MONASH 10/94

MONASH
UNIVERSITY

AUSTRALIA

TAILING THE BID-ASK SPREAD

Paul Kofman and Ton C.F. Vorst

Working Paper No. 10/94

July 1994

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 7326 0399 4

TAILING THE BID-ASK SPREAD

Paul Kofman

Working Paper No. 10/94

July 1994

DEPARTMENT OF ECONOMETRICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

Tailing The Bid-Ask Spread

by

Paul Kofman
Department of Econometrics
Monash University
Clayton, VIC 3168
Australia
Phone: 61 (3) 905-5847

and

Ton C.F. Vorst
Econometric Institute
Erasmus University Rotterdam
P.O.Box 1738
3000 DR ROTTERDAM
The Netherlands
Phone: 31 (10) 408-1285

April, 1994

We gratefully acknowledge comments and suggestions by Laurens de Haan and James Moser. Generous data support was provided by Claudio Capozzi at the London International Financial Futures Exchange.

TAILING THE BID-ASK SPREAD

ABSTRACT

This paper discusses an application of a rather novel technique for the estimation of the tails of return distributions for financial assets. This extreme value approach proves to be particularly useful when assessing characteristics of high frequency (tick-by-tick) transaction data. Tail parameter estimates allow derivation of probabilities of large price changes. These probabilities improve optimal setting of bid-ask spreads based on the order processing component of the bid-ask spread. Estimates for optimal levels are compared to 'observed' bid-ask spreads. The latter, which are estimates in itself, are based on recently developed methods in the literature.

1. INTRODUCTION

Market makers quote bid-ask spreads to get compensation for the provision of liquidity to the market. Glosten [1987] and Glosten and Harris [1988] decompose this spread into two parts: an asymmetric information part and a gross profit part. The last part accounts for inventory holding costs, order processing costs and normal profit. Asymmetric information in particular, as first discussed by Bagehot [1971] and subsequently formalized by, among others Copeland and Galai [1981] and Glosten and Milgrom [1985], imposes a risk component on the market maker in case of sudden large price changes. Trading with better informed investors, without knowing to, entails the risk that shortly or immediately after trading the market moves against the market maker. This would turn his position into a substantial loss. Even without this asymmetric trading type, large price swings indicate potential risks related to the gross profit component. When value changes quickly, market makers often cannot trade on the other side demanding liquidity. Thus, it may take several transactions before open positions can be reversed.

This paper introduces an optimizing technique for setting a 'fair' (assuming zero normal profits) bid-ask spread. We explicitly do not make any claim on distinguishing the above mentioned components, and confine the analysis to the required compensation for large price changes.

The majority of the literature assumes that price processes of financial asset prices can be described by geometric Brownian motions. This implies that, for fixed time intervals, the changes of the logarithms of prices are normally distributed. However, there is an extensive body of literature (see e.g. Mandelbrot [1963], Fama [1963] and Hall, Wade Brorsen and Irwin [1989]) showing that the empirical distribution functions of these so-called returns are characterized by much fatter tails than implied by normal distributions. This is called excess leptokurtosis (a relatively large probability mass at the tails and around the mean). Since market makers particularly fear sudden large price changes, it is obvious that to them the tails are the most relevant part of the distribution. Hence, this paper investigates tail behavior in relation to market makers' risk valuation. Assuming certain regularity conditions for the tail behavior, a tail can be characterized by one single parameter. Several distributions within the same class can have identical tail parameters. This is particularly true for sums of independently and identically distributed random variables (e.g. daily and weekly observations). The tail parameter furthermore allows deduction of probabilities for the underlying random variables to exceed some pre-specified high value without knowledge of the underlying distribution, i.e. they are robust against misspecification in the exact distribution.

As noted before, these tail probabilities obviously affect market makers' behavior. In addition, market makers are not so much concerned with price changes during fixed time intervals, as with price changes between consecutive transactions. Once again assuming a geometric Brownian motion, these changes can no longer be modeled by a (fixed) normal distribution. However, if we assume that transaction times are generated by an exponential distribution the following inference can be made: the tail parameter for the returns between two successive transactions must be identical to the tail parameter of returns for fixed time intervals.

