EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

WORKING PAPER NO. 199

TIME VARIATION IN THE TAIL BEHAVIOUR OF BUND FUTURES RETURNS

BY THOMAS WERNER
AND CHRISTIAN UPPER

DECEMBER 2002

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

WORKING PAPER NO. 199

TIME VARIATION IN THE TAIL BEHAVIOUR OF BUND FUTURES RETURNS

BY THOMAS WERNER
AND CHRISTIAN UPPER

DECEMBER 2002

! "# $ " %
! " &
' ()) " &" * * + , & (()) " ")
& - .%%%%%%
/ ! & 0+ / 0# * " 102 3 14 2 05 0
"(617 37 7833 84) 47 0 (" 9& & " 0 " 9& & "

two indicators may well even send di erent signals at times. For instance, a decline in volatility might be o set by an increase in the fatness of the tail. For this reason, risk estimates based solely on volatility are to be viewed with extreme caution. Observing the tails of the distribution of returns thus provides information relevant to risk management which cannot be derived using conventional methods.

The results of this research project may be viewed as the “rst step towards modelling the fatness of tails. Future work could identify factors a ecting the fatness of the tails in order to gain a better understanding of its development over time and, where appropriate, to enable measures to be taken to in”uence it.

1 Introduction

It has been well known at least since the contribution of Mandelbrot (1963) that the distribution of asset returns is not well approximated by the Gaussian normal. Above all, the return distribution seems to have fatter tails than the normal distribution. The literature on the tail behaviour of returns has grown rapidly over recent years for a number of reasons. First of all, long series of asset prices and higher frequencies have become available. Together with the concurrent increase in computing power, this permitted the use of data intensive methods that previously had been di cult if not impossible to implement. Secondly, several new methods based on extreme value theory have been developed. An important example used in this paper is the bootstrap Hill estimator. Finally, the turbulences in the international "nancial markets in the summer and autumn of 1998 have questioned many of the assumptions of quantitative trading and risk management models, in particular the use of normal distributions.

The literature on the tail behaviour of asset prices has focussed mainly on the foreign exchange (for example Müller, Dacorogna, and Pictet (1998)), and stock markets (for example Lux (2001) for the spot and Cotter (2001) for the futures market). Only few papers deal with bonds or bond futures. Perhaps this is because bond returns are less volatile than stock or forex returns and are therefore believed to pose less risk than other assets. We believe that this argument is wrong and that the omission is not justi"ed. What matters is not the volatility of an asset price per se but the volatility of a position in that asset relative to capital. Even the safest asset becomes risky if leverage is su ciently high.

Our paper addresses this omission. Rather than working with data on the bond market directly, we estimate the tail behaviour of the BUND future. The BUND future contract traded on Eurex has become the main instrument for hedging long term interest rate risk in the Euro area. Trading is much heavier in the futures than in the spot market, and transactions data is available for a longer time span. Furthermore, since futures are traded

electronically on a centralized exchange, the data is also of higher quality than that on the underlying bonds, where trading is more fragmented. Nevertheless, we show in a companion paper (Upper and Werner (2002)) that prices in the futures and spot market move together very closely, so our main "ndings should apply to the bond market as well.

We focus on three questions: (i) Are heavy tails a relevant feature of the distribution of BUND futures returns? (ii) Is the tail behaviour constant over time? (iii) If it is not, then can we use the tail index as an indicator for "nancial market risk and does it add value in addition to classical indicators? The answers to questions (i) and (ii) have important implications for the design of trading and risk management models. If the return distribution has fat tails, then the assumption of normality in many of such models would lead one to seriously underestimate the likelihood of sharp falls and gains. For example, the Value at Risk (VaR) measure often used in risk management corresponds to the maximum loss that can occur with a given probability. In mathematical terms, it refers to a quantile, which depends crucially on the shape of the distribution. Another aspect is the liquidation risk of an asset. In a recent paper Du e and Ziegler (2001) analyse the riskiness of di erent liquidation strategies and show that the riskiness of di erent liquidation strategies depends signi"cantly on the fatness of the tails. But it is not only the fatness of tails alone that is important, also the variation of the tail behaviour over time is of interest. If the fatness of a tail is changing over time it is necessary to recognize this for risk evaluation and modelling.

We "nd that the distribution of high-frequency returns of the BUND future is indeed characterized by heavy tails. At "ve-minute intervals, the tail index estimated over our complete sample is around three. This implies that the distribution has in"nite kurtosis but "nite variance. The tail index increases as one reduces the frequency, although the tails remain signi"cantly heavy even for daily data.

