

Econometric Analysis of Financial Transaction Data: Pitfalls and Opportunities

Nikolaus Hautsch
University of Konstanz
CoFE

Winfried Pohlmeier*
University of Konstanz
CoFE, ZEW

June 18, 2001

Abstract

The recent availability of large data sets covering single transactions on financial markets has created a new branch of econometrics which has opened up a new door of looking at the microstructure of financial markets and its dynamics. The specific nature of transaction data such as the randomness of arrival times of trades, the discreteness of price jumps and significant intraday seasonalities, call for specific econometric tools combining both time series techniques as well as microeconomic techniques arising from discrete choice analysis.

This paper serves as an introduction to the econometrics of transaction data. We survey the state of the art and discuss its pitfalls and opportunities. Special emphasis is given to the analysis of the properties of data from various assets and trading mechanisms. We show that some characteristics of the transaction price process such as the dynamics of intertrade durations are quite similar across various assets with different liquidity and regardless whether an asset is traded electronically or on the floor. However, the analysis of other characteristics of transaction prices process such as volatility requires a careful choice of the appropriate econometric tool. Empirical evidence is presented using examples from stocks traded electronically and on the floor at the German Stock exchange and from BUND future trading at the LIFFE and the EUREX.

JEL classification: C22, C25, C41, G10

Keywords: transaction data, autoregressive conditional duration models, ordered response and count models, electronic and floor trading

*Corresponding author. Department of Economics, Box D124, University of Konstanz, 78457 Konstanz, Germany. Phone ++49-7531-88-2660, Fax -4450, email: winfried.pohlmeier@uni-konstanz.de. We are grateful to Dieter Hess for helpful comments and discussions. For competent research assistance we like to thank Gabriela Schütz and Nadine Warmuth. Financial support by the CoFE is gratefully acknowledged.

1 Introduction

The recent availability of low-cost financial transaction databases has created a new exciting field in econometrics and empirical finance which is able to look at old puzzles in finance from a new perspective and to address a variety of new issues. Transaction data - sometimes referred to as ultra high frequency data - can be seen as the informational limiting case where all transactions are recorded. Since these data contain valuable information on the time between individual transactions (intertrade durations) as well as conditioning information on the type of trades, the issues that can be tackled are manifold. In particular econometric studies based on transaction data can contribute to the empirical assessment of different market designs and institutional settings like the market form (dealership market vs. auction market), the auction mechanism or the trading rules.

Since the seminal work of Engle and Russell (1998) much research effort has been devoted to the econometric analysis of intertrade durations. The time between transactions is an indicator for the speed of the market. Being the reciprocal of the transaction rate, time between trades can serve as a crude measure for volume. In fact, in order driven markets large volumes are split and traded at different prices. If no detailed information on the order book is available, this measure is an attractive alternative to observed volumes that are often split by the matching procedure. One key success factor of an efficient exchange is the provision of liquidity at competitive trading costs. Volume durations, the time it takes to transact a given volume, is an indicator for the time costs of liquidity. On the other hand, price durations, the time it takes to observe a cumulative price change of a given size, can serve as a volatility measure.

Intertrade durations also play a key role in the theoretical understanding of the market microstructure. The informational content of intertrade durations is an issue in the contributions of Diamond and Verrecchia (1987), Admati and Pfleiderer (1988), Foster and Viswanathan (1993), and Easley, Kiefer, and O'Hara (1997) among many others. In Easley and O'Hara (1992) the market maker is a Bayesian learner who infers from intertrade duration whether informed traders are in the market or not. Shorter intertrade durations indicate the market maker that nonidentifiable informed traders are active. Furthermore, from the comparison of cumulated buy and sell orders he may learn the sign of this information. Another important field of investigation relates to the question of how markets perform under stress. In general asset prices adjust to the news events within a short period of time. Since transaction data are time stamped, news events can exactly be assigned to the price process. Therefore they are particularly suited to analyze how financial markets adjust under stress.

Obviously, for a better understanding of the fine structure of the trading process many questions are directly related to the behaviour of transaction prices. For example, in their chapter on market microstructure Campbell, Lo, and MacKinlay (1997) mention the determinants of the bid-ask spread, the relevance of inventory costs relative to adverse selection motive, the price impact of volume and the trader's preferences for round numbers as some of the issues that can be investigated on the basis of transaction data.

The specific stochastic nature of financial transaction data makes the development of appropriate econometric methods an extremely challenging task. The randomness of arrival times of trades as well as the discreteness of price changes and significant intraday seasonalities call for specific econometric methods, combining both time series techniques as well as microeconomic tools. The field of applications of these new econometric tools is, however, much broader. Similar stochastic properties can be found for scanner data from grocery stores and for data from credit rating agencies (e.g. Dunn and Bradstreet, Kreditreform) where information on the firms in the files is updated with every information request of a customer. Being at the outset, much of the current empirical work is mainly of explorative nature. At this stage, the focus of econometric research is on model development and evaluation, i.e. on the search for appropriate econometric models and estimators for specific financial markets and topics of interest.

The goal of this paper is to serve as an introduction to the econometrics of transaction data. We survey the state of the art and discuss the pitfalls and opportunities that are involved with work based on financial transaction data. Thus far, much of the empirical work is based on the NYSE's Trades and Quotes (TAQ) database, which includes all transactions on the NSE, AMEX, NASDAQ and the US regional exchanges. Due to the specific trading mechanism and the comparatively high liquidity of the stocks in the TAQ database, the insights which can be drawn from these studies are only of limited merit for researchers interested in the analysis of European Exchanges. Therefore, special emphasis is given to the analysis of the properties of data from various assets and trading mechanisms. We show that some characteristics of the transaction price process such as the dynamics of intertrade durations are quite similar across various assets with different liquidity and regardless whether an asset is traded electronically or on the floor. However, the analysis of other characteristics of transaction prices process such as volatility requires a careful choice of the appropriate econometric tool.

The structure of the paper is as follows. In Section 2 we discuss the properties of transaction data. Based on data for the highly liquid BUND future and various less

liquid German stocks which are traded on the floor and electronically, we work out the differences that have to be taken into account in the empirical work. Section 3 surveys econometric approaches for the analysis of transaction prices and intertrade durations defined on the calendar time scale. Estimation results are presented in Section 4. Section 5 concludes and gives an outlook on future research.

2 Properties of Transaction Data

2.1 Discreteness of Price Changes

The most prominent feature of transaction data is the discreteness of prices. Since the institutional settings of the great majority of exchanges allow prices to be only multiples of a smallest divisor, called a 'tick', prices and transaction returns take on discrete values. Although not being necessary from a theoretical or practical point of view the basic idea of fixing a minimum price change is to obtain a reasonable trade-off between the provision of an efficient grid for price formation and the possibility to realize price levels that are close to the traders' valuation. The economic aspects of the choice of the tick size has been discussed by Harris (1994).

The minimum tick size varies from asset to asset and also across exchanges. E.g. for the NYSE the minimum tick size is \$ 0.125 for equities, \$ 0.0625 for equity options and \$ 0.05 for futures contracts on the Standard and Poor's 500 index. For equities traded at the Frankfurt stock exchange the minimum tick size varies. As representative examples for transaction data we choose in this paper Allianz and Henkel, which differ substantially in liquidity. Both equities are traded at the German Stock Exchange, Frankfurt, on the floor and by a computer trading system. For the computer based XETRA trading (sample period July 1st to Dec. 30th, 1999) the tick size is 0.01 Euro, while the tick size for the two shares differs on the floor trading. During the sample period Jan. 4th to Dec. 30th, 1999 it is 0.05 Euro for Allianz and 0.01 Euro for the Henkel shares. The third asset we are looking at in this study is the highly liquid BUND future. In our sample period (Nov. 1st to Dec. 5th, 1996) before the denomination in Euro the minimum tick size at the LIFFE and the EUREX (formerly DTB) was 0.01% (one basis point) which corresponds to a face value of DEM 25 (currently it amounts to 10 Euro).

For assets with high transaction rates the discreteness of prices becomes a fundamental feature calling for the application of econometric techniques such as quantal response models or count data models. For instance, the *NYSE Fact Book: 1994 Data* reports that 97.4% of all transactions on the NYSE occurred with either no change or a one-tick change (see Campbell, Lo, and MacKinlay (1997), chapter 4, for more

details on the distribution of transaction prices from the TAQ database). If the transaction rate is low, i.e. only a few transactions are observed within a given time interval, discreteness of transaction prices is less severe. Figures 1 and 2 depict the distribution of absolute price changes for the three assets under investigation distinguished by the type of trading (computer vs. floor trading). Obviously the number of discrete price categories for the less liquid Henkel shares is smaller than for the Allianz share. Comparing the XETRA system with the floor trading system we observe a higher number of relatively large price changes as well as a higher number of zero price changes. E.g. for the XETRA trading of the Allianz stock 27.71% (0.32%) of all observations are zero price changes (1 Euro price changes) while for the floor trading the corresponding numbers are 56.59% and 4.79%. XETRA trading is much more voluminous for these two shares. A similar but more extreme picture arises for the highly liquid BUND future. Here, more than 64 % of all transaction returns at the EUREX are zero. For the floor traded BUND future at the LIFFE the transaction price changes of zero can be found for 46 % of all transactions. For both exchanges price jumps of more than ± 2 ticks are negligible and amount to less than 2 percent of all transaction returns. Comparing the number of observed price categories for the two stocks and the BUND future, we have to conclude that the adoption of a quantal response model would be more appropriate for the latter while the application of a count data approach seems to be a more suitable research strategy for the former.

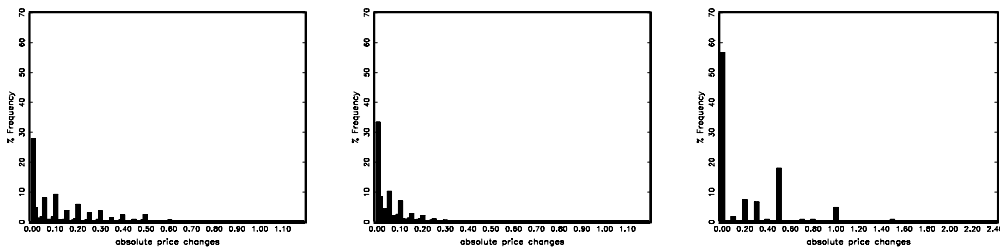


Figure 1: Distribution of absolute price changes. XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96. Left: Allianz, XETRA trading, middle: Henkel, XETRA trading, right: Allianz, floor trading.

