

Haas, Markus; Mittnik, Stefan; Paoella, Marc S.

**Working Paper**

## Modeling and predicting market risk with Laplace-Gaussian mixture distributions

CFS Working Paper, No. 2005/11

**Provided in Cooperation with:**

Center for Financial Studies (CFS), Goethe University Frankfurt

*Suggested Citation:* Haas, Markus; Mittnik, Stefan; Paoella, Marc S. (2005) : Modeling and predicting market risk with Laplace-Gaussian mixture distributions, CFS Working Paper, No. 2005/11, Goethe University Frankfurt, Center for Financial Studies (CFS), Frankfurt a. M., <https://nbn-resolving.de/urn:nbn:de:hebis:30-10872>

This Version is available at:

<https://hdl.handle.net/10419/25453>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



---

No. 2005/11

**Modeling and Predicting Market Risk  
With Laplace-Gaussian Mixture Distributions**

Markus Haas, Stefan Mittnik,  
and Marc S. Paoella

---





## Center for Financial Studies

The *Center for Financial Studies* is a nonprofit research organization, supported by an association of more than 120 banks, insurance companies, industrial corporations and public institutions. Established in 1968 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia.

The CFS Working Paper Series presents the result of scientific research on selected topics in the field of money, banking and finance. The authors were either participants in the Center's Research Fellow Program or members of one of the Center's Research Projects.

If you would like to know more about the *Center for Financial Studies*, please let us know of your interest.

Prof. Dr. Jan Pieter Krahen

Prof. Volker Wieland, Ph.D.



CFS Working Paper No. 2005/11

## **Modeling and Predicting Market Risk With Laplace-Gaussian Mixture Distributions\***

Markus Haas<sup>1</sup>, Stefan Mittnik<sup>2</sup>,  
and Marc S. Paolella<sup>3</sup>

March 2005

### **Abstract:**

While much of classical statistical analysis is based on Gaussian distributional assumptions, statistical modeling with the Laplace distribution has gained importance in many applied fields. This phenomenon is rooted in the fact that, like the Gaussian, the Laplace distribution has many attractive properties. This paper investigates two methods of combining them and their use in modeling and predicting financial risk. Based on 25 daily stock return series, the empirical results indicate that the new models offer a plausible description of the data. They are also shown to be competitive with, or superior to, use of the hyperbolic distribution, which has gained some popularity in asset-return modeling and, in fact, also nests the Gaussian and Laplace.

**JEL Classification:** C16, C50

**Keywords:** GARCH, hyperbolic distribution, kurtosis, Laplace distribution, mixture distributions, stock market returns.

\* The research of Haas and Mittnik was supported by the Deutsche Forschungsgemeinschaft. Mittnik conducted part of the research while visiting the Department of Economics at Washington University in St. Louis with a grant from the Fulbright Commission. Part of the research of M. Paolella has been carried out within the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), which is a research program supported by the Swiss National Science Foundation. The authors are grateful for the constructive comments from an anonymous referee.

<sup>1</sup> Institute of Statistics, University of Munich

<sup>2</sup> Institute of Statistics, University of Munich; Center for Financial Studies; Ifo Institute for Economic Research

<sup>3</sup> Corresponding Author;  
Swiss Banking Institute, University of Zurich, email: paolella@isb.unizh.ch

# Non-technical Summary

In this paper, we analyze the distributional properties of daily returns on 25 stocks belonging to the German stock index DAX. It is widely known that the normal distribution, although it has dominated the field of empirical finance for a long time, cannot serve as an adequate model for this task. This has stimulated considerable efforts to develop more appropriate models. In this context, there is at least as much interest in the conditional as in the unconditional return distribution. The substantial interest in the conditional distribution is a result of the observation that the volatility of asset returns, which can be viewed as a measure of risk, appears to be serially correlated and can, to some degree, be predicted from past observations. Thus, the methods of volatility modeling are indispensable tools for the risk management of banks and other institutions facing short- or medium-term financial risks. Among the models that have been proposed to capture time-varying volatility, the GARCH process is certainly the most popular, as it is not only rather simple but also often quite successful in filtering out the heteroskedasticity from financial return series. Ignoring the highly pronounced volatility dynamics in higher-frequency returns data can give rise to misleading conclusions concerning the underlying distribution governing the return process, because some of its important features are jointly determined by the shape of the conditional distribution and the dynamic properties of its second moment (or some other measure of risk). This draws into question the Laplace-hypothesis that has recently attracted some interest, as is reviewed in the present paper. Developing more general models, we show that the evidence in favor of the Laplace distribution is probably due to not considering the GARCH-type dynamics in the returns. Using the newly developed models, we illustrate in detail the impact of GARCH effects on estimates of important parameters of the return distribution. The specifications we propose are essentially mixtures of the Laplace and the normal distribution, which, as we argue in the paper, makes them attractive models for a number of reasons, as does their good performance in a competition with some mainstream models of financial modelling, the results of which are also reported in the paper. As a consequence, the processes will be further investigated in the future, possibly with a special focus on the pricing of options.

# Nichttechnische Zusammenfassung

Die vorliegende Studie untersucht in systematischer Weise die Verteilungseigenschaften der Aktienkursveränderungen der im DAX 30 enthaltenen Unternehmen für den Zeitraum von 1996 bis 2001. Seit längerer Zeit ist weithin bekannt, dass die Normalverteilung—trotz ihrer langjährigen Dominanz in der empirischen Finanzökonomie—nicht als geeignetes Modell für Renditen im Mittel- und Hochfrequenzbereich in Frage kommt. Noch nicht abgeebbt ist jedoch die Diskussion darüber, wie diese Zeitreihen adäquat beschrieben werden können. Dabei richtet sich das Interesse nicht nur auf die unbedingten Verteilungseigenschaften der Renditen, sondern mindestens ebenso sehr auf die Modellierung der bedingten Verteilung und ihrer Dynamik. Dieser Tatbestand erklärt sich aus der Beobachtung, dass sich an den Finanzmärkten Perioden geringer und solche mit hoher Volatilität abzuwechseln scheinen, so dass sich etwa für das Risikomanagement mit kurz- oder mittelfristigem Horizont die Modellierung der Volatilitätsdynamik mit dem Ziel der Prognose als Anliegen von großer Wichtigkeit darstellt. Das GARCH-Modell mit seinen zahlreichen Modifikationen und Erweiterungen hat sich zu diesem Zweck in Wissenschaft und Praxis mittlerweile fest etabliert. Dabei lässt sich in der empirischen Analyse die Frage, welche Form die bedingte Renditeverteilung hat, nicht von der Frage nach der Renditedynamik trennen, denn beide determinieren gemeinsam die unbedingte Renditeverteilung. Hier setzt die vorliegende Arbeit an. Anhand geeigneter Modelle, die hier entwickelt werden, wird untersucht, in welcher Weise sich die Verteilungseigenschaften beim Übergang von einem statischen zu einem dynamischen Modell verändern. Für die Modellierung der Dynamik vertrauen wir dabei auf den bekannten GARCH-Ansatz, den wir mit speziellen Verteilungsstrukturen kombinieren. Die Analyse zeigt klar, dass bei der Frage nach dem geeigneten Verteilungsmodell für Wertpapierrenditen die Frage der Dynamik nicht außer Acht gelassen werden darf. Im Anschluss an diese Untersuchung werden die entwickelten Modelle einer Reihe von statistischen Tests unterworfen und mit anderen, bereits weit verbreiteten Mainstream-Modellen verglichen, wobei die neuen Modelle stets gut abschneiden. Diese Ergebnisse legen den Schluss nahe, dass es lohnend sein wird, ihr weiteres Studium zu betreiben, etwa im Rahmen von Modellen zur Optionspreisbewertung.

# 1 Introduction

Overwhelming empirical evidence strongly suggests that the normal, or Gaussian, distribution is not appropriate for modeling financial time series such as returns on stocks, foreign currency, and other financial instruments. This finding applies to both the unconditional and conditional returns, where conditional modeling is most often associated with modeling the time-varying volatility of the returns by GARCH-type or stochastic-volatility models. Relative to the normal, the empirical distribution of the returns themselves, or the residuals of conditional (GARCH) models is typically quite fat-tailed and more peaked around the center. As a consequence, alternative distributions possessing these characteristics have been proposed in the literature (see, for example, McDonald, 1996).

One strategy relies on mixtures of distributions, as developed in the seminal papers of Clark (1973) and Epps and Epps (1976), and reviewed in Liesenfeld (2001). As an important case in point, modeling the *unconditional* returns with a mixture of normals allows the *conditional* distribution of the returns to be normal, which is often considered attractive because of the implications of the central limit theorem. In this setting, the variance,  $\tau^2$ , of the conditional normal distribution is the realization of a random variable that may be related to the arrival of relevant new information, which is not distributed uniformly over time. That is, if we specify a density function, say  $h(\tau^2)$ , for the variance, the joint distribution,  $f(r, \tau^2)$ , of returns  $r$  and  $\tau^2$  is  $f(r, \tau^2) = h(\tau^2)f(r|\tau^2) = h(\tau^2)\phi(r; \mu, \tau^2)$ , where  $\phi(r; \mu, \tau^2)$  denotes the normal density function with mean  $\mu$  and variance  $\tau^2$ . The unconditional distribution of returns is found by integrating out  $\tau^2$ , that is,

$$f(r) = \int \phi(r; \mu, \tau^2)h(\tau^2)d\tau^2. \quad (1)$$

Hence, returns over fixed intervals of time follow a mixture distribution with *mixing distribution*  $h(\cdot)$ . The arguably most popular example is the Student's  $t$  distribution, which can be expressed as an inverted gamma-mixture of the normal distribution.