The tail behaviour is not stable over time. What is interesting is that the tail index does not move in parallel with more standard measures for volatil-

ity such as the variance of returns. This suggests that the tail index does indeed provide information not contained in more commonly used volatility measures.

The paper is organized as follows. In section 2 we present the theoretical foundations for the tail index estimation discussed in section 3. This is followed by sections presenting the data and the empirical results. A "nal section concludes.

2 What are heavy tails?

Until now we have used a loose de"nition of heavy tails. It is not easy to de"ne heavy tails precisely. Even the name is not used uniformly. Sometimes heavy-tailed distributions are called fat-tailed, thick-tailed or long-tailed. In the following we use all of these terms as synonyms.

An often used de"nition of heavy-tailness is based on the 4th central moment. If X is a random variable and μ_X and $_X$ are the mean and the standard deviation of X, then X is called heavy-tailed if

$$E\left[\frac{(X - \mu_X)^4}{_X}\right] > 3.$$

This property is called excess kurtosis because the 4th central moment (the kurtosis) of the normal distribution is 3. However, this de"nition can only be applied in a sensible way if the 4th moment of a random variable actually exists. If two variables have in"nite 4th moments, then no discrimination between their distributions is possible on the basis of the kurtosis.

Unfortunately there is no general accepted de"nition of tail-heavyness under which a tail ranking is possible. We obtain such a ranking only for particular classes of distributions. In the following we brie"y discuss "ve classes[1].

E: nonexistence of exponential moments

[1] This classi"cation is borrowed from Bamberg and Dorßeitner (2001).

Figure 1: Di erent classes of heavy-tailed distributions

D: subexponential distributions

C: regular variation with tail index $>$ 0

B: Pareto tails with $>$ 0

A: stable (non-normal) distributions

These classes of distributions are nested as demonstrated in “gure 1. The broadest class E encompasses all distributions with

$$E\ e^{X} = .$$

It is important to note that the normal distribution is not contained in this class as its tail probability $P(X > x) = \overline{F}(x) = 1 - F(x)$ declines faster than exponentially.[2] In this sense all distributions of class E are heavy-tailed with respect to the normal distribution.

[2] $F(x)$ is the cumulative distribution function of X with $F(x) = P(X \quad x)$.

Since the normal distribution has comparatively thin tails, stronger assumptions are possible for heavy tails. The class D contains the subexponential distributions.[3] A distribution is subexponential if

$$\lim_{x} \frac{P(X_1 + \cdots + X_n) > x)}{P(\max(X_1, \ldots, X_n) > x))} = 1. \qquad (1)$$

This condition has a nice interpretation: the sum of n iid subexponential random variables is likely to be large if and only if their maximum is likely to be large. It is possible to show that equation 1 implies

$$\lim_{x} \frac{\overline{F}(x)}{e^{\check{S}\, x}} \qquad > 0.$$

As the name suggests, the tails of a subexponential distribution decrease more slowly than any exponential distribution.

In this paper we focus on the class C of distributions, which are characterized by regular variation in the tails. They form a subclass of the subexponential distributions[4] and satisfy the condition[5]

$$\lim_{t} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{\check{S}} \ . \qquad (2)$$

This condition states that far out in the tail (t) the distribution behaves like a Pareto distribution. As a consequence, the tail probabilities $P(X > x)$ decline according to a power function. The parameter is called Ž tail index Ž and can be used as a measure of tail-heavyness. An important member of class C is the Student-t distribution.

In contrast, distributions in class B have exact Pareto tails. The cumulative distribution function of the Pareto distribution is

$$F(x) = 1 \ \check{S}\ u\ x^{\check{S}} \qquad \text{where } x \quad u \text{ and } u > 0. \qquad (3)$$

[3] A short survey about subexponential distributions is given by Goldie and Klüppelberg (1998).

[4] This is shown, for example, by Goldie and Klüppelberg (1998) pp. 442-443.

[5] See for example Embrechts, Klüppelberg, and Mikosch (1997) p. 335.

The tail probability $P(X > x) = 1 - F(x) = \bar{F}$ of a class B distribution is therefore $u\, x^{-}$. The tail index can be related to the moments of a distribution with Pareto tails. From[6]

$$E[X^k] = u \int_u x^{k-\ -1} dx$$

it follows that only the first k-moments with $k <$ are bounded. This property is important to understand the class A, the class of stable or, as they are also called, -stable distributions. The distributions of this class have Pareto tails with $<$ 2, which implies infinite variance and, as a consequence, very fat tails. In spite of this restriction, class A is of great importance because asymptotic theory similar to central limit laws is possible[7]. Unfortunately, it is only possible to represent the stable distributions in a analytical way by the characteristic function (spectral representation). The density function can only be computed by numerical approximation[8]. In this paper we use a semi-parametric approach based on the Hill estimator that encompasses -stable distributions. However, if the variance of a distribution is known to be infinite, then a parametric approach based on -stable distributions would be more appropriate[9].