After identification of the appropriate tail parameter, the estimate can be usefully employed in deriving the probability of a price change significantly exceeding the bid-ask spread. This provides some insight in the three sources feeding the bid-ask spread. It does, however, not allow discrimination between the different components. The tail approach is applied to transaction data for the Bund futures contract traded at the London International Financial Futures Exchange (LIFFE). After identifying the tail shape, we next compare the tail distribution results with the bid-ask spread. Roll's (1984) well known approach, generalized to continuous time data, is used to estimate this bid-ask spread. In addition, a more efficient method by George, Kaul and Nimalendran (1991) is also tried and found to improve the estimates in the sense that they approach market announced average spreads. LIFFE operates an open outcry (OOC) system for the trading interval 7.30 hours until 16.15 hours. Afterwards an automated pit trading (APT) system takes over from 16.20 hours until 18.00 hours. An interesting question is whether return distributions are driven by the same process for these distinct trading systems. We will restrict ourselves to the question whether the tail behavior differs.

The tail estimation procedure has been applied before to daily settlement prices in e.g. Kofman and de Vries [1989] for futures and in Jansen and De Vries [1991] for stocks and stock indices like the S&P-500 index, but to our knowledge the procedure for transaction data is new.

The paper is organized as follows. The next section explains the most relevant results in the extremal value theory and shows that the tail index is identical for tick-by-tick transaction data and 5-minute time spaced data. Section 3 gives estimates for the tail index while section 4 discusses the implications of the estimates concerning probabilities of large price changes in relation to the bid-ask spread. Section 5 concludes the paper.

2. THEORETICAL TAILS FOR HIGH FREQUENCY DATA

This section presents the most relevant theoretical issues of this paper in relation to extremal value theory. For more details the reader should consult Kofman and De Vries [1989] on which this section heavily draws. Consider a sequence of independent and identically distributed random variables $X_1,\ldots,X_n$ with distribution function F and let M_n be the maximum of these n random variables. F^n is the distribution function of M_n. Consider the following regular variation condition at infinity (see Feller [1971], Ch. VIII.8]).

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)}=x^{-\alpha} \tag{1}$$

for some $0<\alpha<\infty$ and all $x>0$ where, of course, it is assumed that F has no finite upper endpoint. If F fulfills this condition, it can be shown that there exist normalizing constants $a_n>0$ and b_n such that

$$F^n\left[\frac{x}{a_n}+b_n\right]\overset{W}{\rightarrow} G(x) \tag{2}$$

where W stands for weak convergence and

$$\begin{aligned} G(x)&=0 & x\leq 0\\ &=\exp(-x^{-\alpha}) & x>0 \end{aligned} \tag{3}$$

The parameter α is called the tail index of the distribution and G is called a limit law. It follows from De Haan [1976] that $\int_1^\infty t^\beta dF(t)$ is finite for $\beta<\alpha$ and infinite for $\beta>\alpha$. Since all moments exist for normal distributions, α does not have a limit law like

G. In fact, it can be proved that for the normal distribution

$$F^n\left[\frac{x}{a_n}+b_n\right] \overset{w}{\rightarrow} \exp(-\exp^{(-x)}) \tag{4}$$

for suitable constants a_n and b_n. In this case, the tail parameter is defined as $+\infty$.

The same argument holds for mixtures of normals. It illustrates that the distributions with a limit law given by (3) have much fatter tails than normal distributions since

$$\lim_{x\to\infty}\frac{1-\exp(\exp(-x))}{1-\exp(-x^{-\alpha})}=0 \tag{5}$$

However, both Student-t and sum-stable distribution functions satisfy the regular variation condition. Similarly, ARCH processes have a limit law as specified in (3). For the sum-stable distributions $\alpha<2$ and for both Student-t and ARCH processes $\alpha>2$. Hence, most alternative distributions (to the normal distribution) that have been proposed in the literature for stock and futures returns, have a limit law as specified by (3).