The Laplace distribution for returns, with

$$f_X(x) = \frac{1}{2} \exp\{-|x|\}, \quad (2)$$

arises when the mixing distribution in (1) is exponential, i.e.,  $h(\tau^2) = \lambda \exp(-\lambda\tau^2)$ ,  $\lambda > 0$ . The result is well-known; see, for example, Teichrow (1957), Andrews and Mallows (1974); while Linden (2001) also provides a proof. The Laplace distribution is particularly appealing not only because of its emergence from the mixture framework, but also because of other favorable

features, for example, its stability properties (Mittnik and Rachev, 1993; and Kotz, Podgorski and Kozubowski, 2001). The Laplace distribution arises from the geometric summation process, a probabilistic scheme which has some resemblance with Clark's (1973) subordinated process. The role of the Laplace distribution in the family of geometric-stable distributions is analogous to that of the normal distribution among the class of (non-geometric) stable distributions. The process is attractive, because geometric-stable random variables are closed under geometric summation and have domains of attraction, providing certain robustness to model misspecification (for details, see Mittnik and Rachev, 1993).

In conjunction with several conditional GARCH model formulations, the Laplace distributional assumption was shown to be greatly superior to the normal assumption by Granger and Ding (1995) for the S&P 500 index, and by Mittnik, Paoletta and Rachev (1998) for the Nikkei index. It is important to note that use of the Laplace distribution instead of the normal involves the same number of model parameters.

Moreover, the Laplace, as either the unconditional distribution of asset returns, or as the conditional distribution coupled with a GARCH-type structure, has been found to result from a purely *economic* model of stock prices (Reimann, 2005). Simulation results therein show consistent evidence for fat-tailed return distributions, with favor for either hyperbolic returns or, under certain economic model conditions, the Laplace, which is a special case of the hyperbolic (see Section 5 below).

Linden (2001) examines the unconditional distribution of daily returns on Finnish stocks and finds that the Laplace distribution cannot be rejected for roughly half of the stocks. While still informative, conditional models of asset returns are often of greater value than unconditional models when their purpose is risk and volatility prediction. Ignoring the highly pronounced volatility clusters in daily or higher-frequency returns data can give rise to misleading conclusions concerning the underlying distribution governing the return process, because important features, such as excess kurtosis of the unconditional distribution, are jointly determined by the shape of the conditional distribution *and* the dynamic properties of its second moment (or some other measure of volatility).

In this paper, we analyze the daily returns on 25 stocks belonging to the German stock index DAX and show that the Laplace distribution (while still superior to the normal) is still insufficient when GARCH effects are appropriately accounted for. We extend the Laplace model by considering two models which combine both the Laplace and normal distributions. The first proposes a random variable whose probability density function (pdf) is a discrete



mixture of a Gaussian and Laplace pdf. This construction is not new and has its origin in the engineering sciences. The second model involves a random variable defined to be the weighted sum of two independent random variables, one being normal, the other Laplace. Its use appears to be new. Both models provide adequate descriptions of the empirical return distributions and outperform, both in- and out-of-sample, the Laplace and the generalized exponential distribution, or GED, which is another candidate that nests the normal and the Laplace.

The paper is organized as follows. Section 2 tests the Laplace distribution against the GED both in the unconditional and conditional context, resulting in a blatant rejection of the Laplace distribution. Section 3 then introduces the aforementioned normal–Laplace combinations and details their important properties. Section 4 presents empirical estimates for the German stock data, and evaluates their in- and out-of-sample performance. Section 5 augments the comparison with the hyperbolic distribution, which also nests the normal and Laplace. Section 6 summarizes the findings and provides remarks on ways of generalizing the normal–Laplace mixture to support asymmetry.

## 2 Testing the Laplace Model

### 2.1 Unconditional and Conditional Models for Asset Returns

As mentioned above, one of the most striking deviations from normality characterizing the empirical distribution of asset returns is fat-tailedness, or excess kurtosis, relative to the normal. The kurtosis,  $\kappa$ , of a distribution of a zero-mean random variable,  $X$ , (with finite fourth moment) is defined via  $\kappa = E(X^4)/E^2(X^2)$ , where  $E$  denotes the expectation operator. If  $\kappa$  is larger than 3 (the value associated with the normal distribution), the distribution of  $X$  is called leptokurtic, or fat-tailed. A large number of distributions capable of modeling leptokurtosis exist—the Laplace distribution (2) is among them. A detailed account of the Laplace distribution, including applications in finance and economics, is provided by Kotz *et al.* (2001). The raw moments of a Laplace random variable are  $E(X^m) = 0$  for  $m$  odd and  $E(X^m) = m!$  for  $m$  even, so that the coefficient of kurtosis is  $\kappa^{\text{Lap}} = 4!/(2!)^2 = 6$ .

Although it is heavier tailed than the normal, the fact that it lacks an explicit shape parameter (such as the degrees of freedom parameter of the Student’s  $t$  distribution) renders the Laplace relatively inflexible. Its shape is fixed, with the kurtosis being restricted to the constant value of six, irrespective of the kurtosis of the data being modeled. The severeness

of this restriction is illustrated when computing the sample kurtosis,  $\hat{\kappa}$ ,

$$\hat{\kappa} = \frac{\hat{m}_4}{\hat{m}_2^2}, \quad \text{with} \quad \hat{m}_k = T^{-1} \sum_{t=1}^T (r_t - \bar{r})^k, \quad k = 2, 4, \quad \bar{r} = T^{-1} \sum_{t=1}^T r_t, \quad (3)$$

for the 25 daily German stock returns studied in this paper. The values range from 4.13 to 10.45, indicating a considerable variation in the kurtosis across the different assets.

A well-known distribution which nests the Laplace (and normal) via introduction of a shape parameter is the generalized exponential (GED), with density

$$f_X(x; p) = \frac{2^{-(1/p+1)} p}{\Gamma(p-1)} \exp \left\{ -\frac{1}{2} |x|^p \right\}, \quad p > 0, \quad (4)$$

where  $p$  is the shape parameter. Straightforward calculation shows that the kurtosis of (4) is a function of  $p$ , given by

$$\kappa^{\text{GED}}(p) = \frac{\text{E}(X^4)}{\text{E}^2(X^2)} = \frac{\Gamma\left(\frac{5}{p}\right) \Gamma\left(\frac{1}{p}\right)}{\Gamma^2\left(\frac{3}{p}\right)}, \quad (5)$$

where  $\Gamma(\cdot)$  denotes the gamma function,  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ . For  $p = 1$ , (4) reduces to the Laplace distribution, while for  $p = 2$ , the normal is obtained, so that  $\kappa^{\text{GED}}(1) = 6$  and  $\kappa^{\text{GED}}(2) = 3$ . As detailed in the Appendix, the kurtosis is strictly decreasing as  $p$  increases, so that the GED is “somewhere between” the normal and the Laplace for  $1 < p < 2$ .

*A priori*, there is no reason why all stock returns should exhibit the same distributional shape. As such, one would expect that, provided enough data, statistical tests, such as the likelihood ratio (LRT), will favor the GED over the Laplace distribution for most stocks, if not all, as the null hypothesis of  $p = 1$  is a point (i.e., measure zero) hypothesis. Indeed, for the 21 daily time series studied in Linden (2001), the GED was favored over the Laplace in 11 cases, based on LR tests with significance level 0.05. However, Linden (2001) fits the *unconditional* distribution of the returns; that is, he assumes that returns, denoted by  $r_t$ , are adequately modeled as

$$r_t = \mu + \epsilon_t, \quad (6)$$

where  $\mu$  is the (constant) mean of  $r_t$  and  $\{\epsilon_t\}$  is an iid sequence of zero-mean random variables following a Laplace distribution with constant scale-parameter  $\sigma$ . It will be useful in subsequent analysis to write  $\epsilon_t = \eta_t \sigma$ , where  $\{\eta_t\}$  is an iid sequence of standard Laplace random variables with density (2), and  $\sigma$  is a positive constant.

In many, if not most, practical applications involving moderate- to high-frequency financial return data, the *conditional* distribution of  $r_{t+h}$  given the information up to and including

time  $t$ , collected in the information set  $\Psi_t$ , is of most interest. The typical pattern in daily returns is that the autocorrelation in first moments is nearly negligible, while the correlations in second moments are highly significant. The latter phenomenon, known as conditional heteroskedasticity, can often be adequately captured by specifying a GARCH model (see, for example, the survey article of Palm, 1997; and Gouriéroux, 1997). Ignoring first-order dynamics, the basic GARCH model assumes that returns are given by

$$r_t = \mu + \eta_t \sigma_t, \quad (7)$$

where, *vis-à-vis* (6),  $\sigma_t$  is a time-varying scale parameter which evolves according to the GARCH(1,1) recursion

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad \alpha_0 > 0, \alpha_1, \beta_1 \geq 0. \quad (8)$$

The conditional variance of  $r_t$ , given  $\Psi_{t-1} = \{r_{t-1}, r_{t-2}, \dots\}$ , is thus  $\sigma_t^2 E(\eta_t^2)$ . For instance, in the normal case,  $\eta_t \sim N(0, 1)$ , so that (8) represents the conditional variance of  $r_t$ . In the Laplace case,  $E(\eta_t^2) = 2$ , and the conditional variance of  $r_t$  is  $2\sigma_t^2$ . When  $\eta_t$  follows the GED distribution—which Nelson (1991) introduced to the GARCH literature—the variance is  $2^{2/p}[\Gamma(3/p)/\Gamma(1/p)]\sigma_t^2$ .