Let us consider class C in more detail. There is a nice connection between this class and classical extreme value theory. A main topic of extreme value theory is the modelling of the fluctuation of sample maxima. If $X_1, X_2, \ldots$ is a sequence of iid random variables then the sample maximum M_n is defined as

$$M_1 = X_1, \qquad M_n = \max(X_1, X_2, \ldots, X_n), \qquad n \quad 2.$$

One of the most famous theorems of extreme value theory is the Fisher-

[6]The density of the Pareto law is $u\, x^{-\ -1}$. See for example Dacorogna, Müller, Pictet, and de Vries (1998).

[7]Note that the central limit theorems are not applicable to stable distributions because the variance is infinite.

[8]See McCulloch (1998).

[9]For a recent application of -stable distributions to asset pricing models see Kim (2002)

Tippett theorem.[10] It states that if the properly normalised sample maximum converges to a non-degenerate distribution, then this distribution belongs to one of the following three distributions (density functions):

$$\text{Fréchet:} \quad (x) = \begin{cases} 0, & x \quad 0 \\ \exp(\check{S}x^{\check{S}}), & x > 0, \end{cases} \qquad > \quad 0$$

$$\text{Weibull:} \quad (x) = \begin{cases} \exp(\check{S}(\check{S}x^{\check{S}})), & x \quad 0 \\ 0, & x > 0, \end{cases} \qquad > \quad 0$$

$$\text{Gumbel:} \quad (x) = \exp(\check{S}e^{\check{S}x}), \quad x \quad R,$$

If the distribution of the sample maximum of a give distribution converges to one of the three, then it belongs to the maximum domain of attraction of (x), (x) or (x). It is possible to characterize distributions that belong to the maximum domain of attraction of the Gumbel distribution as having thin or moderately heavy tails. Distributions that belongs to the domain of attraction of the Weilbull distribution have a “xed upper end point. Important for our discussion of heavy tails is the class of distributions that belongs to the domain of attraction of the Fréchet distributions. A distribution belongs to this class if and only if its tails are regularly varying.[11] This exactly corresponds to our class C. The heaviness of the tails depends negatively on the tail index . We turn to its estimation in the next section.

3 Estimating and testing the tail index

3.1 The Hill estimator

The tail-index of equation 2 can be estimated with the Hill estimator. If the distribution under consideration is exactly Pareto, then the Hill estimator can

[10] See for example Embrechts, Klüppelberg, and Mikosch (1997) pp. 121-125.

[11] This is proven for example in Embrechts, Klüppelberg, and Mikosch (1997) pp. 131-132.

easily be constructed as a maximum likelihood estimation. The likelihood for the observation of k values $x_1, \ldots, x_k$ is

$$L(x_1, \ldots, x_k; \ ) = \prod_{i=1}^{k} u \ x_i^{\check{S} \ \check{S}1}$$

and the log-likelihood function

$$\log L = \sum_{i=1}^{k} (\log(\ ) + \ \log(u) \check{S} (\ + 1) \log(x_i)) \ . \qquad (4)$$

Maximizing equation 4 with respect to gives the Hill estimator

$$\frac{1}{} = \sum_{i=1}^{k} (\log(x_i) \check{S} \log(u)) \ .$$

The distributions of class C are not exactly Pareto but their tails behave like the tails of a Pareto distribution. Therefore the Hill estimator can be used for the outer parts of the distribution. Let x_i be the ith order statistic such that $x_i \quad x_{i\check{S}1}$ for all $i = 2, \ldots, n$. If we choose to include k observations from the right tail in our estimate of , the Hill estimator becomes

$$(k) = \frac{1}{(k)} = \sum_{i=1}^{k} (\log(x_{n\check{S}i+1}) \check{S} \log(x_{n\check{S}k})) \ . \qquad (5)$$

Whereas the concept and the calculation of the Hill estimator are straightforward, the choice of k is not. On the one hand, the approximation of the tails by the Pareto distribution improves as one moves further out into the tails. On the other hand, this leads to a reduction in the number of data points available, which drives up the variance. No general solution for this trade-o exists and many competing methods are available. An often used heuristical method is the Hill plot.[12] The Hill estimates are plotted for all possible values of k and an optimal k is selected by eye-ball search for a range that is robust with respect to k. Unfortunately, the Hill plot is sometimes erratic and may thus not be very useful.