Instead of estimating the parameter directly, Hill [1975] proposes an estimator for $\gamma=1/\alpha$. Let $X_{(1)}\geq X_{(2)}\geq\ldots\geq X_{(n)}$ be the descending order statistics from a sample $X_1,X_2,\ldots,X_n$ of n empirical observations. Define

$$\hat{\gamma}=\frac{1}{m-1}\sum_{i=1}^{m-1}\left[\log X_{(i)}-\log X_{(m)}\right] \tag{6}$$

If the number (m) of tail observations in a sample of size n approaches infinity, i.e. $m(n)\to\infty$, Mason [1982] shows that $\hat{\gamma}$ is a consistent estimator, as long as the regular variation condition is satisfied. Furthermore, Goldie and Smith [1987] indicate that $(\hat{\gamma}-\gamma)m^{1/2}$ is asymptotically normal with mean zero and variance γ^2. Hence, one can derive confidence intervals for $\hat{\gamma}$. For finite n it is not obvious how to choose m. Kofman and De Vries [1989] propose a procedure to choose an optimal m. The same procedure is applied in this paper.

The tail index itself might not be of too much interest to the market maker. However, it enables him to estimate levels for which there is a small probability that price changes will exceed these particular levels. Let this level $\hat{x}_p$, for small p and given k, be defined as

$$P\{X_1 \le \hat{x}_p, \ldots, X_k \le \hat{x}_p\} = F^k(\hat{x}_p) = 1-p \tag{7}$$

For $p > 1/n$, one can use the empirical distribution function to estimate $\hat{x}_p$. This leads to an unbiased estimator (see e.g. Mood, Graybill and Boes [1974]). However, for probabilities $p < 1/n$, the empirical distribution function makes no sense. In that case Dekker and De Haan [1989] give the following consistent estimator for $\hat{x}_p$, which is based on the tail behavior of the distribution:

$$\hat{x}_p = \frac{(kr/pn)^{\hat{\gamma}}-1}{1-2^{-\hat{\gamma}}}\left(X_{(n-r)}-X_{(n-2r)}\right)+X_{(n-r)} \tag{8}$$

where, as before, n is the number of observations and r=m/2, with m the number of the lowest order statistic used to compute $\hat{\gamma}$. For $p<1/n$, $\hat{x}_p$ is extrapolated outside the domain of the empirical distribution function.

Jansen and De Vries [1990] show that the tail parameters α for data decreasing in frequency (weekly, monthly etc.) are identical to the tail parameter α for daily return data. For the sum-stable distributions with $\alpha<2$ this result is imminent since the aggregated variables have the same kind of distribution. For the other leptokurtic alternatives with $\alpha>2$, like the Student-t distribution, the aggregates no longer share the same distribution. However, the tail behavior of these aggregates is still the same, see Feller [1971, Ch. VIII.8]. This is merely one illustration that, if one is only concerned with the tail distribution, there is additional robustness of the results due to their independence of the exact specification of the total distribution.

To link the tails of transaction return data with e.g. daily data we use another result of Feller [1971] and a model described in Harris [1987]. Assume that each day a series of events takes place, each of which generates information relevant for asset pricing. In succession to an event the values of assets change. However, it is not necessary for a transaction to take place each time an event occurs. Furthermore, assume that all of these changes in value are independently and identically distributed, that the number of events between two transactions is Poisson distributed[1], and that the number of transactions on each day is also Poisson distributed. Let S be the random variable describing the logarithms of the value changes. Then, both the returns between two subsequent transactions and the daily returns follow a compound Poisson process with underlying random variable S. It follows from Feller [1971, Ch. VIII.10 ex. 31] that this compound Poisson distribution has an identical tail index as the random variable S. Transaction returns must therefore have identical tail indices as daily, weekly or monthly returns.

To test whether this is indeed the case (for a particular data set) proceed as follows. Assume there are two series of i.i.d. random variables with, respectively tail parameters α_1 and α_2. Then the statistic Q with

$$Q = \left[\frac{\alpha_1}{\hat{\alpha}_1} - 1 \right]^2 m_1 + \left[\frac{\alpha_2}{\hat{\alpha}_2} - 1 \right]^2 m_2 \tag{9}$$

is $\chi^2(2)$ distributed with $\hat{\alpha}_i = 1/\hat{\gamma}_i$, and $\hat{\gamma}_i$, m_i as specified by (6) for the different samples. One should reject the hypothesis that $\alpha_1 = \alpha_2$ at the 5 percent significance level if there is no $\alpha = \alpha_1 = \alpha_2$ for which Q is below 5.99 (the value which is exceeded by a $\chi^2(2)$ with probability 0.05). If there is an α for which Q is below 5.99 one can not reject the hypothesis of equal tail parameters but one can derive an interval of α's for which Q is below 5.99. This is the confidence interval for the tail parameter that both samples have in common.

