. Partial Ser. Corr. Lag4		1.5145

Table 1: Pooled Regression of LIFFE Bund Future transactions. Gamma Compounded Hazard Model.
Categorization: $[0;1],(1;2],\ldots,(49;50],(50,\infty)$.
Critical value for test-statistic (5% level): 3.841.
N = 38772.

	Regression 2		Regression 3	
variables	coeff.	t-val.	coeff.	t-val.
Hypothesis H1				
Bid ask spread ≥ 2	-0.0504	-3.752	-0.0425	-3.733
Hypothesis H2(b)				
Average volumes 0-5 min.	-0.2477	-4.472	-0.2002	-4.362
Average volumes 5-10 min.	0.0858	1.582	0.1477	2.055
Average volumes 10-15 min.	0.0163	0.267	0.0170	0.267
Hypothesis H3				
Lag 1 of log-duration	0.1284	20.393	0.1295	18.997
Lag 2 of log-duration	0.0704	11.898	0.0782	11.975
Lag 3 of log-duration	0.0606	10.271	0.0624	9.481
Lag 4 of log-duration	0.0650	10.911	0.0634	9.637
Average Lags 5-10	0.2542	19.891	0.2692	18.840
Average Lags 10-20	0.1832	12.397	0.1866	11.292
Average Lags 20-30	0.1050	6.993	0.0996	6.001
Average Lags 30-40	0.0549	3.771	0.0441	2.636
Average Lags 40-50	0.0361	2.499	0.0341	2.056
Average Lags 50-60	0.0517	3.530	0.0602	3.666
Average Lags 60-70	0.0162	1.195	0.0038	0.248
Further explanatory variables				
Volume ≥ 41	-0.0664	-3.006	-0.0425	-1.722
No midq. movements. (5 trades)	0.0856	4.953	0.0969	4.952
Midquote changes ≥ 1.5	-0.2505	-4.645	-0.2964	-5.954
No changes buys-sells (5 trades)	-0.0925	-4.594	-0.0859	-5.954
Seasonal patterns				
Trades between 10:00 and 12:00	0.0642	3.439	0.0440	2.082
Trades between 12:00 and 14:00	0.1387	6.369	0.1043	4.158
Trades between 14:00 and 16:00	-0.0326	-1.874	-0.0403	-2.080
Trades between 16:00 and 17:15	-0.0246	-1.339	-0.0132	-0.642
Trades at Tuesday	-0.0268	-1.327	-0.0159	-0.699
Trades at Wednesday	-0.0425	-2.119	-0.0356	-1.562
Trades at Thursday	-0.0692	-3.466	-0.0640	-2.848
Trades at Friday	-0.0342	-1.600	-0.0333	-1.407
Heterogeneity variance θ^{-1}	0.0671	3.834	0.0641	2.111
$\chi^2(27)$	8183.61		6709.50	
Mc Kelvey-Zavoina-R^2	0.1444		0.1441	
Teststat. Partial Ser. Corr. Lag1	0.0698		0.0122	
Teststat. Partial Ser. Corr. Lag2	0.4081		0.0319	
Teststat. Partial Ser. Corr. Lag3	0.0159		0.0071	
Teststat. Partial Ser. Corr. Lag4	0.2412		0.0068	

Table 2: Pooled Regression of LIFFE Bund Future transactions. Gamma Compounded Hazard Model.
Regression 2: Categorization: [0;1],(1;2],...,(49;50],(50,∞).
Regression 3: Categorization: [0;5],(5;10],(10;20],(20,∞).
Critical value for test-statistic (5% level): 3.841.
N = 38977.

Regression 4		
variables	coeff.	t-val.
Lag 1 of 1og-duration	0.3408	33.606
Heterogeneity variance θ^{-1}	0.3809	10.411
$\chi^2(1)$		2736.96
Mc Kelvey-Zavoina-R^2		0.0561