If GARCH effects are present in return data, then their consideration is crucial for distributional modeling, because important features of the unconditional distribution (for example, unconditional kurtosis), are induced both by volatility clustering *and* conditional non-normality of the return distribution (see, for example, the contributions in Knight and Satchell, 2001; and Rachev, 2003). In fact, the exact relation between these two sources of unconditional kurtosis has only recently been established by Bai, Russell and Tiao (2003), while the fact that (G)ARCH effects account for excess kurtosis even with conditional normality has been known since the pioneering work of Engle (1982).

In view of this relation, tests of the Laplace distribution against the GED based on fitting unconditional distributions are thus expected to produce misleading results, because the kurtosis due to non-normality of the conditional distribution is overestimated when ignoring GARCH volatility dynamics.

We perform the test using daily (closing) returns on 25 German stocks (see Table 1 for details), which are included in the German DAX index, a blue chip index comprising the 30 largest firms in terms of exchange turnover and market capitalization. As the composition of the index changes over time, we use the 25 corporations that were included in the index over the

whole sample period (see Theissen, 2003).<sup>1</sup> The data range from December 1996 to October 2001, yielding 1,220 observations for each stock. Continuously compounded percentage returns are used, i.e.,  $r_t = 100(\log P_t - \log P_{t-1})$ , where  $P_t$  is the price at time  $t$ . Table 1 reports statistical properties of the 25 return series, along with Lagrange multiplier test results for conditional heteroskedasticity, which reveal highly significant ARCH effects of order 5 for all return series but one (Bayer). However, the ARCH effects of order 6 are significant at the 10% level for Bayer.

## 2.2 Empirical Results for the Laplace Model

As the GED density (4) nests the Laplace when  $p = 1$ , the LRT can be used to test for the appropriateness of the Laplace distribution. We do so for the unconditional model (6) as well as the conditional model given in (7) and (8), adopting both the Laplace and GED assumption for the disturbance term  $\epsilon_t$ , respectively  $\eta_t$ . The exact maximum likelihood estimator (MLE) is used for the unconditional models and the conditional MLE for the GARCH-based models (see Mittnik, Paoletta and Rachev, 1998, for details on the latter).

Estimation results are reported in Table 2 and include the log-likelihood values,  $L^{\text{Lap}}$  and  $L^{\text{GED}}$ , of models assuming Laplace and GED distributed innovations, the likelihood ratio test statistic  $LR = -2(L^{\text{Lap}} - L^{\text{GED}})$ , and the estimates of the shape parameter,  $p$ , of the GED distribution. Under the null hypotheses of Laplace, the  $LR$  is approximately  $\chi^2(1)$ -distributed, implying critical values of 2.706, 3.842, and 6.635, for the 10, 5 and 1 percent significance level, respectively.

The results are twofold. First, for all 25 return series, the Laplace distribution is rejected at the 5% level in favor of the GED, both in the unconditional and conditional models, and in most (21 out of 25) cases even at the 1% level. Thus, the data at hand are even less favorable for the Laplace hypothesis than for the Finnish stocks examined in Linden (2001), where the Laplace is not rejected for 10 out of 21 stocks. A second result is that, if we discard the assumption of iid returns and allow for GARCH dynamics in the scale parameter as in (8), then the evidence against the Laplace distribution becomes even stronger. For *all* stocks, the  $LR$ -statistics increase substantially. In the GARCH framework, the LRT rejects the Laplace distribution for all series at the 1% level. At the same time, a comparison to the unconditional log-likelihood values,  $L^{\text{Lap}}$  and  $L^{\text{GED}}$ , in the left part of Table 2, indicates

---

<sup>1</sup> This has, for example, the consequence that DaimlerChrysler is not included because of the merger of Chrysler and Daimler Benz in 1998.

Table 1: Stock Names and Basic Statistical Properties of their Returns<sup>a</sup>

Stock ID#	Name	Mean	Std. Dev.	Min.	Max.	Kurt.	ARCH(1)	ARCH(5)
1	Adidas-Salomon	0.003	2.570	-11.91	12.78	5.34	71.62***	86.29***
2	Allianz	0.054	2.357	-15.13	13.40	8.16	11.33***	46.90***
3	BASF	0.029	2.028	-7.57	13.70	5.43	15.91***	64.91***
4	Bayer	0.019	2.039	-19.40	9.89	10.45	0.15	4.88
5	BMW	0.067	2.680	-12.09	12.86	4.74	27.07***	89.53***
6	Commerzbank	-0.004	2.179	-12.77	10.21	6.88	31.30***	97.94***
7	Dt. Bank	0.049	2.470	-14.09	13.80	6.86	31.92***	68.33***
8	Degussa	-0.008	2.371	-8.49	14.35	4.66	37.28***	76.18***
9	Dt. Telekom	0.014	2.951	-12.46	14.50	4.27	4.53**	52.82***
10	E.ON	0.033	2.105	-8.57	10.45	4.45	7.73***	34.85***
11	Fresenius	0.018	2.359	-11.33	14.38	5.90	13.31***	71.81***
12	Henkel	0.052	2.303	-8.16	12.01	4.80	17.52***	59.49***
13	Hypo-Vereinsbank	0.013	2.628	-11.02	13.93	5.68	59.50***	150.72***
14	Lufthansa	0.017	2.537	-16.37	12.20	6.68	4.61**	138.87***
15	Linde	0.010	2.206	-14.32	11.05	5.67	3.75*	15.31***
16	MAN	0.016	2.344	-9.51	11.13	4.40	29.38***	48.76***
17	Metro	0.045	2.349	-12.10	8.37	4.23	26.90***	46.69***
18	Münchener Rück	0.100	2.511	-17.05	15.65	6.61	2.48	24.07***
19	Preussag	0.046	2.304	-18.42	13.45	8.64	59.15***	97.43***
20	RWE	0.036	2.154	-7.70	11.68	5.69	64.78***	70.93***
21	SAP	0.101	3.707	-19.39	21.66	6.41	12.07***	53.67***
22	Schering	0.089	1.998	-7.88	7.31	4.13	8.70***	30.60***
23	Siemens	0.064	2.653	-12.48	16.19	5.70	1.93	36.54***
24	ThyssenKrupp	0.002	2.396	-17.54	11.65	7.19	3.36*	23.20***
25	Volkswagen	0.029	2.477	-14.28	10.68	5.72	23.38***	85.60***

<sup>a</sup>Listed are the names of the stocks used for the investigation. Statistical properties of the returns are also reported. Std. Dev. is the standard deviation, and Kurt. refers to the sample kurtosis coefficient  $\hat{\kappa}$ , given in equation (3). ARCH(q) refers to the Lagrange multiplier test for ARCH effects, as proposed by Engle (1982). The test is obtained by running the regression  $r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_q r_{t-q}^2 + u_t$ . Then, under the null of no ARCH effects, the quantity  $TR^2$  is approximately distributed as  $\chi^2(q)$ , where  $T$  is the number of observations and  $R^2$  is the coefficient of determination obtained for the regression. Asterisks \*, \*\* and \*\*\* indicate significance at the 10%, 5% and 1% levels, respectively.

dramatic improvements in fit across all assets when allowing for a GARCH structure. Thus, even when focussing on distributional aspects, dynamic structures should be taken into account in financial modeling.

The estimates for the shape parameters of the GED, reported in Table 2, reveal the reason for this shift in evidence towards the GED. It reflects the density’s degree of fat-tailedness for the unconditional ( $\hat{p}_u$ ) and conditional ( $\hat{p}_c$ ) models. All 25 estimates rise when moving from unconditional to conditional models, indicating lower excess kurtosis for the conditional distributions. Obviously, in the unconditional setting, the assumed distribution can capture some of the properties of the unconditional return distribution induced by volatility dynamics. Once these are incorporated into the model, they account for a portion of the unconditional excess kurtosis, and the GED shape parameter estimate moves further away from unity—the value associated with the Laplace. However, given the substantial improvement in fit, these dynamic structures should not be ignored when modeling high-frequency asset returns.

### 3 Incorporating the Laplace Into More Flexible Models

We demonstrated in Section 2.2 that the Laplace distribution, due to its fixed kurtosis, cannot be expected to serve as a tenable model for many financial return series. However, as detailed in the introduction, the Laplace distribution has some appealing properties as a model for asset returns, and we may not want to be abandon it completely. In this section, we propose two models which incorporate the Laplace distribution as a building block. These models, like the GED, have only one additional parameter to be estimated, but, as will be seen, provide better in- and out-of-sample fit than the GED. Specifically, we consider the use of a mixture and the assumption of conditional normality. As seen in Table 2, the Laplace distribution apparently has too high a kurtosis, as all  $p$ -estimates exceed unity significantly, implying that there is still a component in the return process which is more adequately modeled with a Gaussian assumption.