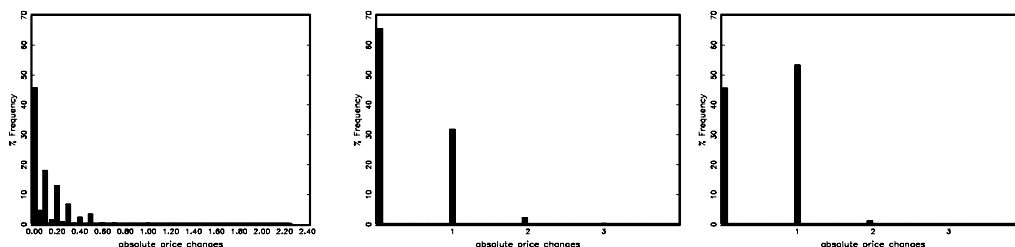


Figure 2: Distribution of absolute price changes. XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96. Left: Henkel, floor trading, middle: BUND future trading, EUREX, right: BUND future trading, LIFFE.

Several authors (e.g. Harris (1990), Dravid (1991), Hasbrouck (1996) and Manrique and Shephard (1998)) have stressed that round numbers for transacted prices systematically occur more often. The multimodal distributions depicted in the figures above are clearly consistent with the hypothesis of preferences for round numbers. In fact, such preferences for round numbers seem to be more pronounced for floor traded assets.

Much attention has also been paid to the implications of the bid-ask bounce of transaction price movements. The bounce effect refers to the phenomenon that transaction returns do not satisfy the weak white noise hypothesis. The negative first order autocorrelations of the transaction price changes reported in Table 3 are quite typical for transaction data. This empirical finding is consistent with the simple model proposed by Roll (1984) who shows that price changes exhibit volatility and negative autocorrelations under randomly trades initiated buy and sell orders even if the fundamental value of the asset is constant. See also Glosten (1987) who derives the impact of adverse selection on the statistical properties of transaction data in a more elaborate theoretical framework. The impact of buyer and seller initiated trades on the dynamics of the transaction price process can easily be detected if autocorrelations from signed trades are computed. If information on the bid-ask quotes is available, the problem of negative first order serial correlation can be resolved by using midquotes.¹ It is needless to stress that the bid-ask bounce should be more severe for low-priced stocks traded a exchanges with comparatively high minimum tick sizes.

¹In the absence of quotes, Ederington and Lee (1995) use 'pseudo-equilibrium prices' by averaging the last two transaction prices.

Table 1: *Autocorrelations of price changes. XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96.*

	A, X	H, X	A, F	H, F	BF, E	BF, L
lag1	-0.259	-0.220	0.005	-0.015	-0.266	-0.326
lag2	-0.020	-0.039	0.006	-0.017	-0.002	0.027
lag3	0.001	-0.014	0.005	-0.011	0.011	0.003
lag4	0.003	-0.006	-0.004	0.002	0.003	0.017

X: XETRA, F: Frankfurt, E: EUREX, L: LIFFE

A: Allianz, H: Henkel, BF: BUND future

Long-term dependence in volatility is a well documented feature of financial data. This pattern, however, cannot always be found at high frequencies. Andersen and Bollerslev (1997) show that while persistence is evident in S&P returns at all frequencies, it cannot be found for the \$-DEM exchange rate at high frequencies.

Looking at simple autocorrelations of absolute price changes in Figure 3 reveals that persistence in volatility is also an issue at the transaction level. The empirical autocorrelations for the three assets are found to be quite small but they are dying out slowly. However, the discreteness of transaction prices might generate these long-run dependence since discrete price jumps may mimic jump effects that are falsely perceived as volatility persistence in models for fractionally integrated time series. Interestingly, the autocorrelations for the floor traded stocks are clearly larger than the ones for the computer traded counterparts. Franke and Hess (2000) argue that traders can learn more about the other traders' strategies, and thus can reduce asymmetric information, when trading on the floor than on anonymous electronic trading systems. This might have an impact on the serial dependence of the time between trades, transaction volumes, and trade-to-trade price changes.

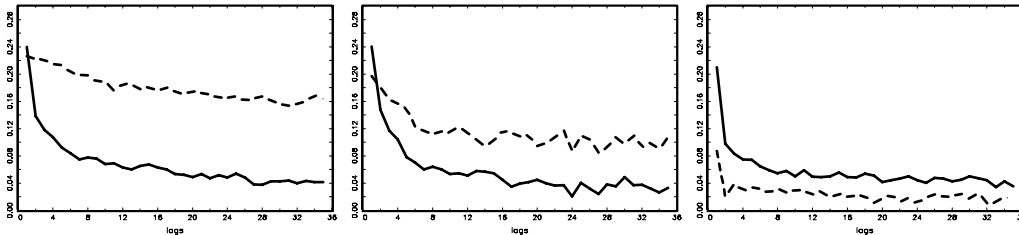


Figure 3: Autocorrelogram of absolute price changes. XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96. Left: Allianz, middle: Henkel, right: BUND future. Solid line: Electronic trading (XETRA or EUREX, respectively), broken line: floor trading (Frankfurt or LIFFE, respectively).

2.2 Intertrade durations

The analysis of intertrade durations is an ongoing topic in the empirical analysis of market microstructure. Intertrade durations measure the speed of the market, and thus, are indicators for the trading activity. Several contributions to the literature of market microstructure, like Easley and O’Hara (1992), Diamond and Verrecchia (1987) or Admati and Pfleiderer (1988) emphasize the importance of intertrade durations for a better understanding of the information processing in financial markets. Within these studies, the timing of trades plays an important role in the learning mechanisms of traders drawing inferences from past market activities. In many theoretical studies, intertrade durations are regarded as means to aggregate information on price signals available to individual traders in an asymmetric information environment (see e.g. Easley and O’Hara (1992)).

In general, researchers analyzing the time between trades are interested in three major aspects. First, analyzing the impact of market microstructure variables, like bid-ask spreads, price changes, transaction volumes, as well as intraday seasonalities allow to check the empirical evidence of market microstructure hypotheses. In particular, such investigations provide deeper insights into traders’ learning and the impact of past and current market activities on traders’ preferences for immediacy. Secondly, modelling the hazard rate of intertrade durations enables one to analyze the informational content of intertrade waiting times. For instance, Gerhard and Hautsch (2000) characterize the economic implications of different shapes of the hazard function and derive a simple relationship between the information process and the resulting trading

process. Third, models of intertrade durations serve as ingredients for multivariate models of the trading process (see e.g. Russell and Engle (1998), Grammig and Wellner (1999), Ghysels and Jasiak (1997) or Gerhard and Pohlmeier (2000)) or as basic models for volatility and liquidity estimation.

Especially in electronic trading systems the accuracy of the recorded trade arrival times is hundredths of seconds. Particular attention should be paid to the treatment of extremely small intertrade durations.² Often such observations correspond to 'split-transactions'. Such observations arise when the volume of one order exceeds the capacities of the first queue of the other side of the limit order book. In this case the order is automatically matched against several opposite order book entries. Typically, the recorded time between the 'sub-transactions' is extremely small and the corresponding transaction prices are equal or show an increasing (or decreasing, respectively) sequence.³ In some studies the particular sub-transactions are treated separately by fixing the corresponding inter-trade durations synthetically on one second. In this paper we consolidate 'split-transactions' by applying an algorithm proposed by Grammig and Wellner (1999). According to this rule a trade is identified as a split-transaction when the durations between the sub-transactions are smaller than one second and the sequence of the prices are either non-increasing (non-decreasing) implying a split transaction on the bid (ask) side of the order book. For simplicity the time stamp and corresponding price of the split-transaction is determined by the last sub-transaction. An alternative and slightly more precise method would be to treat the corresponding inter-trade duration as left-censored and to compute the price as the (volume weighted) average of the prices of the sub-transactions. Note that such a proceeding would lead to a disappearance of the discreteness of the price process.

²In electronic trading systems, even zero intertrade durations are recorded.

³Another way to dealing with extremely small intertrade durations is proposed by Veredas, Rodriguez-Poo, and Espasa (2001). They argue that the occurrence of such observations is due to the fact that the limit orders of many traders are set for being executed at round prices and, thus trades executed in the same second do not belong to the same trader.

Table 2: *Descriptive statistics of intertrade durations. XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96. Durations are measured in seconds.*

	A, X	H, X	A, F	H, F	BF, E	BF, L
obs	89346	28161	43131	12936	64682	53720
Mean	44.543	141.078	168.226	557.466	13.409	14.233
Std.dev.	66.739	207.700	199.271	620.961	24.638	18.501
Min	1	1	1	1	1	1
Max	1075	4219	408	4308	588	887

X: XETRA, F: Frankfurt, E: EUREX, L: LIFFE

A: Allianz, H: Henkel, BF: BUND future

The strong relation to market liquidity is an important feature of intertrade durations. Table 2 shows descriptive statistics of the intertrade durations of the assets analyzed in this study. The results illustrate that the liquidity of the particular assets and the particular markets is quite different. The significantly shortest intertrade durations occur for the two BUND future markets where we observe on average 4 trades per minute. On the XETRA market a relatively liquid stock, like the DAX stock Allianz, is traded on average every 45 seconds while a comparatively illiquid asset like Henkel provides intertrade durations of about 140 seconds on average. In contrast to the XETRA trading, considerably longer intertrade durations can be observed for the Frankfurt floor trading which is nevertheless the most liquid floor trading exchange in Germany.