[12] See for example Reiss and Thomas (2001) chapter 5.

Regression based estimator

A recently developed alternative to the Hill plot by Huisman, Koedijk, Kool, and Palm (2001) is especially useful for small samples. Their regression-based approach is based on an approximation of the asymptotic expected value of the Hill estimator as a linear function ofk

$$E(\ (k)) \quad \frac{1}{} \check{S}\ ck. \tag{6}$$

Here c is a constant depending on parameters of the distribution and the sample size. Ifk becomes small, the bias goes down and the expectation goes to the true value $= \frac{1}{}$. The variance of the estimator increases with small k

$$\text{var}(\ (k)) \quad \frac{1}{k^{\ 2}} \tag{7}$$

The idea of Huisman, Koedijk, Kool, and Palm (2001) is to use equation 6 in a regression analysis and regress the(k) values (computed with an ordinary Hill estimator) against k as follows:

$$(k) = \ _0 + \ _1 k + \ (k), \qquad k = 1, \ldots, \ .$$

The estimated $_0$ is an estimator of $= \frac{1}{}$. The authors propose to choose $= n/2$ wheren is the sample size[13]. Furthermore they propose a weighted least squares method to improve the e ciency of the estimator by use of equation 7 and show that the resulting estimator has good small sample properties.

We are especially interested in because it can be related directly to the existence of moments. One can show, by monte-carlo simulation, that the distribution of is asymmetric. To construct con“dence intervals for the point estimation of this parameter we use therefore a nonparametric percentile bootstrap proposed and tested in a similar context by Caers, Beirlant, and Vynckier (1998).

[13]The authors show that the results are robust with respect to the choice of .

Bootstrap based estimator

Another method to determine the optimal value of k has been developed by Danielsson, de Haan, Peng, and de Vries (1998). It is based on an evaluation of the mean squared error of de“ned as

$$MSE(k) = E\ (\ (k) \check{S}\ )^2\ . \qquad (8)$$

To evaluate this value the authors have proposed a bootstrap approach. The idea is to randomly draw with replacement from the original data set and to compute from this arti“cial sample. If this procedure is repeated for a large number of bootstrap samples the MSE can be calculated in principle. The optimal value of k is then found be minimizing equation 8. The problem is the value of in this equation. It is unknown and more problematic, and estimation is only possible if k is know in advance. Danielsson, de Haan, Peng, and de Vries (1998) have solved this problem by a method based on a combination of subsample bootstraps and the use of asymptotic theory. The details are summarized by Matthys and Beirlant (2000).[14]

3.2 Testing for structural breaks in the tail behaviour

Although it is well known that the volatility of asset returns varies over time, not much is known about the stability of the tail behaviour. To deal with this question, Quintos, Fan, and Phillips (2001) have recently developed a test for structural change in tail behaviour. If a priori a change from thinner to thicker tails is suspected then a recursive version of the Hill estimator (equation 5) is proposed by the authors. The test is based on

$$Y_T(t) = \frac{tm_t}{T}^{1/2} \quad \frac{t}{T} \check{S}\ 1\ . \qquad (9)$$

[14] This method should not be confused with the bootstrapping of conÞdence intervals. We use it only to asses the robustness of the tail index estimations and do not provide conÞdence intervals because bootstrapping a bootstrap method is computational infeasible in our case.

The test is performed by the computation of equation 9 for a recursively increasing sample size t. In equation 9, $_T$ is the Hill estimator for the whole sample size T and $_t$ is the same estimator for the sample up to time t. m_t is equal to k in the Hill estimator. In addition, Quintos, Fan, and Phillips (2001) have proposed a modi"cation to this test to deal with GARCH type dependency. It is known[15] that the Hill estimator is a consistent estimator for a large class of dependent processes but the variance is e ected by dependency. This problem is solved by the authors by a variance correction. In this paper we use the more general version allowing for GARCH type dependency.

4 Data

Our analysis is based on tick data for the Bund future obtained from Deutsche Börse AG. It covers all transactions from January 1997 to December 2001, a total of 12.7 million trades. For each transaction, we have a time stamp (up to the centisecond), volume and price as well as the expiry date of the contract. Our data does not, however, include quotes.[16]

Contracts expire in March, June, September and December of each year. Trading is concentrated on the nearby maturity and switches to the next contract within days just before expiry. Since contracts with di erent expiry dates tend to di er in price, we look only at the most actively traded maturity on each trading day. Since the changeover occurs very rapidly, we loose only about 5 % of the observations. We then link the individual contracts to a long series covering the whole sample period.