As such, we propose two models that combine the normal and the Laplace. The first is a discrete mixture of the normal and Laplace distributions. Such models have been proposed by Kanji (1985) and Jones and McLachlan (1990) in an engineering context. The second model appears to be new and assumes that the (conditional) returns can be described by a weighted sum of a Gaussian and a Laplace random variable.

Table 2: Unconditional and Conditional Tests of Laplace against GED Distribution<sup>a</sup>

Stock ID#	Unconditional Models				Conditional Models			
	$L^{\text{Lap}}$	$L^{\text{GED}}$	LR	$\hat{p}_u$	$L^{\text{Lap}}$	$L^{\text{GED}}$	LR	$\hat{p}_c$
1	-2835.1	-2830.8	8.649***	1.176 (0.0639)	-2800.7	-2788.3	24.76***	1.324 (0.0735)
2	-2718.8	-2712.2	13.30***	1.199 (0.0584)	-2689.1	-2673.6	30.98***	1.325 (0.0645)
3	-2579.8	-2564.5	30.54***	1.352 (0.0722)	-2560.8	-2535.1	51.32***	1.520 (0.0884)
4	-2555.8	-2548.0	15.61***	1.234 (0.0640)	-2542.7	-2532.9	19.54***	1.257 (0.0631)
5	-2916.9	-2903.8	26.08***	1.326 (0.0723)	-2884.0	-2861.1	45.82***	1.461 (0.0815)
6	-2614.5	-2611.9	5.245**	1.127 (0.0582)	-2560.3	-2543.0	34.48***	1.386 (0.0760)
7	-2777.3	-2772.3	10.01***	1.176 (0.0596)	-2735.5	-2720.5	30.05***	1.331 (0.0675)
8	-2772.3	-2758.3	27.97***	1.349 (0.0754)	-2738.2	-2711.4	53.65***	1.529 (0.0880)
9	-3028.0	-3019.9	16.15***	1.266 (0.0739)	-2965.8	-2949.0	33.66***	1.396 (0.0797)
10	-2621.2	-2610.5	21.41***	1.309 (0.0753)	-2590.1	-2571.5	37.30***	1.424 (0.0817)
11	-2729.5	-2724.7	9.676***	1.184 (0.0627)	-2702.0	-2691.1	21.74***	1.297 (0.0699)
12	-2707.5	-2702.8	9.537***	1.188 (0.0658)	-2668.6	-2654.9	27.45***	1.346 (0.0763)
13	-2848.2	-2845.2	5.898**	1.135 (0.0585)	-2780.2	-2757.6	45.27***	1.451 (0.0792)
14	-2818.1	-2812.9	10.31***	1.189 (0.0632)	-2785.8	-2768.2	35.33***	1.415 (0.0816)
15	-2668.0	-2657.6	20.75***	1.279 (0.0678)	-2637.4	-2621.6	31.67***	1.352 (0.0683)
16	-2752.5	-2741.2	22.72***	1.310 (0.0736)	-2724.9	-2701.5	46.70***	1.489 (0.0866)
17	-2766.4	-2750.9	30.95***	1.379 (0.0789)	-2743.8	-2722.8	42.10***	1.446 (0.0809)
18	-2827.0	-2814.8	24.44***	1.301 (0.0673)	-2797.8	-2775.4	44.93***	1.417 (0.0706)
19	-2682.6	-2679.2	6.722***	1.144 (0.0583)	-2650.8	-2638.4	24.71***	1.319 (0.0731)
20	-2622.4	-2617.0	10.87***	1.198 (0.0645)	-2589.0	-2574.9	28.32***	1.348 (0.0742)
21	-3260.2	-3257.6	5.068**	1.121 (0.0565)	-3204.8	-3192.4	24.81***	1.310 (0.0711)
22	-2568.6	-2552.9	31.52***	1.378 (0.0777)	-2552.7	-2531.6	42.17***	1.452 (0.0823)
23	-2890.5	-2880.3	20.31***	1.272 (0.0666)	-2845.6	-2824.1	42.93***	1.420 (0.0730)
24	-2748.7	-2742.3	12.72***	1.207 (0.0622)	-2720.5	-2706.9	27.28***	1.315 (0.0671)
25	-2773.7	-2771.3	4.939**	1.123 (0.0582)	-2728.8	-2715.9	25.81***	1.321 (0.0717)
Mean	-2763.3	-2755.3	16.06	1.237	-2728.0	-2710.5	34.91	1.386

<sup>a</sup>Standard errors are given in parentheses.

$L^{\text{Lap}}$  and  $L^{\text{GED}}$  denote the maximum value of the log-likelihood function of the Laplace and GED models, respectively. LR is the likelihood ratio statistic,  $\text{LR} = -2(L^{\text{Lap}} - L^{\text{GED}})$ , which is approximately distributed as  $\chi^2(1)$ . Asterisks \*, \*\* and \*\*\* indicate significance at the 10%, 5% and 1% levels, respectively. The critical values at the 10%, 5% and 1% levels are 6.635, 3.842, and 2.706, respectively.  $\hat{p}_u$  and  $\hat{p}_c$  denote the estimates of the GED shape parameter  $p$  in the unconditional and conditional (GARCH) setting, respectively.

### 3.1 The Gauss–Laplace Mixture (GLaM)

The discrete mixture of a normal and a Laplace distribution, which will be referred to as Gauss–Laplace mixture, or GLaM, was introduced in Kanji (1985) as a model for wind shear data. The mixture is characterized by the *mixing proportion*  $\theta_m$ ,  $0 \leq \theta_m \leq 1$ , which denotes the fraction of normality in the distribution. Because this fraction is generally unknown a priori, it has to be estimated from the data. The standardized density of a zero–mean random variable  $X$  following a GLaM with mixing proportion  $\theta_m$  is

$$f_X(x; \theta_m) = \frac{\theta_m}{\sqrt{2\pi}} e^{-x^2/2} + \frac{1 - \theta_m}{2} e^{-|x|}, \quad \theta_m \in [0, 1]. \quad (9)$$

Its kurtosis, given by

$$\kappa^{\text{GLaM}}(\theta_m) = 3 \left( 1 + \frac{(\theta_m + 4)(1 - \theta_m)}{(2 - \theta_m)^2} \right), \quad (10)$$

varies with  $\theta_m$  and is bounded between 3 (for  $\theta_m = 1$ ) and 6.125 (for  $\theta_m = 2/7$ ). For  $0 < \theta_m < 1/2$ , the kurtosis exceeds that of the Laplace distribution (albeit by a small amount). The dependence of the GLaM kurtosis on the “fraction of normality”,  $\theta_m$ , is shown in Figure 1.

If the sample kurtosis,  $\hat{\kappa}$ , is between 3 and 6.125, then an initial estimate of  $\theta_m$  can easily be computed by the method of moments. Solving (10) for  $\hat{\theta}_m$  implies

$$\hat{\theta}_m = \frac{4c - 3 \pm \sqrt{25 - 24c}}{2(c + 1)}, \quad c = \frac{\hat{\kappa} - 3}{3}. \quad (11)$$

While the positive root is admissible for all values of  $\hat{\kappa}$  between 3 and 6.125, the negative root yields admissible solutions, i.e.,  $\theta_m \in [0, 1]$ , only for  $\hat{\kappa} \in [6, 6.125]$ . Hence, the solution is not unique in this interval. However, it may be preferable to use the negative root there, as the positive solution implies an unrealistic high fraction of normality. For example, if sample kurtosis is 6, which is the kurtosis of the Laplace distribution, then the moment estimates for  $\theta_m$  are 0 and 0.5 using the negative and the positive root, respectively.

### 3.2 The Gauss–Laplace Sum (GLaS)

Let  $Z$  be the weighted sum of a normal random variable,  $N$ , and a Laplace random variable,  $L$ , i.e.,

$$Z = \theta_s N + (1 - \theta_s) L, \quad \theta_s \in [0, 1]. \quad (12)$$

We refer to (12) as a Gauss–Laplace sum, or GLaS. The derivation of the probability density function (pdf) and the cumulative distribution function (cdf) of a GLaS random variable is