Another prevalent feature of transaction data is a stochastic clustering of the transaction arrival indicated by a strong serial dependency in the intertrade duration process. The market microstructure theory provides several explanations for this phenomenon. One string of the literature focuses on the existence of two different types of traders: informed traders who trade after price signals which are not publicly available and non-informed traders (liquidity or noise traders) who trade because of exogenous reasons. A common assumption is the existence of an uninformed specialist who updates the quote setting in response to the order flow. If informed traders seek to take advantage of their information, one should observe a clustering of transactions following an information event because of an increased number of informed traders. Another explanation is provided by Admati and Pfleiderer (1988). They partition liquidity traders in 'discretionary' traders who have some choice over the timing of trades and 'nondiscretionary' traders whose trading time is randomly chosen. It is shown that 'discretionary' liquidity trading and thus informed trading is typically concentrated.

However, relatively little is known about the impact of market settings on the duration dynamics. Figure 4 shows the duration correlograms of the particular assets used in this study. The pictures depict higher autocorrelations for intertrade durations of floor trading systems. Hence, the anonymity of traders in electronic trading systems seems to weaken the strength of dynamics in the trading intensity. As in the case of the autocorrelations for absolute price changes long-term persistence is also an issue in the analysis of intertrade durations. Jasiak (1999) argues that the slowly decaying shape of the autocorrelation function might be associated with a fractionally integrated duration process. For this reason she introduces a fractionally integrated ACD model for the analysis of intertrade durations.

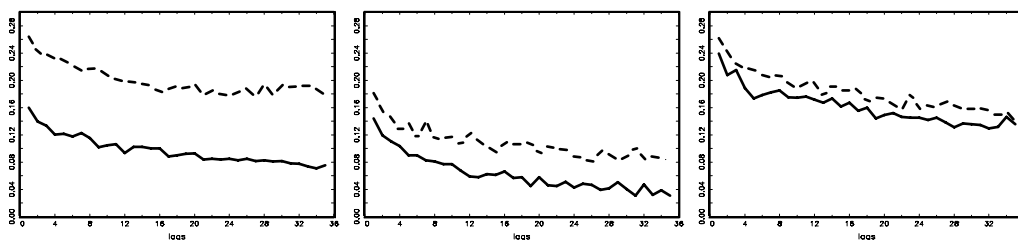


Figure 4: Correlogram of intertrade durations. XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96. Left: Allianz, middle: Henkel, right: BUND future. Solid line: Electronic trading (XETRA or EUREX, respectively), broken line: floor trading (Frankfurt or LIFFE, respectively).

Figure 5 shows the distributions of intertrade durations based on the electronic trading systems and on the floor trading systems. While the density function of durations based on order book trading systems monotonically declines, we observe a slightly hump-shaped pattern for the duration density based on the floor trading. The little hump is a known feature and a typical phenomenon for floor trading transactions. This property is often associated with a certain reaction time caused by the manual registration of the transaction process on the floor.

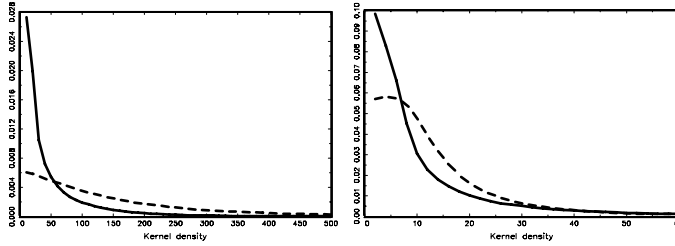


Figure 5: Kernel density plots of intertrade durations (Epanechnikov kernel). XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96. Left: Allianz, right: BUND future. Solid line: Electronic trading (XETRA or EUREX, respectively), broken line: floor trading (Frankfurt or LIFFE, respectively).

2.3 Price and volume durations

While intertrade durations play an important role in market microstructure issues, the aggregation of durations is a valuable means to analyze intraday market activities on an aggregated level. The most common types of aggregated durations are price and volume durations. Price durations are generated by thinning the marked point process with respect to a predetermined minimum price change. Therefore, price durations are defined as the time until a predetermined cumulative price change is realized. As illustrated by Engle and Russell (1998), Gerhard and Hautsch (1999) and Giot (2000b) price durations are strongly related to the intraday volatility process. Since they use an aggregation scheme which is based on the price process such models are a valuable alternative for standard GARCH procedures. These relationships are briefly illustrated as follows. Let τ_i the (calendar) time of transaction i , then the volatility per time at τ_i is defined as

$$\sigma^2(\tau_i) = \text{E} \left[\frac{1}{\Delta} \left(\frac{p(\tau_i) - p(\tau_i - \Delta)}{p(\tau_i)} \right)^2 \right], \quad (2.1)$$

where $p(\tau_i)$ denotes the price at τ_i and Δ corresponds to a certain time interval. Standard GARCH-type procedures are based on equidistant and thus aggregated observations. Therefore, the use of GARCH models implies to fix the time interval Δ , e.g. on intervals of 1 minute, 5 minutes or 30 minutes, hence $p(\tau_i - \Delta)$ corresponds to the price level Δ minutes before the current trade i . Such a procedure raises the question of an optimal aggregation level. Andersen and Bollerslev (1998) illustrate that the choice of an appropriate aggregation scheme is very crucial for these models

and has consequences for the resulting volatility estimates.

A straightforward alternative procedure implies not to fix the time interval Δ but the price change $c \equiv p(\tau_i) - p(\tau_i - \Delta)$, e.g. on 5 ticks. Then Δ is the time until a cumulative price change of 5 ticks is realized and, thus Δ is a random variable. The economic motivation behind this approach is to assume a decision maker who associates a certain cumulative price change with a certain risk. By predetermining the size of the price change c he accounts for a predetermined risk and thus gives the tuning parameter for volatility estimation. While eq. (2.1) gives a volatility per time which is constant within the corresponding price duration, Engle and Russell (1998) derive the instantaneous volatility per second as

$$\sigma^2(\tau_i) = \lim_{\Delta \rightarrow 0} E \left[\frac{1}{\Delta} \left(\frac{p(\tau_i) - p(\tau_i - \Delta)}{p(\tau_i)} \right)^2 \right] = \left(\frac{c}{p(\tau_i)} \right)^2 \lambda(\tau_i - \tau_{i-1}), \quad (2.2)$$

where τ_{i-1} denotes the (calendar) time of the most recent trade and $\lambda(t)$ denotes the hazard rate associated with the corresponding price duration at $t_i = \tau_i - \tau_{i-1}$.

Table 3: *Descriptive statistics of price and volume durations, Allianz, XETRA trading, 07/99-12/99.*

	$\Delta p = 0.5$	$\Delta p = 1.0$	$\Sigma v = 10,000$	$\Sigma v = 20,000$
obs	10777	3063	5407	2824
Mean	367.68	1305.05	732.874	1403.139
Std.dev.	576.56	1859.04	591.166	1026.526
Min	1	1	1	3
Max	11369	17346	4599	7210

Descriptive statistics of price durations based on different price changes for the Allianz stock (XETRA trading) are given in Table 3. For the extreme case ($c=1.0$ Euro) we observe, on average, 25 trades per day. If investors predetermine the Value-at-Risk (VaR) associated with a given large (negative) price movement the expected time until the occurrence of such a price change can be interpreted as a risk measure.⁴

⁴For more details concerning the use of high-frequency data for Value-at-Risk estimation see e.g. Giot (2000a).

A different way of aggregating durations arises from the use of volume durations. Volume durations are defined as the time until a certain aggregated volume is traded on the market. Gouriéroux, Jasiak, and LeFol (1999) illustrate that volume durations provide reasonable liquidity measures. Taking the usual definition of liquidity, an asset is considered as liquid if it can be traded quickly, in large quantities and with little impact on the price. This implies that liquidity is associated with three dimensions of the transaction process. Since the measurement of the price impact is quite difficult and requires detailed order book information, volume durations account for the time and volume dimension and may serve as building blocks for reasonable liquidity measures based on transaction data.

Consider e.g. an investor who wants to trade a large volume as quick as possible. In a dealership market he has the possibility to trade with the market maker and hence executes his transaction immediately. The investor has to bear liquidity costs which arise through the difference between the market price and the ask or bid quote, respectively. This price increment above (below) the market price can be interpreted as the price (liquidity costs) for immediacy of a transaction. If the investor wants to avoid these liquidity costs he has to distribute the volume over time, i.e. he has to wait until movements on the demand or supply side of the market allow to trade with lower transaction costs. For electronic trading systems liquidity is characterized in a similar fashion. Here, the absorptive capacities of the order queues in the limit order book determine the liquidity. Thus, the larger the volume an investor wants to trade, the higher the probability that it exceeds the capacity of the first queue of the limit order book leading to a price impact and, thus costs for immediacy. Therefore, liquidity is also strongly related to the depth of the market. These costs can be reduced by splitting the order and trading lower volumes. Hence, the waiting time necessary to execute an order of a given size admits a reasonable interpretation as the (time) costs of liquidity.

Descriptive statistics of particular volume durations based on volume aggregates of 10,000 and 20,000 shares are given in Table 3. Note that the average volume per trade for the Allianz stock is about 680 shares. Thus a liquidity measure based on 10,000 shares corresponds to relatively short term measure while volume durations based on 20,000 shares capture quite long market phases of, on average, about 20 minutes.

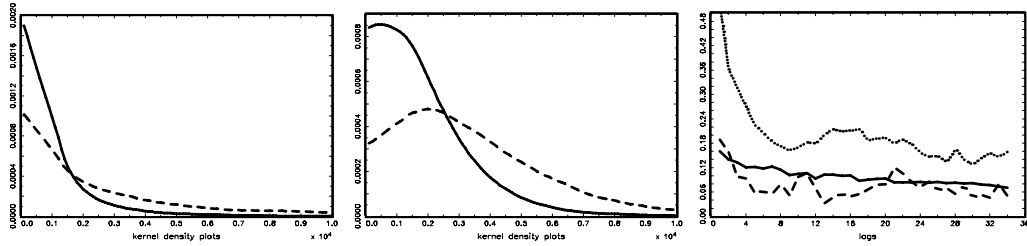


Figure 6: Aggregated durations of Allianz, XETRA trading, 07/99-12/99. Left: Kernel density plots of price durations. Solid line: 0.50 EURO price changes, bars: 1.00 EURO price changes. Middle: Kernel density plots of volume durations. Solid line: 10,000 shares, bars: 20,000 shares. Right: Correlogram of aggregated durations: Solid line: Intertrade durations, bars: 1.00 DEM price durations, dots: 20,000 shares volume durations.