Table 3:   Pooled Regression of LIFFE Bund Future transactions. Gamma Compounded Hazard Model.
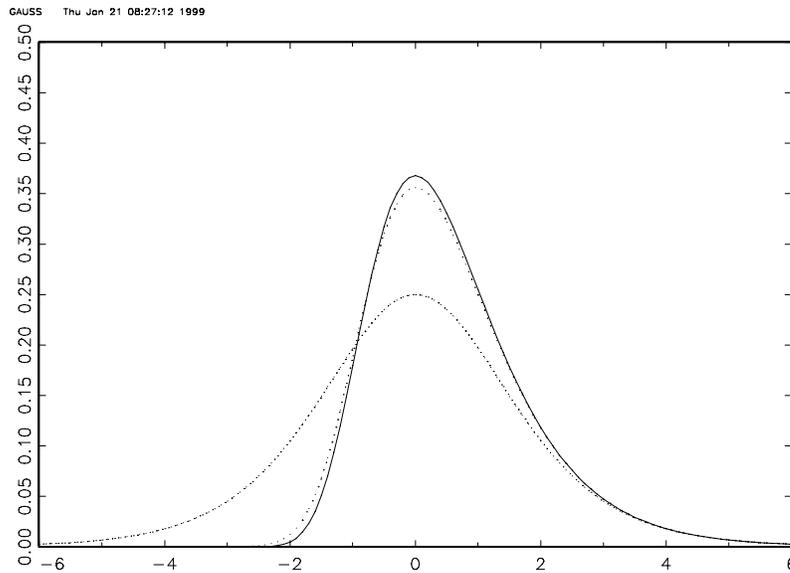Categorization: [0;1],(1;2],...,(49;50],(50,∞).
N = 38977.

Figure 8: Logistic (closely spaced dots), Burr(14.903) (dots) and extreme value distribution (solid).[18]

---

[18]The BurrII(14.903) density is calculated at the values $x - ln(14.903)$.
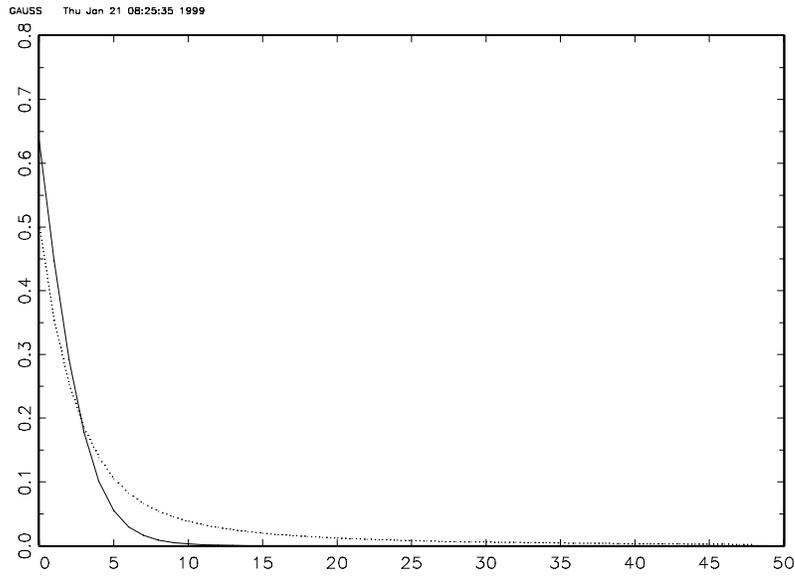
Figure 9: Baseline survivor functions based on the ordered logit (dots) and gamma comp. model (solid).
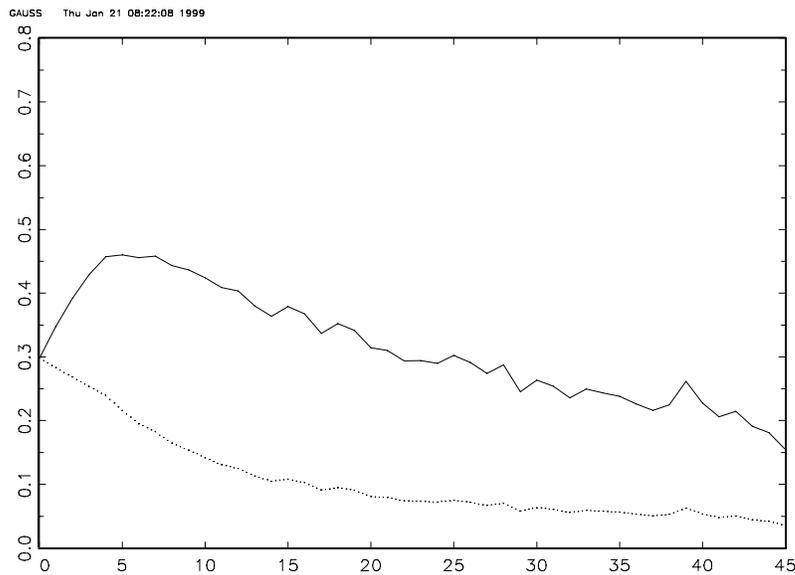
Figure 10: Baseline hazard rates based on the ordered logit (dots) and gamma comp. model (solid).