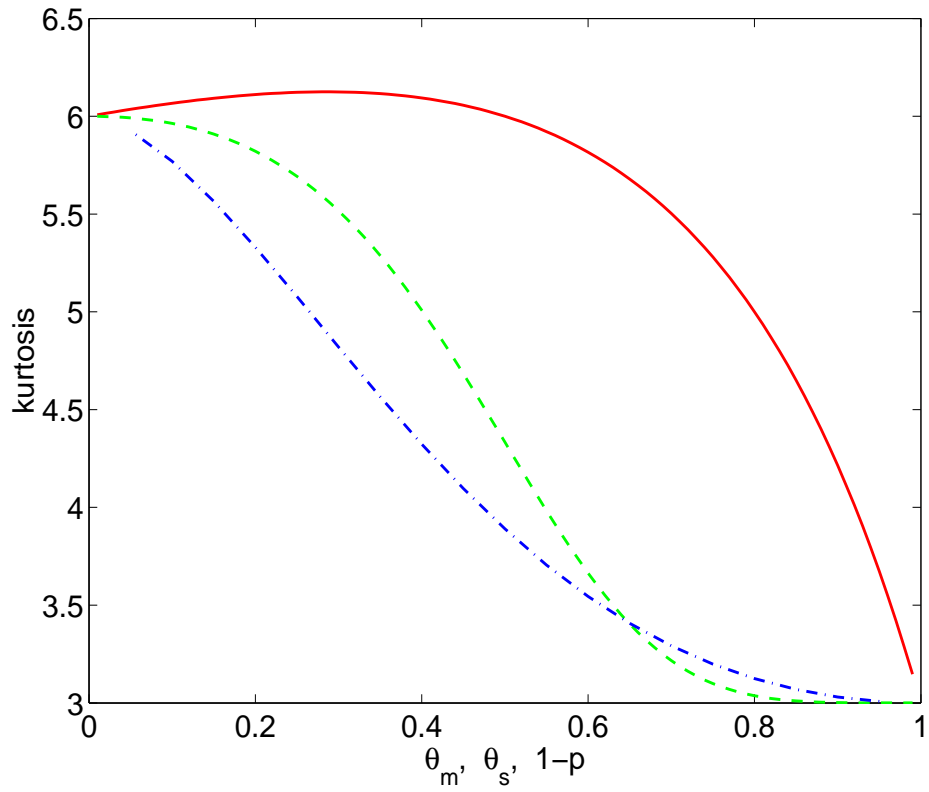


Figure 1: Kurtosis of the Gauss–Laplace mixture (solid line) as a function of  $\theta_m$ , the Gauss–Laplace sum (dashed line) as a function of  $\theta_s$ , and the hyperbolic (dash-dot line), as a function of one minus its shape parameter  $p$ .

detailed in the Appendix. The pdf is needed for maximum likelihood estimation and the cdf for Value-at-Risk estimation. The Appendix also shows that the odd moments of  $Z$  are zero; and the even moments are given by

$$E(Z^m) = (1 - \theta_s)^m m! e_{m/2} \left( \frac{\theta_s^2}{2(1 - \theta_s)^2} \right), \quad (13)$$

where  $e_n(x) = \sum_{k=0}^n x^k/k!$  denotes the exponential sum function.

Using (13), the kurtosis of the GLaS is found to be

$$\kappa^{\text{GLaS}}(\theta_s) = 3 \left[ 1 + \left( \frac{2}{r^2 + 2} \right)^2 \right], \quad r = \frac{\theta_s}{1 - \theta_s}, \quad (14)$$

which exceeds 3 for  $\theta_s \in [0, 1)$ . Note that  $\kappa^{\text{GLaS}}(\theta_s)$  is bounded between 3 and 6, the respective  $\kappa$ -values of the Gaussian and the Laplace. A plot of the kurtosis as a function of  $\theta_s$  is shown in Figure 1.<sup>2</sup>

If the sample kurtosis,  $\hat{\kappa}$ , is between 3 and 6, then (14) can be used to obtain a simple method-of-moments estimate for  $\theta_s$ . From (14), we obtain

$$\hat{\theta}_s = \frac{\sqrt{b}}{1 + \sqrt{b}}, \quad \text{with } b = 2 \left( \sqrt{\frac{3}{\hat{\kappa} - 3}} - 1 \right). \quad (15)$$

## 4 Empirical Results for Gauss–Laplace Models

In this section, the two Gauss–Laplace models introduced above are fitted to the 25 German stock return series. The estimates for the “normality weights”  $\theta_m$  and  $\theta_s$  are discussed. We emphasize the conditional (GARCH) models, because of their clear superiority over the unconditional specifications. The models’ in- and out-of-sample performance is studied in some detail.

The results from maximum likelihood estimation are shown in Table 3. We first consider the estimates  $\hat{\theta}_m$  and  $\hat{\theta}_s$  for the normal weights in the GLaM and GLaS models, and defer the discussion of likelihood values to the next subsection. The means of  $\hat{\theta}_m$  and  $\hat{\theta}_s$ , taken over the 25 series, are reported at the bottom of the table. The average  $\hat{\theta}_m$  for the unconditional and conditional models are 0.503 and 0.662, respectively, in line with the common findings that conditional (GARCH) innovations exhibit less kurtosis than that of their unconditional counterpart, and that conditional normality is (still) not appropriate. The corresponding values for  $\hat{\theta}_s$  are 0.387 and 0.489.

---

<sup>2</sup> Although kurtosis is monotonic in  $\theta_s$ , the moments are not. For example, the variance  $E(Z^2) = \theta_s^2 + 2(1 - \theta_s)^2$  is a quadratic function in  $\theta_s$ , which has its minimum at  $\theta_s = 2/3$ , namely  $E(Z^2) = 2/3$ .

We now concentrate on the conditional models, illustrated in the right part of the table, because the shape–parameter estimates obtained for the unconditional models are expected to have a systematic downward bias (for reasons explained at the end of Section 2.2). The estimates are relatively stable across all stocks, with  $\hat{\theta}_m$  ranging roughly from 0.5 to 0.8, and  $\hat{\theta}_s$  ranging from 0.4 to 0.55. This was to be expected, as the stocks are relatively similar; i.e., blue–chip stocks with high market capitalization traded in the same market, so the fraction of normality arising perhaps from the pattern of the inflow of relevant new information should not be too different.

In order to examine whether this pattern is stable across time, we split the sample into two non-overlapping subsamples of equal length 610, and estimate parameters separately for each of them. The results are illustrated in Figure 2, where the left panel gives the estimates for  $\hat{\theta}_m$  in the Gauss–Laplace mixture GARCH model, and the right panel those for  $\hat{\theta}_s$  in the Gauss–Laplace sum GARCH model. Due to sampling error, the estimates of  $\theta_m$  and  $\theta_s$  are slightly more “volatile” in the shorter subperiods. However, the general pattern observed over the whole sample period also holds in both subperiods, indicating considerable stability of the normal fraction in the return process over the period considered.

#### 4.1 In–Sample Performance

To compare the in–sample performance of the Gauss–Laplace models with the GED distribution, we rely on the likelihood–criterion. Note that the models are not nested, so that a likelihood ratio test is not applicable. However, the log–likelihood value obtained from maximum likelihood estimation may be viewed as an overall measure of fit and allows to judge which candidate is more likely to have generated the data. Consequently, standard model selection criteria such as the AIC (Akaike, 1973) or the BIC (Schwarz, 1978) are widely used for comparing non–nested models, and there is evidence that these measures are useful for choosing among GARCH–models with respect to out–of–sample prediction (see Mittnik and Paoletta, 2000, and the references therein). These model selection criteria use the maximum log–likelihood value of each model but punish the use of additional free parameters. As the competing models under examination, i.e., GED, GLaM and GLaS, all have the same number of parameters both when estimated unconditionally and conditionally, ranking them according to any of these criteria is equivalent to ranking them according to only their likelihood values.

Table 4 summarizes the results of a likelihood–based comparison of the three models both for the unconditional and the conditional (GARCH) models. It lists, for each model, the