Figure 6 presents the correlograms of price and volume durations based on different aggregation levels as well as the corresponding kernel density plots. While the distribution of price durations is relatively similar to the distribution of intertrade durations the density function of volume durations is quite different. Comparing price and volume durations three main differences can be summarized: First, while the distribution of price durations is relatively similar to the distribution of intertrade durations, the density function of volume durations is hump-shaped. Secondly, while price durations as well as inter-trade durations reveal overdispersion, volume durations show a strong underdispersion. Both aspects have important consequences for the choice of distributional assumptions when the density function has to be modelled. Thirdly, volume durations show a significantly higher autocorrelation at the first lags while the long-term persistence seems to be lower.

2.4 Intraday Seasonality

Financial markets exhibit a strong seasonality within a trading day. Figure 7 shows typical intradaily seasonal patterns for intertrade durations based on spline regressions⁵. There is little trading around noon leading to longer intertrade durations. This 'lunch time'- effect appears to be more pronounced for the floor trading. There is also evidence that floor trading starts off somewhat more relaxed. Attention should be paid to the occurrence of intraday auctions which typically take place around noon (13.00) and at the end of a trading day.

⁵We use cubic splines based on knots of 1 hour.

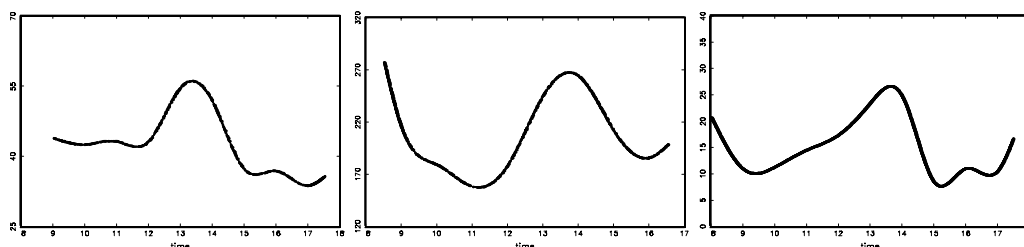


Figure 7: Intraday seasonality of intertrade durations. Estimation based on cubic splines (1 hour knots). XETRA trading, 07/99-12/99, floor trading, Frankfurt, 01/99-12/99, BUND future trading, EUREX, Frankfurt, and LIFFE, London, 11/96-12/96. Left: Allianz, XETRA trading, middle: Allianz, floor trading, right: BUND future trading, LIFFE.

Figure 8 shows the typical pattern for the intraday seasonalities of absolute price changes and volume. While there is hardly any seasonality of volatility within a trading day, the seasonality of the transaction volume cannot be ignored. The impact of intraday auctions at noon and before the closure of the exchange on traded volumes is nicely documented for the Allianz stocks below.

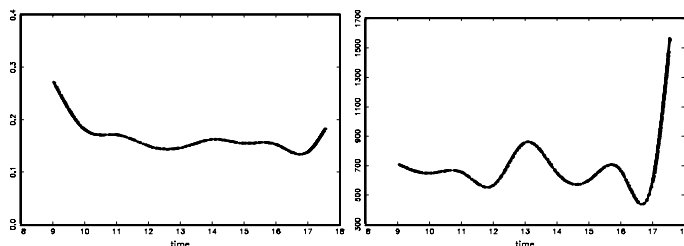


Figure 8: Left: Intraday seasonality of absolute price changes and transaction volumes, Allianz, XETRA trading, 07/99-12/99. Left: Seasonality of absolute price changes, right: seasonality of transaction volumes.

While the problem of accounting for seasonalities is well explored for the case of equidistant observations, there is little experience with the modelling of seasonalities in the context of transaction data. Engle and Russell (1998) remove intraday seasonalities from the intertrade durations by applying a piecewise cubic spline. Due to numerical problems that arise by estimating the parameters of an autoregressive conditional duration model jointly with the seasonal effects it is common to apply a

two-step estimation approach. In the first step the intertrade durations are seasonally filtered and in the second step the parameters are estimated on the basis of the deseasonalized dependent variable. Alternatively, kernel estimates (Gourieroux, Jasiak, and LeFol (1999)) and Fourier series approximation (Gerhard and Hautsch (1999)) have been used to remove intraday seasonalities in the first step of the estimation. Veredas, Rodriguez-Poo, and Espasa (2001) point out that two-step procedures can lead to serious misspecifications unless seasonal and non-seasonal components depend on some deterministic time index and the non-seasonal components are linear in the parameters to be estimated. They propose a semiparametric estimator where the seasonal components are jointly estimated non-parametrically with the parameters of the ACD model.

Obviously for quantal response and count data models deseasonalization by a two-step procedure is infeasible. In this case the joint estimation of seasonal and non-seasonal components is inevitable. Pohlmeier and Gerhard (2001), for instance, use the Fourier series approximation proposed by Andersen and Bollerslev (1998) based on the work of Gallant (1981) in their ordered probit model with conditional heteroskedasticity to account for seasonalities in the volatility of transaction price changes. Assuming a polynomial of degree Q the nonstochastic seasonal trend term is of the form

$$s(\delta, t_i^*, Q) = \delta \cdot t_i^* + \sum_{q=1}^Q (\delta_{c,q} \cos(t_i^* \cdot 2\pi q) + \delta_{s,q} \sin(t_i^* \cdot 2\pi q)), \quad (2.3)$$

where δ , $\delta_{c,q}$, and $\delta_{s,q}$ are the seasonal coefficients to be estimated and $t_i^* \in [0, 1]$ is a normalized intradaily time trend defined as the number of seconds from opening of the exchange until occurrence of transaction i divided by the length of the trading day.

3 Models of the Transaction Price Process

The treatment of non-equidistant time series data has generated two types of research strategies which differ in their use of the time scale. Calendar time τ , e.g. measured in seconds from the opening, takes on continuous nonnegative values while intrinsic time z (sometimes also called deformed time, market time or business time) takes on discrete nonnegative values. The directing process that maps calendar time to intrinsic time is denoted by $Z : \tau \in \mathcal{R}^+ \rightarrow Z_\tau \in \mathcal{N}$. If $(Y_\tau, \tau \in \mathcal{R}^+)$ is the variable of interest in calendar time, then $(Y_z^*, z \in \mathcal{N})$ would be the corresponding time transformed variable in intrinsic time. Usually, time deformation models assume that the directing process Z and the process in intrinsic time Y^* are independent, see e.g. Stock (1988) and Ghysels, Gourieroux, and Jasiak (1997). Even under this independence assumption time deformation is by no means harmless since the stochastic properties of the two variables in calendar time and intrinsic time may differ. Gourieroux and Jasiak (2000)

show that time deformation has serious implications for the market efficiency hypothesis. In particular, if it holds in the calendar time framework it may not necessarily hold in the intrinsic time framework and vice versa.

In the following we restrict our attention to approaches defined in calendar time. Let a single trade i be characterized by a number of different jointly occurring marks. Well-known examples are transaction prices, quotes or midquotes that go along with other covariates such as volume, absolute price changes and intertrade durations. Let Y_i be a m -dimensional vector of marks of a single trade i . This trade takes place at real time τ_i . Thus the time between two consecutive trades, the intertrade duration is given by $T_i \equiv \tau_i - \tau_{i-1}$. Let the transaction process be fully described by the sequence:

$$\{(Y_i, T_i)\}_{i=1, \dots, n}. \quad (3.1)$$

Following the framework outlined in Engle (2000) it is meaningful to start from defining the joint density of marks and intertrade durations conditional on the past filtration \mathcal{F}_{i-1} as:

$$(Y_i, T_i) | \mathcal{F}_{i-1} \sim f_{Y,T}(y_i, t_i | \check{y}_{i-1}, \check{t}_{i-1}; \theta), \quad (3.2)$$

where $\theta \in \mathcal{R}_k$ and \check{x}_i and \check{t}_i , respectively, represents the values of the variable x and t , respectively, up to the i -th transaction. Much of the current work on the econometric modelling of the transaction process focuses on the analysis either of the components of Y_i or -like the vast majority of studies - on the transaction durations. The link between these two approaches becomes obvious by decomposing the joint density as the product of the conditional density of the marks given the intertrade durations and the marginal density of the intertrade durations:

$$f_{Y,T}(y_i, t_i | \check{y}_{i-1}, \check{t}_{i-1}; \theta) = f_{Y|T}(y_i | \check{y}_{i-1}, \check{t}_i; \theta_1) f_T(t_i | \check{y}_{i-1}, \check{t}_{i-1}; \theta_2), \quad (3.3)$$

where the parameters of the conditional and the marginal density are related to θ by a transformation $g(\theta)' \equiv [\theta'_1, \theta'_2]$. Using this decomposition the log likelihood is given by

$$\mathcal{L}(\theta_1, \theta_2) = \sum_{i=1}^n \ln f_{Y|T}(y_i | \check{y}_{i-1}, \check{t}_i; \theta_1) + \sum_{i=1}^n \ln f_T(t_i | \check{y}_{i-1}, \check{t}_{i-1}; \theta_2). \quad (3.4)$$

There is no obvious reason to assume before hand that θ_1 and θ_2 are variation free so that intertrade durations can be treated as being weakly exogenous and subsets of θ can be estimated by solely concentrating on the conditional density of the marks. Although most of the econometric studies are based on decompositions such as (3.3), there is no reason from an economic or statistical point of view to use a decomposition in terms of the conditional duration model and a marginal density of the marks. In order to reduce the complexity of the model, the conditioning set is often restricted

to impose some type of non-causality. For instance, Darolles, Gouriéroux, and LeFol (1998) assume that only past prices contain relevant information for the joint process of marks and intertrade durations such that a non-causality from intertrade durations to prices is assumed: $f_{Y|T}(y_i|\check{y}_{i-1}, \check{t}_i; \theta_1) = f_{Y|T}(y_i|\check{y}_{i-1}; \theta_1)$. Symmetrically, the literature on intertrade durations imposes an unidirectional non-causality from prices to intertrade durations. In this case lagged durations are fully informative to predict current durations $f_T(t_i|\check{y}_{i-1}, \check{t}_{i-1}; \theta_2) = f_T(t_i|\check{t}_{i-1}; \theta_2)$. Obviously, both non-causality assumptions can be made subject to specification testing.