---

1 The same argument holds in case transactions do take place after each event.

3. ESTIMATED TAILS FOR HIGH FREQUENCY DATA

The estimation techniques of the previous section are now applied to prices of Bund futures traded at the London International Financial Futures Exchange (LIFFE). LIFFE's market audit trail (TAS, time and sales specification) files are used for this purpose. Day trading operates by open outcry (OOC) and lasts from 7.30 until 16.15 hours. After closing of this regular market, trading continues from 16.20 onwards until 18.00 hours by automated pit trading (APT). Time-stamped transaction data, for both OOC and APT, are selected for the nearby futures contract on the notional German Government Bond with a 6% coupon. Two successive weeks, November 11-15, and 17-22 in 1991, are covered in estimating tails and exceedance probabilities. The nearby delivery month of the futures contract is December. In addition to the continuously recorded prices, a 5 minute time-spaced subset has been constructed by selecting the price nearest to each 5 minute point. In order to obtain a sufficient number of observations, these series are extended by including the week immediately preceding and the week immediately following the mentioned two week period. All estimations are based on logarithmic transformations of first differenced prices.

Table I: Data Characteristics

sets	n	mean	variance	skewness	kurtosis
ALL(5m)	2420	$-5.550*10^{-7}$	$3.057*10^{-8}$	-0.191	6.141
ALL1	6314	$9.354*10^{-7}$	$7.648*10^{-9}$	-0.500	23.628
ALL2	5949	$6.932*10^{-7}$	$7.169*10^{-9}$	-0.041	6.806
OOC(5m)	2040	$-1.588*10^{-7}$	$3.275*10^{-8}$	-0.194	-6.099
OOC1	5683	$7.329*10^{-7}$	$7.969*10^{-9}$	-0.523	23.938
OOC2	5276	$9.232*10^{-9}$	$7.040*10^{-9}$	0.012	2.966
APT(5m)	360	$0.422*10^{-5}$	$0.183*10^{-7}$	-0.106	4.061
APT1	626	$2.765*10^{-6}$	$4.328*10^{-9}$	-0.074	2.255
APT2	668	$-6.186*10^{-7}$	$8.029*10^{-9}$	-0.316	29.975

ALL = Total data set,
OOC = Open outcry subset APT = Automated pit trading subset
1,2,(5m) = Respectively indicating week 1, week 2 and 5 minute intervals

Table I summarizes the standard characteristics: mean, variance, skewness and kurtosis, of the different data sets. The first column specifies the sets. Three groupings of data are distinghuised: open outcry data (OOC), automated pit trading data (APT) and the combination of both sets (ALL). For each group, parameters are estimated for 5 minute period data (5m) for the (extended) four week period, transaction data for the first week (1) and transaction data for the second week (2), of the middle two week period. This gives nine different data sets. The number of observations in each set is specified in the second column. For the 5 minute data the number of observations is six times as large for open outcry sessions as it is for automated pit sessions (the first period lasts six times as long as the second period). However, tick-to-tick transaction data indicate that trading is more active during open outcry sessions. From the kurtosis column we can observe that the assumption of normal kurtosis is strongly rejected, except for the second week of open outcry data. For this set normal kurtosis can not be rejected at a 5% significance level. However, for all other sets it is clear that there is excessive kurtosis and that they all have fat tails. Although some sets exhibit skewness, this is not as strong an aberration from normality as the kurtosis.

Table II gives results for the tail estimation procedure outlined in the previous section applied to the distinguished data samples. The second column specifies estimates of the tail parameters for the positive returns (the upper tail), while column five gives these parameters for negative returns (the lower tail). The third and six column give respectively maximum and minimum observations for each sample. The fourth and seventh column give the 95% confidence intervals for the estimated tail parameters. These confidence intervals are constructed symmetrically around $\hat{\alpha}=1/\hat{\gamma}$, assuming that $(\hat{\gamma}-\gamma)m^{1/2}$ is asymptotically normal. Most tail parameters are larger than 2 which indicates that Student-t distributions are better fit to describe these futures returns than sum-stable distributions. However, the latter distribution class can not be dismissed in all cases.