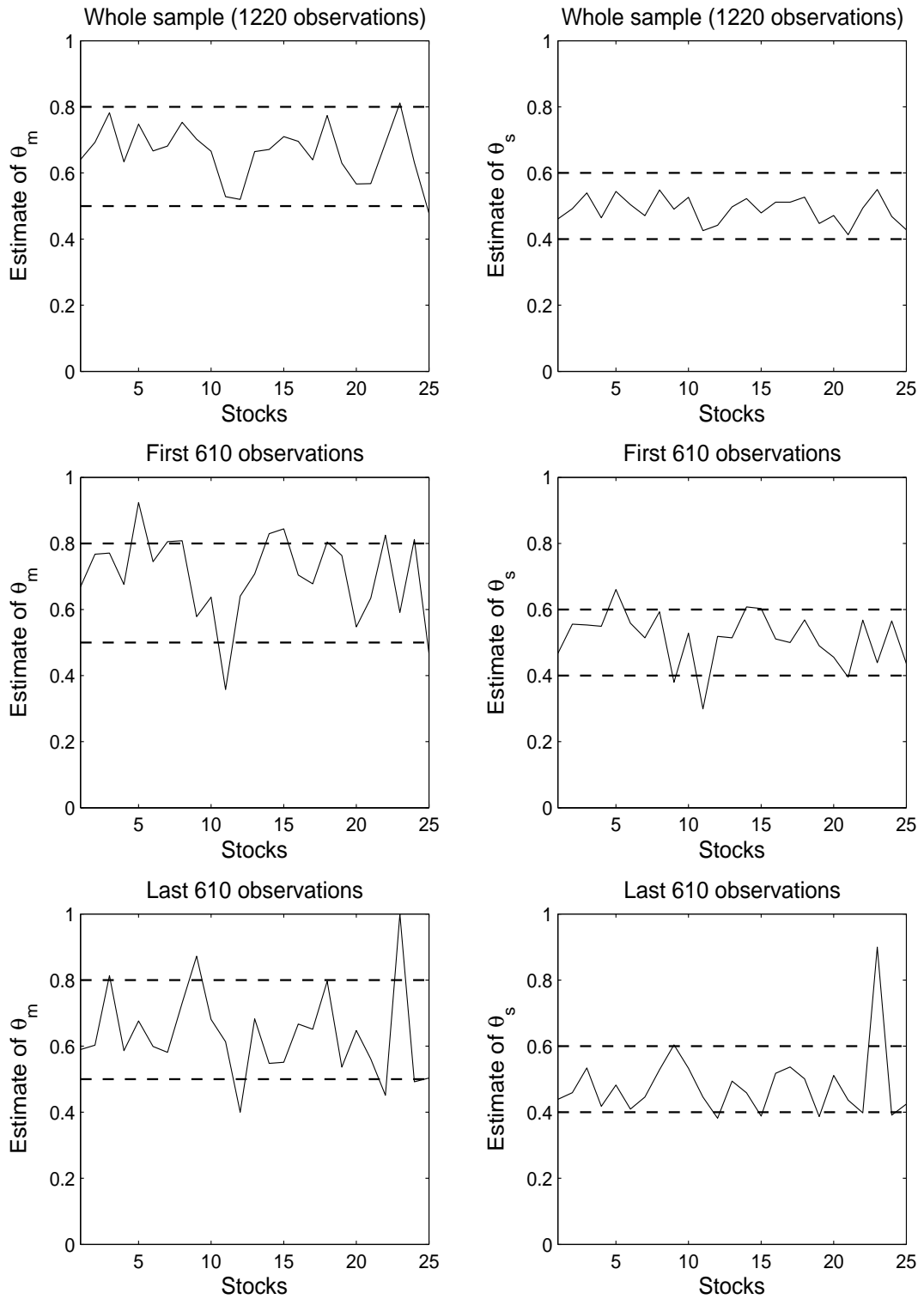


Figure 2: Estimates of the normality parameters  $\theta_m$  (GLaM) and  $\theta_s$  (GLaS) over the whole sample period, as well as over non-overlapping subperiods of equal length.

Table 3: Likelihood Values and Parameter Estimates for GLaM and GLaS Models<sup>a</sup>

Stock ID#	Unconditional Models				Conditional Models			
	$L^{\text{GLaM}}$	$\hat{\theta}_m$	$L^{\text{GLaS}}$	$\hat{\theta}_s$	$L^{\text{GLaM}}$	$\hat{\theta}_m$	$L^{\text{GLaS}}$	$\hat{\theta}_s$
1	-2828.6	0.406 (0.102)	-2830.2	0.345 (0.0515)	-2784.1	0.640 (0.0854)	-2786.5	0.461 (0.0431)
2	-2697.9	0.621 (0.0746)	-2703.9	0.415 (0.0400)	-2665.3	0.692 (0.0686)	-2662.7	0.492 (0.0345)
3	-2562.2	0.647 (0.0858)	-2562.2	0.464 (0.0441)	-2537.7	0.782 (0.0795)	-2536.1	0.540 (0.0475)
4	-2543.0	0.630 (0.0980)	-2546.8	0.429 (0.0508)	-2530.4	0.633 (0.0972)	-2529.2	0.464 (0.0450)
5	-2902.8	0.610 (0.0939)	-2902.8	0.440 (0.0457)	-2858.2	0.748 (0.0759)	-2855.8	0.544 (0.0415)
6	-2606.7	0.440 (0.0989)	-2610.3	0.331 (0.0533)	-2540.5	0.666 (0.0782)	-2541.5	0.504 (0.0417)
7	-2766.9	0.499 (0.0962)	-2768.1	0.362 (0.0459)	-2718.7	0.681 (0.0858)	-2715.4	0.471 (0.0431)
8	-2761.7	0.622 (0.111)	-2759.7	0.426 (0.0533)	-2716.6	0.753 (0.0806)	-2712.0	0.549 (0.0478)
9	-3024.6	0.361 (0.133)	-3024.3	0.374 (0.0625)	-2945.0	0.702 (0.0803)	-2946.7	0.490 (0.0421)
10	-2611.7	0.561 (0.107)	-2613.0	0.438 (0.0519)	-2574.3	0.666 (0.0879)	-2573.0	0.526 (0.0468)
11	-2722.8	0.422 (0.105)	-2722.5	0.344 (0.0472)	-2694.2	0.528 (0.0993)	-2689.8	0.425 (0.0504)
12	-2705.3	0.270 (0.122)	-2703.5	0.294 (0.0550)	-2657.1	0.520 (0.0917)	-2655.6	0.442 (0.0465)
13	-2843.8	0.303 (0.0977)	-2841.5	0.309 (0.0446)	-2759.5	0.664 (0.0821)	-2755.2	0.497 (0.0407)
14	-2809.6	0.528 (0.107)	-2813.2	0.361 (0.0601)	-2773.5	0.671 (0.0948)	-2769.4	0.522 (0.0482)
15	-2657.0	0.562 (0.103)	-2655.8	0.400 (0.0476)	-2615.5	0.710 (0.0822)	-2616.1	0.479 (0.0406)
16	-2743.6	0.517 (0.111)	-2740.7	0.405 (0.0463)	-2707.0	0.695 (0.0957)	-2702.2	0.511 (0.0466)
17	-2753.8	0.621 (0.105)	-2752.4	0.466 (0.0467)	-2727.5	0.640 (0.0825)	-2720.7	0.511 (0.0391)
18	-2809.2	0.655 (0.0874)	-2809.7	0.443 (0.0415)	-2762.0	0.774 (0.0653)	-2766.7	0.527 (0.0363)
19	-2674.8	0.448 (0.102)	-2676.5	0.344 (0.0483)	-2639.1	0.630 (0.0978)	-2637.8	0.447 (0.0487)
20	-2612.3	0.474 (0.0924)	-2615.8	0.399 (0.0448)	-2575.8	0.567 (0.0872)	-2573.4	0.471 (0.0409)
21	-3254.2	0.347 (0.0935)	-3253.4	0.313 (0.0463)	-3190.2	0.568 (0.0917)	-3188.4	0.413 (0.0410)
22	-2555.4	0.596 (0.0975)	-2552.9	0.448 (0.0466)	-2536.6	0.690 (0.0923)	-2531.4	0.494 (0.0465)
23	-2875.1	0.592 (0.0862)	-2876.9	0.426 (0.0430)	-2814.4	0.812 (0.0691)	-2817.3	0.550 (0.0386)
24	-2736.3	0.533 (0.0914)	-2738.7	0.397 (0.0437)	-2702.3	0.631 (0.0867)	-2700.2	0.468 (0.0375)
25	-2769.4	0.307 (0.0989)	-2768.7	0.300 (0.0498)	-2721.2	0.478 (0.0917)	-2714.9	0.428 (0.0451)
Mean	-2753.1	0.503	-2753.7	0.387	-2709.9	0.662	-2707.9	0.489

<sup>a</sup>Standard errors are given in parentheses. $L^{\text{GLaM}}$  and  $L^{\text{GLaS}}$  denote the maximum value of the log-likelihood function of the Gauss-Laplace mixture and Gauss-Laplace sum, respectively.

Table 4: Frequency of Likelihood-Based Rankings of Unconditional and Conditional Distributional Models

Rank	Unconditional models			Conditional Models		
	GED	GLaM	GLaS	GED	GLaM	GLaS
1	5	11	9	6	6	13
2	3	8	14	7	6	12
3	17	6	2	12	13	0
Avg. Rank	2.48	1.80	1.72	2.24	2.28	1.48

number of stocks for which a model is ranked as the best (Rank 1), second best or worst (Rank 3). The last row reports the average ranks over the 25 stocks for each model, which may be viewed as the overall goodness-of-fit criterion. According to this, the GLaS would be the preferred model both in the unconditional and the conditional case. Another possible criterion is the number of times a model is selected as the best. In this case, GLaM dominates somewhat in the unconditional case, while GLaS is clearly preferred in the GARCH context. *The GED is not favored by any of these criteria.* In particular, the GLaS is clearly superior to the GED and the GLaM when used as conditional distribution in a GARCH model. The latter seems more important, given the strong evidence for GARCH effects in the data.

However, the foregoing analysis does not take into account the fact that small differences in the log-likelihood value—and hence in AIC or BIC—may be viewed as negligible. For example, according to Kass and Raftery (1995), a BIC difference of less than two corresponds to “not worth more than a bare mention”; while differences between two and six imply positive evidence; differences between six and ten give rise to strong evidence; and differences greater than ten invoke very strong evidence. It turns out, however, that most of the differences in log-likelihood used to produce Table 4 are in the region of positive to very strong evidence according to Kass and Raftery (1995), as is illustrated in Figure 3. As two times the difference in log-likelihood is used in the computation of AIC and BIC, Figure 3 shows the differences in the values of  $L$  multiplied by 2 for the 25 stocks for each pairwise model comparison.