3.1 Transaction Price Models

Transaction price models in real time concentrate on the modelling of some type of price variable conditional on the current intertrade duration and past filtration. Hausman, Lo, and MacKinlay (1992) propose the ordered probit model with conditional heteroskedasticity to model discrete price movements at the transaction level for stock prices quoted at the NYSE. Bollerslev and Melvin (1994) apply the same approach to the analysis of quotes on FX markets. The ordered probit is particularly suitable for the case of (highly liquid) markets where transaction price changes take on only a few distinct values. The inclusion of conditioning information is straightforward in order to account for factors assumed to drive the price process. This is a substantial advantage compared to the rounding models of Ball (1988), Cho and Frees (1988), or Harris (1990). Like in any other threshold crossing ordered response model, the model of interest is defined in terms of a continuous latent dependent variable y_i^* , e.g. the price pressure, that is only observable through an ordered response variable y_i , i.e. the categorized price change variable:

$$\begin{aligned}
 y_i^* &= x_i' \beta + \epsilon_i & (3.5) \\
 \text{with } E[\epsilon_i | x_i] &= 0, \\
 \epsilon_i &\sim \text{i.n.i.d.} N(0, \sigma_i^2), \\
 \sigma_i &= \sigma_0 \exp(w_i' \gamma),
 \end{aligned}$$

where the $(K \times 1)$ and $(L \times 1)$ vectors x_i and w_i contain the conditioning variables for the mean and the variance function. Conditional on the set of explanatory variables, we assume that the latent variable is mutually independent. Since no restrictions are imposed on the stochastic process of x_i and w_i , the price process may well reveal unconditional serial dependence. Gerhard (2000) shows that the model can be extended to include an ARMA-dynamics in terms of the latent dependent variable. Contrary to the coefficients of dynamic generalized linear model the coefficients of the latent ARMA-model are easily interpretable. Latent price pressure and observable discrete

price change are related by the following observation rule:

$$y_i = \begin{cases} -k_l & \text{if } y_i^* \in (-\infty; \alpha_1) \\ \vdots & \\ -1 & \text{if } y_i^* \in (\alpha_{l-1}; \alpha_l] \\ 0 & \text{if } y_i^* \in (\alpha_l; \alpha_{l+1}] \\ +1 & \text{if } y_i^* \in (\alpha_{l+1}; \alpha_{l+2}] \\ \vdots & \\ k_u & \text{if } y_i^* \in (\alpha_{l+k_u}; \infty), \end{cases} \quad (3.6)$$

where the α'_j s are unknown threshold parameters that separate the state space of y_i^* . Ticks larger or equal to a given size k_u are gathered in the uppermost category of y_i . Correspondingly, ticks smaller or equal to size $-k_l$ are captured in the category $y_i = -k_l$. Assuming no intercept in the mean function and $J + 1$ categories the model consists of $J + K + L + 1$ parameters. However, since the parameters of an ordered response model are only identifiable up to a factor of proportionality, only the parameter vector $\lambda = [\alpha'/\sigma_0 \quad \beta'/\sigma_0 \quad \gamma']$ is identifiable without any additional restrictions. Besides the problem of shifting the focus of interest from the observable discrete price change variable to the continuous latent counterpart which has no straightforward economic interpretation the identification issue may be regarded as one of the major drawbacks in the application of quantal response models, particularly, if the focus of interest is on estimating volatility at the transaction level. Gerhard, Hess, and Pohlmeier (1998) propose a minimum distance approach applied to the intraday estimates to identify σ_0 and the volatility of the latent price variable relative to a benchmark period. Alternatively, Pohlmeier and Gerhard (2001) suggest a nonlinear restriction to identify σ_0 that relates the variance of the observable discrete price change variable to the variance of the latent counterpart.

As indicated in the previous section, transaction price changes at many European stock markets are not really characterized by a small number of discrete price jumps. In this case, adopting a count data approach or a multinomial model seems to be a reasonable research strategy. Contrary to quantal response models no threshold parameters have to be estimated, which may become numerically difficult in the case of many categories. A natural starting point is the Poisson regression model for the analysis of absolute price changes. Since the Poisson distribution belongs to the linear exponential family, the Pseudo ML (PML) estimates of the mean function are robust against distributional misspecification. Because of the large number of observations efficiency considerations can usually be neglected. Moreover, count data models can easily be interpreted and can be extended in many ways to capture the nature of the

underlying data generating process.⁶ In the following section we also present estimates of the zero-inflated Poisson model (ZIP) which adds additional mass to the pdf at zero counts and nests the simple Poisson model. The ZIP can easily be extended to account also for additional mass at realizations of the count variable in order to capture the higher probabilities of round transaction returns. The density of the ZIP is given by:

$$\begin{aligned}\Pr[y_i = 0 | x_i] &= \varphi_i + (1 - \varphi_i) \exp(-\lambda_i), \\ \Pr[y_i = j | x_i] &= (1 - \varphi_i) \frac{\exp(-\lambda_i) \lambda_i^j}{j!}, \quad j = 1, 2, \dots\end{aligned}\tag{3.7}$$

where $\lambda_i = \exp(x_i' \beta)$. In the application below the inflation parameter φ_i is treated as constant across all observations but it can also be modelled as a function of covariates. For $\varphi_i > 0$ the ZIP implies overdispersion which usually can be found in transaction count data. The ZIP model boils down to the Poisson model if $\varphi_i = 0$.

3.2 Models of the Trade Intensity

A well-known feature of transaction based durations is the occurrence of clustering, i.e. short (long) durations tend to be followed by short (long) durations. For this reason, financial duration models originate from traditional time series concepts. The most popular autoregressive duration approach was proposed by Engle (1996)⁷ and Engle and Russell (1998). They specify a model for point processes with serial dependent arrival rates which shows a strong resemblance to the GARCH approach for price processes. The main principle of the Autoregressive Conditional Duration (ACD) model is a dynamic parameterization of the conditional mean function

$$\begin{aligned}\Psi_i &= E[t_i | \mathcal{F}_{i-1}; \theta_2] \\ &= \int s \cdot f_T(s | \check{t}_{i-1}, \check{y}_{i-1}; \theta_2) ds.\end{aligned}\tag{3.8}$$

In its standard form the ACD(p, q) model is defined as

$$\Psi_i = \omega + \sum_{j=1}^p \phi_j t_{i-j} + \sum_{j=1}^q \psi_j \Psi_{i-j},\tag{3.9}$$

with the error term ϵ_i entering multiplicatively:

$$t_i = \Psi_i \cdot \epsilon_i.\tag{3.10}$$

⁶See Cameron and Trivedi (1998) for an comprehensive survey of count data models.

⁷The paper is now published as Engle (2000).

Engle and Russell (1998) illustrate that the $ACD(p, q)$ model can be written in terms of an ARMA model

$$t_i = \omega + \sum_{i=1}^{max(p,q)} (\phi_i + \psi_i)t_{i-j} - \sum_{i=1}^q \psi_j \eta_{i-j} + \eta_i, \quad (3.11)$$

where $\eta_i \equiv t_i - \Psi_i$ is a martingal difference.

The most obvious choice for the distribution of the error terms ϵ_i is the standard exponential distribution which allows for parsimony and yields robust parameter estimates since the maximum likelihood estimates of the Exponential ACD model have PML properties. Assuming exponentially distributed errors, the pseudo-true log likelihood function is given by

$$\ln \mathcal{L}_{PLM} = - \sum_{i=1}^n \ln \Psi_i + \frac{t_i}{\Psi_i}. \quad (3.12)$$

Several recent contributions to the literature deal with the extension of the ACD framework. One string of the literature focuses on alternative specifications of the conditional mean function. Bauwens and Giot (2000) and Lunde (2000) introduce a logarithmic ACD model which allows for the inclusion of explanatory variables without accounting for parameter restrictions due to non-negativity conditions.

Bauwens and Veredas (1999) specified the conditional mean function (3.9) stochastically leading to the Stochastic Conditional Duration (SCD) model which implies more flexible dynamics and, because of its property as mixture model, more flexible distributions of the underlying duration process. Zhang, Russell, and Tsay (1999) propose a Threshold Autoregressive Conditional Duration (TACD) model to account for state dependent dynamics and to capture nonlinear relationships between the conditional expected duration and past information variables. Dufour and Engle (2000) propose more general functional forms, like a Box-Cox-ACD model or an ACD models based on a piecewise linear conditional mean function. They show that the functional form of the conditional mean function has a strong influence on the predictive performance of ACD models.

Another string of contributions deals with ACD models based on more flexible distributions, like Grammig and Maurer (2000) who introduce the Burr-ACD model, Lunde (2000) who specifies the Generalized Gamma ACD model or Hautsch (2001) who proposes an ACD model based on the Generalized F distribution. A semiparametric ACD model has been suggested by Drost and Werker (2000).

The inclusion of explanatory variables requires to account for two aspects: First, a direct inclusion of explanatory variables in the conditional mean function requires to ensure parameter restrictions due to the non-negativity condition. Thus, a straightforward specification arises by an exponential form. Secondly, explanatory variables may enter the model dynamically or statically. Therefore, by using the Log-ACD form of Bauwens and Giot (2000), the conditional mean function becomes

$$\Psi_i = \exp \left(\omega + \sum_{j=1}^p \phi_j \ln t_{i-j} + \sum_{j=1}^q \psi_j \ln \Psi_{i-j} + x'_i \beta \right) \quad (3.13)$$

for the dynamic case. Specifying the explanatory variables outside the dynamics one obtains

$$\begin{aligned} \tilde{\Psi}_i &= \frac{\Psi_i}{\exp(x'_i \beta)}, \\ \text{with: } \tilde{\Psi}_i &= \exp \left(\omega + \sum_{j=1}^p \phi_j \ln t_{i-j} + \sum_{j=1}^q \psi_j \ln \tilde{\Psi}_{i-j} \right). \end{aligned} \quad (3.14)$$

Note that (3.13) implies that the explanatory variables enter the model with an infinite lag structure.