Table II: Tail Estimates

sets	α up	max.	range	α low	min.	range
ALL(5m) m=35 n=2420	2.726	10.4	1.82-3.62	5.979	-11.6	4.00-7.94
ALL1 m=125 n=6314	2.338	11.6	1.93-2.75	2.593	-13.9	2.14-3.05
ALL2 m=115 n=5949	2.674	8.1	2.18-3.16	2.190	-9.2	1.79-2.59
OOC(5m) m=30 n=2040	2.495	10.4	1.60-3.39	5.100	-11.6	3.28-6.94
OOC1 m=110 n=5683	2.247	11.6	1.83-2.67	2.535	-13.9	2.06-3.01
OOC2 m=105 n=5276	3.144	3.4	2.54-3.75	2.602	-3.4	2.11-3.10
APT(5m) m=5 n=360	2.143	4.7	0.66-3.62	12.075	-4.7	1.49-22.72
APT1 m=20 n=626	6.896	2.3	3.88-9.90	3.159	-4.6	1.78-4.55
APT2 m=20 n=668	1.987	8.1	1.12-2.86	1.731	-9.2	0.97-2.49

m= number of ordered observations used in estimating α.
n= number of observations.
maximum and minimum return columns are given in basis points.

As shown theoretically in the previous section, 5 minute data and continuously recorded data should have an identical tail parameter. The Q-test statistic can be used to assess whether this is indeed the case. Furthermore, it is interesting to see whether the different trading systems generate data with the same tail index.

Table III: Equality Tests

Subsets	upper tail range	lower tail range
ALLOUT(5m)-ALLIN(5m)	3.38-6.35	2.27-4.26
ALLOUT(5m)-ALL1	rejected	2.24-3.15
ALLOUT(5m)-ALL2	rejected	2.01-2.62
OOC(5m)-APT(5m)	1.43-3.43	3.38-7.22
OOC1-APT1	rejected	2.07-3.13
OOC2-APT2	2.47-3.07	2.01-2.67
ALL1-ALL2	2.12-2.83	2.05-2.68
OOC1-OOC2	2.45-2.63	2.14-2.99
APT1-APT2	rejected	1.49-2.63

ALLOUT(5m) = Two weeks surrounding ALL1+ALL2 at 5-minute interval
ALLIN(5m) = Two weeks overlapping ALL1+ALL2 at 5-minute interval

Table III gives results of the Q-tests on the equivalence between tail parameters for different combinations of periods and data sets. The first part, rows 2-4, tests whether it matters to observe on a time spaced basis or on a continuous basis. To avoid overlapping observations (potentially inducing a bias in the test) in this experiment, the 5 minute set is split into two parts: ALLIN containing overlapping observations with ALL1 and ALL2 continuously recorded prices, and ALLOUT containing the nonoverlapping returns (the weeks preceding ALL1 and following ALL2). The second part, rows 5-10, confronts separate weeks and trading systems.

Equality of the tail parameter is tested at the 5% significance level. In case equality is not rejected, the confidence interval for the common parameter is given in the table. Whereas theoretical equality between lower tails of continuous and five minute data can not be rejected, it is rejected for the upper tails. In the week 1 versus 2 comparison, rejections are obtained for the upper tail of automated pit trading. A final rejection occurs in comparing trading systems for week 1 (OOC1 versus APT1).

4. TAIL PROBABILITIES AND BID-ASK SPREADS

To evaluate the risk component in the bid-ask spread, one first has to introduce a measure for this spread. Since continuously recorded data series do not usually contain bid and ask quotes[2] (in addition to transaction prices), the majority of the literature investigated ways to estimate the spread from the available data. The most common estimator, using only transaction data, is the one developed by Roll (1984). Well known problems with this estimator, like positive auto-covariances are documented in e.g., Glosten (1987). However, Harris (1990) shows that the estimator tends to be relatively well behaved for high frequency (like 5 minutes) data. In table IV, the Roll estimator S for the percentage bid-ask spread based on continuously compounded returns R_t:

$$S = 200\sqrt{-Cov(R_t, R_{t-1})} \tag{10}$$

is given for the same samples we have used so far.