## 4.2 Out-of-Sample Performance

Comparing the performance of the distributional models with respect to out-of-sample forecasting, we shall only consider GARCH models, given the overwhelming evidence for conditional heteroskedasticity in the return series. Also, we focus on the evaluation of density

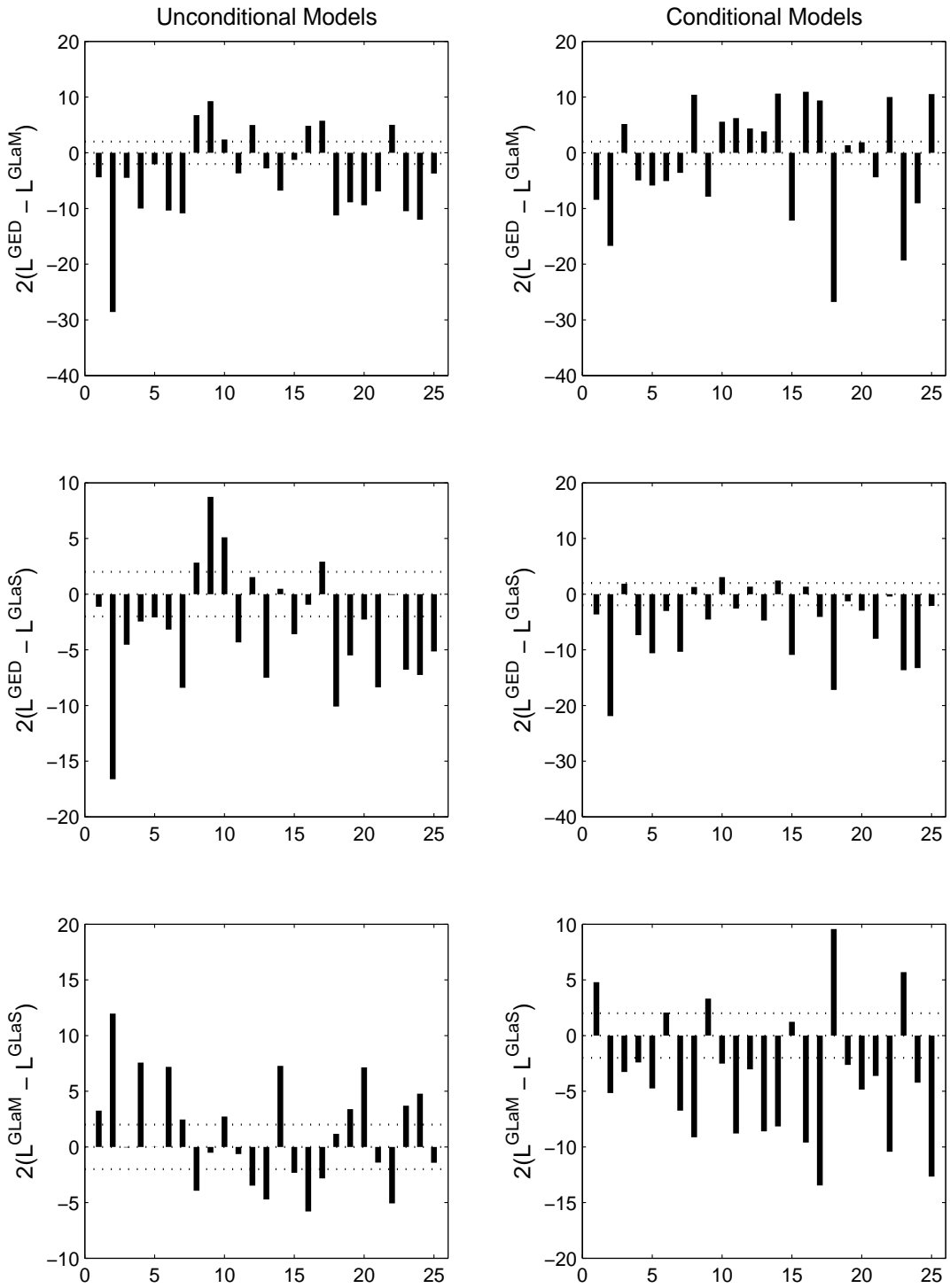


Figure 3: The left panel shows two times the difference in log-likelihood of unconditional models, the right panel is the similar, but for conditional (GARCH) models.

forecasts, since they are the cornerstone for reliable Value-at-Risk (VaR) predictions, and, thus, of great importance in modern risk management (Diebold, Gunther and Tay, 1998). Compared to VaR analysis, density forecasts are concerned not only with selected quantiles of the conditional distribution, but with its characteristics over the whole support, and, thus, their evaluation seems appropriate when fully specified distributional models are subject to testing. Employing the first half of the sample (i.e., the first 610 observations) for estimation, we produce density forecasts for the 610 observations in the second half in the sample, reestimating the models after each forecast.

The evaluation of density forecasts is based on the series of estimated  $\epsilon_t$  values in (6), denoted by  $\hat{\epsilon}_t$ ,  $t = 611, \dots, 1220$ . They are computed as  $\hat{\epsilon}_t = r_t - \hat{\mu}$ , where  $\hat{\mu}$  is the estimate of the mean. If the model used to generate the forecast error series were correctly specified, then the series of standardized forecast errors,  $\hat{\epsilon}_t/\hat{\sigma}_t$ , where  $\hat{\sigma}_t$  is the (estimated) conditional standard deviation obtained from (8), should follow the postulated zero mean and unit scale distribution.

However, as these standardized residuals depend on the distribution used, we employ the probability integral (or Rosenblatt) transformation method in order to make the forecast errors of the different models comparable. This produces a series of transformed residuals, denoted  $\hat{u}_t$ , which are independently uniformly distributed over the unit interval—provided the underlying model is correctly specified (Rosenblatt, 1952; see also Diebold *et al.*, 1998). For a density forecast with cdf  $\hat{F}(\cdot | \Psi_{t-1})$ , the probability integral transform is simply defined as

$$\hat{u}_t = \hat{F}(\hat{\epsilon}_t | \Psi_{t-1}). \quad (16)$$

The cdf of the Laplace distribution is given in (22) in the Appendix, where we also derive the cdf of the GLaS. The cdf of the GLaM is simply a linear combination of the cdf of the normal and the Laplace distribution; and the cdf of the GED distribution is given by (see Johnson, Kotz and Balakrishnan, 1995)

$$F_X(x; p) = \begin{cases} \frac{1}{2} \left[ 1 - \Gamma^{\text{inc}}\left(\frac{1}{p}, \frac{|x|^p}{2}\right) \right], & \text{if } x \leq 0, \\ \frac{1}{2} \left[ 1 + \Gamma^{\text{inc}}\left(\frac{1}{p}, \frac{x^p}{2}\right) \right], & \text{if } x > 0, \end{cases}$$

with  $\Gamma^{\text{inc}}(\cdot, \cdot)$  being the incomplete gamma function given below in (24).

Palm and Vlaar (1997) propose a Pearson goodness-of-fit test to test for iid uniformity of (16); see also Stuart, Ord and Arnold (1999, Ch. 25) for further account of this test. One drawback associated with this test is the fact that the choice for the number of bins may affect



the test results when dealing with continuous distributions. To circumvent this problem, we apply the inverse normal cdf as a second transformation, as advocated by Palm and Vlaar (1997); that is

$$z_t = \Phi^{-1}(\hat{u}_t), \quad (17)$$

where  $\Phi(\cdot)$  is the standard normal cdf. The  $z_t$  are then iid  $N(0,1)$  distributed, provided the underlying model is correct, so that the  $\hat{u}_t$  values are indeed uniformly distributed. Berkowitz (2001) shows that inaccuracies in the density forecast will be preserved in the transformed data.

To test for standard normality of the  $z_t$  values, a joint test involving the first four moments is employed using the likelihood ratio approach proposed by Berkowitz (2001). It is implemented as follows: Under the alternative hypothesis, we let  $z_t$  be distributed according to the skewed exponential power distribution (SEP) of Fernandez, Osiewalski and Steel (1995), with density

$$f(z_t; \mu, \sigma, d, \theta) = K \begin{cases} \exp \left\{ -\frac{1}{2} \left( \frac{|z_t - \mu| \theta}{\sigma} \right)^d \right\}, & \text{if } z_t < \mu, \\ \exp \left\{ -\frac{1}{2} \left( \frac{z_t - \mu}{\sigma \theta} \right)^d \right\}, & \text{if } z_t \geq \mu, \end{cases} \quad (18)$$

where  $K = [\sigma(\theta + \theta^{-1})2^{1/d}d^{-1}\Gamma(d^{-1})]^{-1}$ . This distribution nests the normal for  $\theta = 1$  and  $d = 2$ . For  $\theta < 1$  ( $\theta > 1$ ) the density is skewed to the left (right); and is fat-tailed for  $d < 2$ . If  $\hat{\mu}$ ,  $\hat{\sigma}$ ,  $\hat{d}$ , and  $\hat{\theta}$  are the values which maximize the log-likelihood

$$L(\mu, \sigma, d, \theta) = \sum_{t=1}^T \log f(z_t; \mu, \sigma, d, \theta), \quad (19)$$

then the likelihood ratio statistic

$$\text{LR}_B = 2[L(\hat{\mu}, \hat{\sigma}, \hat{d}, \hat{\theta}) - L(0, 1, 2, 1)] \quad (20)$$

is asymptotically distributed as  $\chi^2(4)$  if the  $z_t$  are  $N(0,1)$ -distributed.

The test results are summarized in Table 5. For each distributional model, the table reports the percentage of stock return series for which the models are not rejected at the 0.1, 0.05 and 0.01 significance level, respectively. The results in Table 5 show that the GED distribution is rejected more often than the GLaM and GLaS, although the results are not as clear-cut as in the in-sample comparison. Also, we note a slight superiority of the GLaM over the GLaS model. Finally, an interesting result is that, although the pure Laplace distribution is rejected for most of the time series, there are some stocks that seem to be reasonably described by the Laplace distribution.