Engle and Russell (1998) illustrate that the standard ACD model (3.9) implies a conditional hazard rate of the duration t_i which can be written in terms of the hazard rate λ_0 of the ACD residual ϵ_i

$$\lambda(t_i | \mathcal{F}_{i-1}) = \lambda_0 \left(\frac{t_i}{\Psi_i} \right) \frac{1}{\Psi_i}, \quad (3.15)$$

which can be estimated nonparametrically, due to the PML properties of the Exponential ACD model.

Eq. (3.15) illustrates that duration models of the ACD type belong to the class of accelerated failure time models, i.e. explanatory variables (here the duration history) accelerate or decelerate the time to failure, and in this context the time between trades.

A more traditional way of specifying the duration process is to characterize it directly in terms of the hazard rate, e.g. by using the popular proportional hazard specification. While covariates in accelerated failure time models deform the time scale, they change the hazard rate in the proportional hazard framework. In financial studies based on transaction data estimates of hazard rates and survivor functions can be used to characterize the trading activity (see Gouriéroux, Jasiak, and LeFol (1999)). The application of this class of models is useful if one has to deal with censored

observations. While the inclusion of censoring effects in hazard rate models is relatively straightforward, it is quite cumbersome in the ACD framework. In order to account for serial dependencies in financial durations Gerhard and Hautsch (2001) propose a method to include autoregressive structures in semiparametric proportional hazard models. They illustrate that such specifications serve as valuable tools to construct individual and specific risk measures. Especially estimates of the probability to observe a certain price change within a given time interval can be used for risk management and portfolio strategies.

3.3 Multivariate Specifications

Because of the complexity of the estimation problem, only a few studies focus on modelling the joint transaction process for marks and intertrade durations. Since the marks are usually qualitative variables and intertrade durations are continuous variables defined over the positive domain there is no obvious probability density function which also allows for a straightforward introduction of dynamics in the marks and the intertrade durations. Rydberg and Shephard (1998) employ a discrete time framework by decomposing the transaction duration into a sequence of binary indicators for periods in which a trade occurs or not. Given the occurrence of a trade, positive price changes are modelled on the basis of a dynamic count data model. Trade occurrence and price changes conditional on trade occurrence are both modelled within the generalised linear modeling framework. Russell and Engle (1998) propose to combine the ACD with a generalized Markov Chain for the transition from a price change of one category to another in order to obtain a joint model of price changes and time between transactions in a continuous time framework. The model proposed by Gerhard and Pohlmeier (2000) combines the two approaches sketched above by accounting for the discrete nature of price changes while remaining in the continuous time framework and keeping the advantage of a parsimonious transaction duration model. A bivariate system of trade intensities and limit order arrival times is analyzed by Russell (1999). Since the multivariate filtration of arrival times is somewhat arbitrary to parameterize in terms of a likelihood decomposition his approach is to derive a joint likelihood for the bivariate process.

4 Some Empirical Illustrations

4.1 Estimating the Process of Price Changes

Based on the simple Poisson model and the ZIP model we present empirical evidence for the determinants of volatility at the transaction level under two different trading mechanisms. Our dependent variable is the absolute price change from transaction

$i - 1$ to transaction i measured in the number of ticks. As explanatory variables we use lagged absolute price changes, intertrade duration and volume including lagged values as well as intradaily seasonality dummies⁸.

Table 4 contains the estimation results for the simple Poisson regressions where heteroskedasticity and autocorrelation robust standard errors have been computed as GMM standard errors (see e.g. Cameron and Trivedi (1998), chap. 7.3). The results illustrate that past absolute price changes have a strong positive impact on volatility. The overall impact of intertrade durations on volatility is positive for all shares. For the XETRA trading we find that longer intertrade durations lead to higher and larger price jumps which are somewhat dampened for the following trades. The instantaneous effect of intertrade durations can also be found for the floor trading. However, a clear pattern for the impact of lagged intertrade durations on volatility cannot be found. Our results nicely confirm that liquidity is an issue for both shares at both market places. Volume has a significant impact on absolute price changes in all four cases. For Allianz, the more liquid share, the price impact of volume is only contemporaneous at both trading platforms. Contrarily, prices for the Henkel shares are also affected by lagged volumes. As already indicated graphically in Figure 8, left panel, the intradaily seasonal pattern of absolute price changes is only mildly pronounced so that three time dummies suffice to capture volatility differences during the trading day. The highest volatility can be observed for the opening phase. This time period is even more volatile than the closing phase. Apparently, there are no large differences in the seasonal volatility patterns with regard to the trading mechanism.

⁸Because absolute price changes show only weak intraday seasonality patterns (compare to Figure 8) we restrict the seasonality variables to three dummies capturing the trading period until 10.00, 10.00-13.00 and 13.00-15.00

Table 4: *Regression results for the Poisson Model. Dependent variable: Absolute price changes in number of ticks ($|\Delta p|_i$). P-values based on robust standard errors. Data sets: (a) Allianz, XETRA trading, (b) Henkel, XETRA trading, (c) Allianz, Frankfurt, (d) Henkel, Frankfurt. Diagnostics: ACF and Ljung-Box statistic of the first three lags of the standardized Poisson residuals.*

	(a)		(b)		(c)		(d)	
	coeff.	p-val.	coeff.	p-val.	coeff.	p-val.	coeff.	p-val.
const	1.616	0.000	0.256	0.000	-2.479	0.000	-1.186	0.000
$ \Delta p _{i-1}$	0.751	0.000	1.547	0.000	0.279	0.000	0.768	0.000
$ \Delta p _{i-2}$	0.224	0.000	0.556	0.000	0.270	0.000	0.438	0.000
$ \Delta p _{i-3}$	0.365	0.000	0.645	0.000	0.257	0.000	0.414	0.000
$\log t_i$	0.165	0.000	0.204	0.000	0.201	0.000	0.192	0.000
$\log t_{i-1}$	-0.056	0.000	-0.043	0.000	-0.004	0.020	-0.013	0.000
$\log t_{i-2}$	-0.032	0.000	-0.013	0.000	-0.002	0.238	0.021	0.002
$\log t_{i-3}$	-0.034	0.000	-0.023	0.000	0.009	0.000	-0.000	0.831
$\log v_i$	0.126	0.000	0.100	0.000	0.499	0.000	0.386	0.000
$\log v_{i-1}$	-0.002	0.193	0.038	0.000	0.003	0.107	-0.016	0.000
until 10.00	0.141	0.000	0.190	0.000	0.223	0.000	0.169	0.000
10.00-13.00	-0.055	0.000	-0.082	0.000	-0.060	0.000	-0.070	0.000
13.00-15.00	-0.011	0.002	-0.015	0.042	0.091	0.000	-0.076	0.000
Obs	89346		28161		43131		12936	
LLH	2.72e7		2.01e6		1.56e6		2.54e6	
AC Lag1	0.049		0.062		0.025		0.032	
AC Lag2	0.031		0.040		0.012		0.026	
AC Lag3	-0.005		0.015		0.010		0.017	
LB(3)	310.09		162.32		38.55		26.95	

In order to evaluate the goodness-of-fit with respect to the dynamics found in the process of absolute price changes we check for serial correlation in the (standardized) Poisson residuals. The first three lags of the autocorrelation function as well as the corresponding Ljung-Box statistic indicate a negligible serial dependence in the Poisson residuals. Nevertheless, the autocorrelation coefficients are clearly significant which is not surprising, given the considerable sample sizes.

Table 5 gives the estimation results of the ZIP model. The sign pattern of the estimated coefficients is rather similar to the one for the simple Poisson model. The estimates of the inflation parameter φ are between .43 and .72 and significantly different from zero indicating overdispersion in the absolute price changes for all four shares. With respect to the explanatory variables we find coefficients which are very similar to the Poisson estimates. The diagnostics based on the standardized ZIP residuals given

by

$$z_i \equiv \frac{y_i - (1 - \varphi_i)\lambda_i}{\sqrt{(1 - \varphi_i)(\lambda_i + \varphi_i\lambda_i^2)}},$$

with $\lambda_i = \exp(x_i'\beta)$, show slightly higher serial correlations in the residual process than for the Poisson regression. However, the serial dependence is quite small, indicating that the inclusion of past absolute price changes seem to be sufficient to capture the main body of the process dynamics.

Table 5: *Regression results for the Zero-Inflated Poisson Model. Dependent variable: Absolute price in number of ticks ($|\Delta p|_i$). P-values based on robust standard errors. Data sets: (a) Allianz, XETRA trading, (b) Henkel, XETRA trading, (c) Allianz, Frankfurt, (d) Henkel, Frankfurt. Diagnostics: ACF and Ljung-Box statistic of the first three lags of the standardized ZIP residuals.*

	(a)		(b)		(c)		(d)	
	coeff.	p-val.	coeff.	p-val.	coeff.	p-val.	coeff.	p-val.
const	2.494	0.000	1.719	0.000	0.774	0.000	1.613	0.000
$ \Delta p _{i-1}$	0.495	0.000	1.206	0.000	0.231	0.000	0.563	0.000
$ \Delta p _{i-2}$	0.223	0.000	0.472	0.000	0.235	0.000	0.408	0.000
$ \Delta p _{i-3}$	0.229	0.000	0.538	0.000	0.217	0.000	0.319	0.000
$\log t_i$	0.088	0.000	0.114	0.000	0.088	0.000	0.093	0.000
$\log t_{i-1}$	-0.028	0.000	-0.022	0.000	-0.016	0.000	-0.006	0.000
$\log t_{i-2}$	-0.021	0.000	-0.014	0.000	0.000	0.000	0.019	0.002
$\log t_{i-3}$	-0.021	0.000	-0.017	0.000	0.007	0.000	-0.004	0.831
$\log v_i$	0.046	0.000	0.029	0.000	0.121	0.000	0.087	0.000
$\log v_{i-1}$	-0.001	0.556	-0.009	0.000	0.035	0.403	0.005	0.656
until 10.00	0.094	0.000	0.133	0.000	0.156	0.000	0.154	0.000
10.00-13.00	-0.046	0.000	-0.073	0.000	-0.027	0.000	-0.018	0.000
13.00-15.00	-0.016	0.041	-0.024	0.042	0.026	0.213	-0.054	0.407
φ	0.722	0.000	0.666	0.000	0.435	0.000	0.542	0.000
Obs	89346		28161		43131		12936	
LLH	6.20e6		1.16e6		9.30e5		4.88e5	
AC Lag1	0.070		0.095		0.077		0.072	
AC Lag2	0.035		0.069		0.075		0.070	
AC Lag3	0.010		0.038		0.074		0.046	
LB(3)	564.80		435.59		746.66		160.94	

Table 6 shows the regression results for the ordered probit with conditional heteroskedasticity for the BUND Future trading of the EUREX and the LIFFE. As dependent variable we use (raw) price changes based on the categorization $(-\infty; -2]$, $-1, 0, 1, [2; \infty)$.