Table IV: Bid-Ask Spread Estimates*

Type	ALL (5m)	ALL1	ALL2	OOC (5m)	OOC1	OOC2	APT (5m)	APT1	APT2
Roll	0.98	0.96	0.93	1.06	0.97	0.93	0.68	0.72	1.08
GKN	1.41	1.36	1.46	1.32	1.25	1.22	1.85	2.03	2.29

* Roll's spread estimates are given in the first row. Adjusted spread estimates, using the GKN method are presented in the second row. All estimates are in basis points (1/100th percent).

In contrast with Roll we do not find any negative spreads. As in Roll's case, the estimates of the bid-ask spread depend on the observation interval. Longer periods (in this case the 5-minute data) provide higher estimates than shorter periods (tick-by-tick data), except for the APT-series. However, the differences are not as large as in Roll's daily versus weekly case.

In addition, we give estimates for an adjusted bid-ask spread measure taking

---

2 Even if (as in our data set) these quotes are included, they are usually not measured simultaneously. Furthermore, transactions are not designated to be at the bid or at the ask. Thus, a revised bid or ask quote may immediately be hit, without the new bid/ask ever appearing in the data records.

account of asymmetric information, as proposed by George, Kaul and Nimalendran (1991)[3] - GKN further on -, using transaction data and either bid or ask quotes. The autocovariance aspect is identical to the Roll approach:

$$S_{GKN} = 200 * \sqrt{-COV(\Delta X_{BT,t}, \Delta X_{BT,t+1})}$$
$$with \tag{11}$$
$$\Delta X_{BT,t} = (X_t - X_{t-1}) - (X_{B,t} - X_{B,t-1})$$

with bid (or ask) quote $X_{B,t}$ measured subsequent to each transaction price X_t. This incorporates the fact that quotes are revised prior to transactions based on expected price changes. Effectively, it means that market makers protect themselves by adjusting (widening or narrowing) bid/ask quotes with respect to an expected price trajectory. Such a price path would, by definition, be detected first by informed traders. Roll does not take this information component into account, and therefore underestimates the true spread. News releases indicate that the quoted spread at LIFFE was somewhere near 1.5 ticks (or 1.5 basis points). That figure seems much better approximated by our adjusted estimates in Table IV. It also illustrates the impact of the information component in the quoted spread (some 25%) which weighs heavily in the spread compensation, especially for the APT-samples where the gap between Roll and GKN is largest. This reflects the sometimes suggested anonymity problem with automated trading systems. Since market makers do not know whether there counterparty is informed or not, quoted bid-ask spreads will widen.

As noted before one can compare the bid-ask spread with the large price changes that make up the tail of the distribution, and are thus described by the tail parameter. Table V gives the estimates for price change levels as specified by equation (8) for all continuously quoted transaction data in the first and in the second week. Table IV shows that the bid-ask spread for the first week based on all transaction data (ALL1) is 1.36 basis points, while Table V indicates that the probability is 0.0005 that the price change is more than 14 basis points. This is ten times as high as the bid-ask spread.

---

3 George et al. operationalize Glosten's (1987) work.

Table V: Exceedance levels

ALL1 data α=2.338		ALL2 data α=2.764	
p	$\hat{X}_p$	p	$\hat{X}_p$
0.00005	41	0.00005	32.4
0.0001	29.9	0.0001	24.5
0.00013	26.5	0.00013	22.1
0.00015	24.8	0.00015	20.8
0.00017	23.4	0.00017	19.8
0.00019	22.2	0.00019	18.9
0.0002	21.7	0.0002	18.5
0.0003	17.9	0.0003	15.5
0.0004	15.6	0.0004	13.7
0.0005	14	0.0005	12.4

p = probability

$\hat{X}_p$ = exceedance level

Hence, if a market maker shorts this security, the probability that he will loose ten times the bid-ask spread when reversing his position in the next transaction is very small. The one twentieth percent probability indicates that it occurs on average once every 2000 transactions. If the market maker shorts, he mainly fears the upper tail of the distribution of price changes. If he is long, the lower tail is the more relevant part. The lower tail probabilities have been calculated as well and do not differ significantly from those for the upper tail. The riskiness of his position, as illustrated in Table V in absolute terms, is approximately identical whether he is short or long. It also follows from equation (8) that $\hat{x}_p$ is homogeneous in k and p. If he trades 100 transactions within a certain period, the probability that he will loose more than ten times the bid-ask spread (0.140%) is 100 times as large (i.e. 0.05, which is 5%). Of course he is then compensated with 50 times the bid-ask spread. One might conclude that, given these low probabilities, the bid-ask spread provides a rather generous compensation for the risk the market maker is exposed to. Even with a large number of transactions, this risk of at least one very large adverse price movement is well paid for. Hence, the bid-ask spread is merely used to compensate for small adverse changes, processing costs and a normal profit.