Trades occur at irregular intervals whereas the statistical methods used in this paper require equally spaced data. We create such a series by recording

[15] See for example Resnick and Stùcà (1996).

[16] Until the December 1998 contract, each Bund future refered to a notional German government bond with a face value of 250,000 DM and a coupon of 6 %. The Euro Bund future, which replaced the Bund future in the transition to EMU, has a contract value of 100,000 Euro.

Year	Method	5 minute		1 hour		1 day	
		Left-tail	Right-tail	Left-tail	Right-tail	Left-tail	Right-tail
1997-2001	Regression	3.01	3.33	3.36	3.90	4.55	6.28
		[2.91, 3.09]	[3.20, 3.46]	[3.06, 3.73]	[3.52, 4.35]	[3.35, 6.98]	[4.61, 10.16]
	Bootstrap	3.16	3.44				
1997	Regression	2.46	2.88	2.91	3.40		
		[2.34, 2.61]	[2.72, 3.09]	[2.43, 3.58]	[2.74, 4.34]		
	Bootstrap	3.16	3.75				
1998	Regression	2.46	2.88	3.18	3.77		
		[2.33, 2.62]	[2.70, 3.09]	[2.66, 4.12]	[3.18, 5.02]		
	Bootstrap	2.66	2.77				
1999	Regression	3.75	4.41	3.51	3.91		
		[3.49, 4.04]	[4.06, 4.81]	[2.94, 4.55]	[3.23, 5.07]		
	Bootstrap	3.46	3.86				
2000	Regression	3.22	3.19	4.09	4.59		
		[2.98, 3.46]	[2.96, 3.45]	[3.29, 5.37]	[3.67, 6.15]		
	Bootstrap	2.72	3.12				
2001	Regression	2.78	2.91	3.38	4.37		
		[2.60, 2.97]	[2.74, 3.11]	[2.79, 4.57]	[3.62, 5.87]		
	Bootstrap	2.66	3.17				

Table 1: The tail index of Bund futures returns

price of the last transaction in each time bracket. Intervals where no transactions take place are treated as missing data. This means that we do not “ll in with the last available price, as is often done in the literature. Overnight returns are discarded for frequencies higher than a day. We also construct a series with daily data. In this case, we take the last trade on or before 5 p.m., when trading activity is at its peak.

5 Empirical results

5.1 Tail estimates

The results of the tail index estimations are summarized in table 1. Remember from section 2 that declines as the tails of the distribution become thicker and that a tail index of less than 4 implies in“nite forth moments and therefore an in“nitely high kurtosis. Let us “rst discuss the results for the complete sample (1997-2001) shown at the top of table 1. For the “ve-

minute returns, both methods estimate α to be greater than 3 but smaller than 4. The values in square brackets are the upper and lower bounds of 95% bootstrap con“dence intervals for the regression-based tail index[17]. High-frequency returns thus appear to have fat tails with in“nite kurtosis, possibly in“nite third moments but de“nitely “nite variance. At lower frequencies, the tail index increases, and the tails become thinner[18]. For daily returns, the estimates for α obtained by the regression-based method are above 4, suggesting a bounded fourth moment. But it is important to note that even in this case the tails remain heavier than those of the normal distribution, which has a very large, in theory in“nite, tail index. Therefore our results suggest that the “rst of our questions can be answered in the a rmative: Yes, the returns on the Bund future do have thick tails. Moreover, the left tails, corresponding to negative returns, tend to be slightly thicker than the right tails irrespective of the frequency.

Let us now turn to the estimates for the di erent years of our sample. Unfortunately, only results for frequencies higher than a day are available as there are not su cient data points at the daily level. We “nd that the tails seem to be particularly fat during 1998 and 2001, and less so in 1999 and 2000. This is not surprising, given that 1998 saw some of the worst turbulence in the international “nancial markets in living memory, and 2001 was marred by the September 11 shock. In contrast, 1999 and 2000 were rather tranquil years. The con“dence intervals of the 5 minute losses (left-tail) are non-overlapping for 1998 to 2001. This “nding is a “rst sign of time variation in the tail behaviour, which we shall explore in the following subsection.

[17] Computing conÞdence intervals for the bootstrap method would involve bootstrapping the bootstrap. In our case, it has turned out to be unfeasible

[18] We use the bootstrap method only for the 5 minute returns because it is known from the literature that it requires a very large sample size (> 5000), see Matthys and Beirlant (2000).