In the mean function we include the first three lags of price changes. The conditional variance function depends on the contemporaneous and lagged inter-trade duration and transaction volume as well as seasonality variables based on the flexible Fourier form (eq. 2.3) of order $Q = 5$.

Concerning the past price changes we find significantly negative coefficients which strongly is in support of the existence of a bid-ask bounce effect. Almost all variables in the variance function are highly significant. The positive coefficient of the contemporaneous volume is economically quite reasonable. It indicates that volatility and thus the uncertainty is higher the larger the time between two consecutive trades. For the lagged intertrade durations we find a significantly negative impact on the variance, thus the higher past market activities the higher the current conditional variance. A converse effect is observed for the transaction volume, thus the higher the contemporaneous as well as the past trading volume, the lower the volatility. This result is quite interesting since it indicates that high volumes tend to go along with less volatile market periods.⁹ For the seasonality variables we find highly significant coefficients, and thus empirical evidence for the existence of strong seasonalities in the intraday variance process. The plot of the estimated flexible Fourier form for the seasonalities (not presented here) shows an inverted U-shape for the daily pattern of the volatility per transaction.

To check for dynamic misspecification we use a test on first order serial correlation in the error term of the latent model based on generalized residuals. Gouriéroux, Monfort, and Trognon (1987) generalize the concept of the residual to qualitative and limited dependent variable models with a latent dependent variable that belongs to the exponential family and show that a wide range LM-tests can be expressed in terms of their concept of generalized residuals. Following this idea we compute the autocorrelations of the resulting generalized residuals and the corresponding Ljung-Box statistic. The diagnostics indicates that serial dependencies are negligible.

⁹Note that we omit the volume variables in the LIFFE regression. This is quite reasonable since the recording of the trading volume at the LIFFE is extremely inaccurate.

Table 6: Regression results for the ordered probit with conditional heteroskedasticity. Dependent variable: (Categorized) price changes in ticks. Categories: $(-\infty; -2]$, -1 , 0 , 1 , $[2; \infty)$. Data sets: (a) Bund-Future trading, EUREX, (b) Bund-Future trading, LIFFE. Diagnostics: ACF and Ljung-Box statistic of the first three lags of generalized residuals.

	(a)		(b)	
	coeff.	p-val.	coeff.	p-val.
Mean function				
Δp_{i-1}	-0.458	0.000	-0.433	0.000
Δp_{i-2}	-0.108	0.000	-0.122	0.002
Δp_{i-3}	-0.003	0.307	-0.024	0.000
Thresholds				
μ_1	-3.835	0.000	-3.932	0.000
μ_2	-2.597	0.000	-2.264	0.000
μ_3	-0.838	0.000	-1.229	0.000
μ_4	0.424	0.000	0.377	0.000
Variance function				
$\log t_i$	1.182	0.000	-0.560	0.000
$\log t_{i-1}$	-0.248	0.000	-0.151	0.001
$\log t_{i-2}$	-0.282	0.000	-0.131	0.003
$\log t_{i-3}$	-0.222	0.000	-0.157	0.000
$\log v_i$	-0.359	0.000		
$\log v_{i-1}$	-0.061	0.007		
δ	-0.209	0.000	-0.059	0.067
$\delta_{1,s}$	-0.118	0.000	-0.059	0.000
$\delta_{2,s}$	-0.022	0.002	-0.019	0.017
$\delta_{3,s}$	-0.014	0.011	0.000	0.473
$\delta_{4,s}$	-0.030	0.000	-0.026	0.000
$\delta_{5,s}$	0.016	0.002	0.004	0.285
$\delta_{1,c}$	0.083	0.000	0.059	0.000
$\delta_{2,c}$	0.016	0.000	0.002	0.330
$\delta_{3,c}$	0.004	0.198	-0.014	0.020
$\delta_{4,c}$	0.014	0.002	-0.002	0.347
$\delta_{5,c}$	-0.009	0.035	-0.018	0.002
Obs	64679		89346	
LLH	5.97e4		6.20e6	
AC Lag1	0.012		0.001	
AC Lag2	0.002		0.001	
AC Lag3	0.000		0.001	
LB(3)	10.77		9.98	

4.2 Estimating Intertrade Durations

Table 7 gives the estimation results of the Log-ACD model applied to the intertrade durations of the Henkel and the Allianz shares for XETRA and Frankfurt floor trading. Unlike the vast majority of empirical studies on intertrade durations which explain trade intensities solely by past filtration of trade intensities rather than accounting for the impact of other marks of the trading process, we include volume and lagged absolute price changes in a dynamic version of an Log-ACD(1,1) specification. To adjust for seasonalities we use a two-step procedure where in the first step the seasonalities are estimated based on cubic splines (1 hour knots). With some reservations we may conclude that both variables contribute to the explanation of trade intensities and should not be ignored if information on these variables is available. The sign pattern of the corresponding coefficients, however, is ambiguous so that a clear economic interpretation of these effects is premature.

Table 7: *Regression results for the Log-ACD model with dynamically included explanatory variables. P-values based on robust standard errors. Data sets: (a) Allianz, XETRA trading, (b) Allianz, floor trading Frankfurt, (c) Bund-Future trading, EUREX, Frankfurt, (d) Bund-Future trading, LIFFE, London. Diagnostics: Mean and standard deviation of ACD residuals. Ljung-Box statistics of the first 20 lags of ACD residuals.*

	(a)		(b)		(c)		(d)	
	coeff	p-value	coeff	p-value	coeff	p-value	coeff	p-value
ω	0.108	0.000	0.0515	0.0000	0.0486	0.1672	0.1065	0.0000
ϕ	0.048	0.000	0.0397	0.0000	0.0389	0.0000	0.0753	0.0000
ψ	0.939	0.000	0.9225	0.0000	0.9485	0.0000	0.9167	0.0000
$\log v_i$	-0.012	0.000	-0.0034	0.0491	0.0028	0.0034	-0.0218	0.0000
$ \Delta p _{i-1}$	0.012	0.001	0.0873	0.0004	-0.0181	0.0329	-0.0194	0.0000
Obs	89346		28161		12936		64682	
LLH	82138.904		27254.694		12398		57052.693	
Mean $\hat{\epsilon}_i$	1.000		1.002		1.013		1.000	
Std.dev. $\hat{\epsilon}_i$	1.354		1.374		1.115		1.303	
LB(20) $\hat{\epsilon}_i$	730.237		234.283		700.787		268.905	
LB(20) t_i	-15024.551		2323.730		3106.386		29535.599	

Interestingly, the coefficients ϕ and ψ , which pick up the dynamics of the model, are quite similar across trading mechanisms and shares. We find that not only the dynamics of the trade intensity for the same asset traded on the floor and electronically is

surprisingly similar, but also that the dynamics of assets with considerable differences in liquidity (BUND future vs. Allianz) are resembling. The two coefficients nearly add up to one indicating that long-term persistence is an issue in the modelling of intratrade durations. Given that the mean function of the Log-ACD model is correctly specified, our estimates are robust against dynamic misspecification in the residuals. To obtain correct inference, even under dynamic misspecifications, we use Newey and West (1987) robust standard errors. However, the diagnostics based on the ACD residuals indicate that correlations in the residual are still present and thus calling for richer dynamics or alternative specifications of the conditional mean function. Since these estimates serve primarily as illustrations, more sophisticated specifications are beyond the scope of the paper.

5 Conclusions

This paper presents a partial survey on the econometrics of transaction data. We focus on models defined in the calendar time dimension which combine microeconomic and time series tools. Complementary directions of research, not surveyed here, are dedicated to the development of appropriate econometric models in continuous (calendar) time and in intrinsic time (time deformation models). Much of the current work is concerned with the proper modelling of the underlying stochastics of the transaction price process. Generalizations of these approaches with respect to estimation methods, functional form and dynamics are necessary. A demanding but profitable research task will be the development of multivariate specifications for the analysis of the joint dynamics of markets and of several marks of a trading process. Multivariate duration models can account for the arrival times of different types of trades (e.g. buyer and/or seller initiated trades). Such extensions may also serve as the methodological basis for an analysis of order book dynamics and the relationship between the trade and quote process.

Experience in applied work concerning the performance and the benefits of particular model specifications and estimators is still limited. Thus future research needs to stress the comparison and evaluation of existing duration models and other microeconomic tools with respect to goodness-of-fit, prediction performance and robustness.

In order to accomplish the full value of transaction data more research has to be dedicated to tackle questions raised by the literature on market microstructures. Studies on the quote formation and the price process can generate insights about adverse-selection costs of market-making. The analysis of the limit order book dynamics based on transaction data may provide valuable insights on the cost and benefits of a particular market design. The measurement of liquidity is rendered difficult, as it requires

accounting for the dimensions time, volume and price. For instance, the econometric techniques surveyed above could be valuable for the development of liquidity measures useful for specific market scenarios (e.g. when options expire, after block trades or in extremely volatile market phases like after announcements of news releases). Another application is related to Value-at-Risk (VaR) concepts. Value-at-Risk is an important quantitative tool used to assess financial risks in terms of the potential trading loss of a trader or a bank. There is little experience with respect to the development of VaR measures based on transaction data. The specification of VaR concepts appropriate for investors or market makers operating on an intraday basis seems a promising research topic.