5. CONCLUSION

This paper focuses on the distributions of returns for financial assets and in particular on their tails. These areas with 'extreme' price changes may be most relevant for market makers that fear large (adverse) price changes in succession to trading with better informed traders. A novel procedure to investigate the tails of tick-by-tick transaction data is proposed. From estimates of the tail parameter it must be concluded that the appropriate distributions do not belong to the sum-stable distribution class but rather to the Student-t type class. In addition, the tail parameter allows inference on probabilities of large price changes even if no such changes are encountered in the sampling distributions. These changes cum probabilities are next compared to the bid-ask spread as observed in and estimated from transaction data. Given the small probabilities of very large price changes, the risk of strong adverse market movements is rather limited. This risk-component should therefore not be very influential on the bid-ask spread, except for the case where market makers are very risk adverse, which certainly is not the case. Assessing the relative importance of conventional risk measures (like the standard deviation) versus the tail measure can probably be achieved by means of a cross-sectional study.

References

Bagehot, W., 1971, The Only Game in Town, *Financial Analysts Journal*, 27, March/-April, 12-14.

Copeland, T.E. and Galai, D., 1983, "Information Effects on the Bid-ask Spread", *Journal of Finance*, 38, December, 1457-1469.

Dekkers, A.L.M., and L. de Haan, 1989, "On the Estimation of the Extreme-Value Index and Large Quantile Estimation", *Annals of Statistics*, Vol. 17, 1795-1832.

Fama, E.F., 1963 "Mandelbrot and the Stable Paretian Hypothesis", *Journal of Business*, Vol. 36, 420-429.

Feller, W., 1971, *An Introduction to Probability Theory and Its Applications*, Vol. 2, New York: Wiley.

George, T.J., G. Kaul and M. Nimalendran, 1991, "Estimation of the Bid-Ask Spread and its Components: A New Approach", *Review of Financial Studies*, 4, 623-656.

Glosten, L.R., 1987, "Components of the Bid-ask Spread and the Statistical Properties of Transaction Prices", *Journal of Finance*, 42, 1293-1308.

Glosten, L.R. and L.E. Harris, 1988, "Estimating the Components of the Bid/Ask Spread", *Journal of Financial Economics*, 21, 123-142.

Glosten, L.R. and P. Milgrom, 1985, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders", *Journal of Financial Economics*, 14, March, 21-42.

Haan, L. de, 1976, "Sample Extremes: An Elementary Introduction", *Statistica Neerlandica*, 161-72.

Hall, J.A., B. Wade Brorsen, and S.H. Irwin, 1989, "The Distribution of Futures Prices; a Test of the Stable Paretian and Mixture of Normals Hypotheses", *Journal of Financial and Quantitative Analysis*, Vol. 24, 105-116.

Harris, L.E., 1987, "Transaction Data Tests of the Mixture of Distribution Hypothesis", *Journal of Financial and Quantitative Analysis*, Vol. 22, 127-142.

Harris, L.E., 1990, "Statistical Properties of the Roll Serial Covariance Bid/Ask Spread Estimator", *Journal of Finance*, Vol. 45, 579-590.

Hill, B.M., 1975, "A Simple General Approach to Inference About the Tail of a Distribution", *Annals of Statistics*, Vol. 3, 1163-73.

Jansen, D., and C.G. de Vries, 1991, "On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspectives", *Review of Economics and Statistics*, Vol. 73, 18-24.

Kofman, P., and C.G. de Vries, 1989, "Potato Futures Returns: A Tail Investigation, *The Review of Futures Markets*, Vol. 8, No. 2, 244-258.

Mandelbrot, B., 1963, "The Variation of Certain Speculative Prices", *Journal of Business, Vol. 36, 394-419.*

Mood, A.M., T.A. Graybill, and D.C. Boes, 1974, *Introduction to the Theory of Statistics*, McGraw-Hill.

Roll, R., 1984, "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market", *The Journal of Finance*, Vol. 39, No. 4, 1127-1139.