As has been pointed out in detail empirical studies on transaction data can help to assess the differences between specific markets and their trading systems and the linkages between different markets. The pure description price discovery process on these markets by econometric studies using transaction data may yield a typology of financial markets in terms of risk and liquidity. In the light of the current trend towards merging and restructuring of exchanges in Europe and elsewhere the comparison of market designs and institutional settings on the micro level is of high relevance not only for financial economists but also for macroeconomists and policy makers.

References

- ADMATI, A., AND P. PFLEIDERER (1988): “A Theory of Intra-Day Patterns: Volume and Price Variability,” *Review of Financial Studies*, 1, 3–40.
- ANDERSEN, T. G., AND T. BOLLERSLEV (1997): “Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns,” *Journal of Finance*, 52, 975–1005.
- (1998): “Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies,” *Journal of Finance*, 53, 219–265.
- BALL, C. (1988): “Estimation Bias Induced by Discrete Security Prices,” *Journal of Finance*, 43, 841–865.
- BAUWENS, L., AND P. GIOT (2000): “The Logarithmic ACD Model: An Application to the Bid/Ask Quote Process of two NYSE Stocks,” *Annales d’Economie et de Statistique*, 60, 117–149.
- BAUWENS, L., AND D. VEREDAS (1999): “The Stochastic Conditional Duration Model: A Latent Factor Model for the Analysis of Financial Durations,” Discussion paper, CORE.
- BOLLERSLEV, T., AND M. MELVIN (1994): “Bid-Ask Spreads and Volatility in the Foreign Exchange Market - An Empirical Analysis,” *Journal of International Economics*, 36, 355–372.
- CAMERON, A. C., AND P. K. TRIVEDI (1998): *Regression Analysis of Count Data*. Cambridge University Press, Cambridge.
- CAMPBELL, J. Y., A. W. LO, AND A. C. MACKINLAY (1997): *The Econometrics of Financial Markets*. Princeton University Press.
- CHO, D. C., AND E. W. FREES (1988): “Estimating the Volatility of Discrete Stock Prices,” *Journal of Finance*, 43, 451–466.
- DAROLLES, S., C. GOURIEROUX, AND G. LEFOL (1998): “Intraday Transaction Price Dynamics,” Discussion paper, CREST.
- DIAMOND, D. W., AND R. E. VERRECCHIA (1987): “Constraints on Short-Selling and Asset Price Adjustment to Private Information,” *Journal of Financial Economics*, 18, 277–311.

- DRAVID, A. (1991): “Effects of Bid-Ask Spread and Price Discreteness on Stock Returns,” Discussion paper, Wharton School.
- DROST, F. C., AND B. J. M. WERKER (2000): “Efficient Estimation in Semiparametric Time Series: The ACD Model,” Discussion paper, CentER, Tilburg University.
- DUFOUR, A., AND R. F. ENGLE (2000): “The ACD Model: Predictability of the Time Between Consecutive Trades,” Discussion paper, ISMA Centre, University of Reading.
- EASLEY, D., N. M. KIEFER, AND M. O’HARA (1997): “One Day in the Life of a Very Common Stock,” *Review of Financial Studies*, 10, 805–835.
- EASLEY, D., AND M. O’HARA (1992): “Time and the Process of Security Price Adjustment,” *Journal of Finance*, 47, 577–607.
- EDERINGTON, L. H., AND J. H. LEE (1995): “The Short-Run Dynamics of the Price Adjustment to New Information,” *Journal of Financial and Quantitative Analysis*, 30, 117–134.
- ENGLE, R. (2000): “The Econometrics of Ultra-High-Frequency Data,” *Econometrica*, 68, 1, 1–22.
- ENGLE, R. F. (1996): “The Econometrics of Ultra-High Frequency Data,” pp. 1–22.
- ENGLE, R. F., AND J. R. RUSSELL (1998): “Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data,” *Econometrica*, 66, 1127–1162.
- FOSTER, F. D., AND S. VISWANATHAN (1993): “Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models,” *Journal of Finance*, 48, 187–211.
- FRANKE, G., AND D. HESS (2000): “Electronic Trading Versus Floor Trading,” *Journal of Empirical Finance*, 7, 455–478.
- GALLANT, R. A. (1981): “On the Bias in Flexible Functional Forms and an Essential Unbiased Form: The Fourier Flexible Form,” *Journal of Econometrics*, 15, 211–245.
- GERHARD, F. (2000): *Empirical Models of the Intraday Process of Price Changes and Liquidity - A Transaction Level Approach*. Ph.D. Thesis, University of Konstanz.
- GERHARD, F., AND N. HAUTSCH (1999): “Volatility Estimation on the Basis of Price Intensity,” Discussion Paper 99/19, CoFE, Konstanz.

- (2000): “Determinants of inter-trade durations and hazard rates using proportional hazard ARMA models,” Discussion Paper 00/20, CoFE, Konstanz.
- (2001): “Semiparametric dynamic proportional hazard models,” *Mimeo*.
- GERHARD, F., D. HESS, AND W. POHLMEIER (1998): “What a Difference a Day Makes: On the Common Market Microstructure of Trading Days,” Discussion Paper 98/01, CoFE, Konstanz.
- GERHARD, F., AND W. POHLMEIER (2000): “On the Simultaneity of Absolute Price Changes and Transaction Durations,” Discussion paper, Paper prepared for the CoFE Conference on Intertemporal Finance, Konstanz October 6 and 7, 2000.
- GHYSELS, E., C. GOURIEROUX, AND J. JASIAK (1997): “Trading Patterns, Time Deformation and Stochastic Volatility in Foreign Exchange Markets,” in *Nonlinear Modelling in Foreign Exchange Markets*, ed. by C. Dunis, and B. Zhou. Wiley, New York.
- GHYSELS, E., AND J. JASIAK (1997): “GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model,” Discussion Paper 97s-06, CIRANO, Montreal.
- GIOT, P. (2000a): “Intraday Value-at-Risk,” Discussion paper, Mimeo.
- (2000b): “Time Transformations, Intraday Data and Volatility Models,” Discussion Paper 9944, CORE.
- GLOSTEN, L. (1987): “Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices,” *Journal of Finance*, 42, 1293–1307.
- GOURIEROUX, C., AND J. JASIAK (2000): *Financial Econometrics*. forthcoming.
- GOURIEROUX, C., J. JASIAK, AND G. LEFOL (1999): “Intra-Day Market Activity,” *Journal of Financial Markets*, 2, 193–226.
- GOURIEROUX, C., A. MONFORT, AND A. TROGNON (1987): “Generalized Residuals,” *Journal of Econometrics*, 34, 5–32.
- GRAMMIG, J., AND K.-O. MAURER (2000): “Non-Monotonic Hazard Functions and the Autoregressive Conditional Duration Model,” *Econometrics Journal*, 3, 16–38.
- GRAMMIG, J., AND M. WELLNER (1999): “Modeling the Interdependence of Volatility and Inter-Transaction Duration Processes,” *Discussion Paper, Humboldt-University, Berlin*, 21.

- HARRIS, L. (1990): “Estimation of Stock Variances and Serial Covariances from Discrete Observations,” *Journal of Financial and Quantitative Analysis*, 25, 291–306.
- HARRIS, L. E. (1994): “Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes,” *Review of Financial Studies*, 7, 149–178.
- HASBROUCK, J. (1996): “Modeling Market Microstructure Time Series,” in *Handbook of Statistics, Vol. 14*, ed. by G. S. Maddala, and C. R. Rao, chap. 22, pp. 647–692. Elsevier Science.
- HAUSMAN, J. A., A. W. LO, AND A. C. MACKINLAY (1992): “An Ordered Probit Analysis of Transaction Stock Prices,” *Journal of Financial Economics*, 31, 319–379.
- HAUTSCH, N. (2001): “Analyzing liquidity dynamics using ACD models,” *Mimeo*.
- JASIAK, J. (1999): “Persistence in Intraday Durations,” Discussion paper, York University, Montreal.
- LUNDE, A. (2000): “A Generalized Gamma Autoregressive Conditional Duration Model,” Discussion paper, Aalborg University.
- MANRIQUE, A., AND N. SHEPHARD (1998): “Simulation Based Likelihood Inference for Limited Dependent Processes,” Discussion paper, Nuffield College, Oxford.
- NEWBY, W. K., AND K. D. WEST (1987): “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix,” *Econometrica*, 55, 703–708.
- POHLMAYER, W., AND F. GERHARD (2001): “Identifying Intraday Volatility,” in *Beiträge Zur Mikro- und Zur Makroökonomik, Festschrift Zum 65. Geburtstag Von Hans Jürgen Ramser*, ed. by S. Berninghaus, and M. Bräulke. Springer.
- ROLL, R. (1984): “A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market,” *Journal of Finance*, 39, 1127–1139.
- RUSSELL, J. R. (1999): “Econometric Modeling of Multivariate Irregularly-Spaced High-Frequency Data,” Discussion paper, University of Chicago Graduate Business School.
- RUSSELL, J. R., AND R. F. ENGLE (1998): “Econometric Analysis of Discrete-Valued, Irregularly-Spaced Financial Transactions Data Using a New Autoregressive Conditional Multinomial Model,” Discussion paper, presented at Second International Conference on High Frequency Data in Finance, Zurich, Switzerland.

- RYDBERG, T. H., AND N. SHEPHARD (1998): “Dynamics of Trade-by-Trade Price Movements: Decomposition and Models,” Discussion paper, Nuffield College, Oxford, presented at Symposium on Microstructure and High Frequency Data, Paris, France.
- STOCK, J. (1988): “Estimating Continuous Time Processes Subject to Time Deformation,” *Journal of the American Statistical Association*, 83, 77–84.
- VEREDAS, D., J. RODRIGUEZ-POO, AND A. ESPASA (2001): “On the (Intradaily) Seasonality and Dynamics of a Financial Point Process: A Semiparametric Approach,” Discussion Paper CORE Discussion Paper 2001/XX, Université catholique de Louvain.
- ZHANG, N. Y., J. RUSSELL, AND R. S. TSAY (1999): “A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data,” Discussion paper, University of Chicago.