Table 5: Density Forecasting Performance of GARCH Models Driven by Alternately Distributed Innovations<sup>a</sup>

Level	Distributional Model			
	Lap	GED	GLaM	GLaS
0.01	48	80	96	88
0.05	24	68	88	68
0.1	20	56	72	68

<sup>a</sup>Percentage of stocks for which the respective distributional model is not rejected at the indicated significance levels according to the likelihood-ratio test defined by Equation (20).

## 5 Use of the Hyperbolic Distribution

There has been a growing interest in hyperbolic processes for modeling financial data in recent years; see, e.g., Bibby and Sørensen (2003) for an overview and account of the historical development of the hyperbolic and related distributions, and Schoutens (2003, Sec. 5.3) for their use in Lévy and Ornstein–Uhlenbeck processes. A useful parameterization of the location–zero, scale–one hyperbolic density function is

$$f(z; p, q) = C_1 \exp \left\{ C_2 \left( p\sqrt{1+z^2} - qz \right) \right\}, \quad 0 < p \leq 1, \quad |q| < p, \quad (21)$$

where

$$C_1 = \frac{\sqrt{p^2 - q^2}}{2pK_1(p^{-2} - 1)}, \quad C_2 = \frac{p^2 - 1}{p^2\sqrt{p^2 - q^2}},$$

and  $K_1$  is the modified Bessel function of the third kind with index  $\nu = 1$ . Integral expressions for  $K_1(x)$  and related functions necessary for working with the hyperbolic distribution are given in Bibby and Sørensen (2003), and can be reliably computed using, for example, the built-in functions provided in Matlab.

For  $q = 0$ , density (21) is symmetric; for  $q < 0$  ( $q > 0$ ) it is left (right) skewed. As  $p \rightarrow 0$ ,  $f(z)$  approaches a normal density, and with  $p = 1$ ,  $q = 0$ , it coincides with a scaled Laplace. Thus,  $p$  and  $q$  can be interpreted as measures of kurtosis and skewness, respectively. The kurtosis as a function of  $p$  is illustrated in Figure 1;<sup>3</sup> it is seen to be bound between 3 and 6, which correspond to the limiting cases of normal and Laplace, respectively.

We repeat the in-sample comparison conducted above with the hyperbolic distribution and also the symmetric hyperbolic, obtained by restricting  $q = 0$  in (21). The latter provides

---

<sup>3</sup> The values were computed via numeric integration. Complicated expressions for the moments do exist however; see, e.g., Küchler, Neumann, Sørensen, and Streller (1999, p. 5).

a fairer comparison to the GLaM and GLaS, as these cannot exhibit skewness. Note also that the symmetric hyperbolic has the same number of shape parameters as GLaM and GLaS (namely one), so that log likelihoods at the respective MLEs can be directly compared. We only summarize the results for the unconditional models: Out of 25 stocks, the hyperbolic is superior to the GED in 20 cases; it is better than the GLaS in 13 cases, and is better than the GLaM in 12 cases. This reinforces the above finding that the GED is a relatively poor candidate distribution, and shows that the GLaS and GLaM are “worthy competitors” to the hyperbolic.

More relevant is the conditional (GARCH) setting. Table 6 shows the likelihood results in this case, as well as the estimated shape parameters, of the symmetric and asymmetric hyperbolic distribution when coupled with the GARCH structure (7) and (8). From the left panel of Table 6, the values of  $\hat{p}$  clearly rule out use of the special cases of normal or Laplace. From the right panel, based on the difference of the likelihood values, we see that only for series 21 and 23 is there mild asymmetry; otherwise, skewness is not a prominent feature in this data. This is reinforced from the mean of the likelihood values and parameter estimates, which are shown in the last column.

Consider comparing the achieved log-likelihood values between the symmetric hyperbolic and those of GLaM and GLaS. In order to ascertain if the latter two new models are useful, we construct the difference

$$D_i = \max(L_i^{\text{GLaM}}, L_i^{\text{GLaS}}) - L_i^{\text{HypS}}, \quad i = 1, \dots, 25,$$

for each of the 25 time series (and, again, restricting attention to the conditional, i.e., GARCH, models). Negative values of  $D$  thus lend favor to the hyperbolic. In our case, 11 of the 25 series have a negative  $D$ , with an average of only  $-0.59$ , while 14 are positive, with an average of  $1.9$ . While this result indeed implies favoring the GLaM and GLaS distributions, matters are more clear when considering the following. For 20 of the series, the absolute difference,  $|D|$  is less than  $2.0$ , while for 5 (namely series 2, 1, 23, 9 and 18 in increasing order of magnitude),  $D$  is larger than  $2.0$ , ranging from  $2.5$  to  $6.5$ . Thus, in light of Kass and Raftery’s (1995) aforementioned guidelines for judging such differences, we conclude that there is positive evidence for favoring the new GLaS and GLaM distributions over the symmetric hyperbolic.

Table 6: Likelihood Values and Selected Parameter Estimates for Symmetric and Asymmetric hyperbolic GARCH Models<sup>a</sup>

Stock ID#	Symmetric Hyperbolic		Asymmetric Hyperbolic		
	$L^{\text{HypS}}$	$\hat{p}$	$L^{\text{Hyp}}$	$\hat{p}$	$\hat{q}$
1	-2786.7	0.638 (0.0596)	-2786.2	0.628 (0.0608)	-0.021 (0.0223)
2	-2665.2	0.607 (0.0508)	-2665.1	0.612 (0.0514)	0.014 (0.0235)
3	-2535.5	0.534 (0.0463)	-2534.6	0.521 (0.0481)	-0.028 (0.0199)
4	-2529.3	0.644 (0.0606)	-2529.2	0.643 (0.0609)	-0.003 (0.0256)
5	-2856.1	0.518 (0.0612)	-2855.8	0.513 (0.0628)	-0.012 (0.0178)
6	-2541.7	0.586 (0.0625)	-2541.7	0.585 (0.0627)	-0.005 (0.0192)
7	-2715.8	0.624 (0.0566)	-2715.6	0.619 (0.0582)	-0.016 (0.0218)
8	-2711.4	0.513 (0.0696)	-2711.2	0.510 (0.0712)	-0.010 (0.0172)
9	-2950.5	0.491 (0.0401)	-2949.8	0.512 (0.0040)	-0.009 (0.0186)
10	-2572.8	0.504 (0.0764)	-2572.1	0.546 (0.0736)	-0.015 (0.0188)
11	-2689.7	0.659 (0.0501)	-2689.1	0.670 (0.0457)	0.026 (0.0244)
12	-2655.1	0.654 (0.0638)	-2655.0	0.654 (0.0640)	-0.010 (0.0231)
13	-2755.4	0.570 (0.0607)	-2755.3	0.572 (0.0606)	0.006 (0.0178)
14	-2768.7	0.562 (0.0685)	-2768.6	0.558 (0.0645)	-0.009 (0.0201)
15	-2616.6	0.609 (0.0571)	-2616.6	0.609 (0.0578)	0.000 (0.0215)
16	-2701.5	0.545 (0.0693)	-2700.6	0.573 (0.0678)	0.030 (0.0231)
17	-2721.6	0.589 (0.0430)	-2721.6	0.588 (0.0434)	-0.004 (0.0220)
18	-2768.5	0.549 (0.0544)	-2768.2	0.554 (0.0544)	0.016 (0.0199)
19	-2637.4	0.648 (0.0628)	-2637.1	0.661 (0.0640)	0.019 (0.0259)
20	-2573.7	0.621 (0.0596)	-2573.6	0.620 (0.0600)	-0.007 (0.0216)
21	-3186.4	0.604 (0.0274)	-3183.6	0.598 (0.0224)	-0.045 (0.0190)
22	-2530.9	0.575 (0.0664)	-2530.8	0.581 (0.0667)	0.010 (0.0238)
23	-2818.4	0.520 (0.0588)	-2815.4	0.508 (0.0592)	-0.040 (0.0170)
24	-2701.6	0.624 (0.0537)	-2701.3	0.614 (0.0561)	-0.017 (0.0213)
25	-2714.7	0.669 (0.0595)	-2714.6	0.672 (0.0596)	0.010 (0.0235)
Mean	-2708.2	0.586	-2707.7	0.589	-0.005

<sup>a</sup>Standard errors are given in parentheses.

$L^{\text{HypS}}$  and  $L^{\text{Hyp}}$  denote the maximum value of the log-likelihood function of the symmetric hyperbolic and unrestricted (asymmetric) hyperbolic, respectively.

## 6 Conclusions

An appealing approach to modeling returns on financial assets and for explaining the excess kurtosis observed in their empirical distribution is to view the return distribution as a finite-variance mixture of normal distributions. It also reconciles the assumption of conditional normality with the observed—even conditional—fat-tailedness. Recently, Linden (2001) argued that the Laplace distribution is a promising candidate, as it fits into the mixture framework and is justifiable on empirical grounds. In this paper, we have demonstrated that the applicability of the Laplace distribution is not as general as perhaps suggested in Linden (2001). In the analysis of 25 German blue-chip stocks, the Laplace-hypothesis was rejected against the GED distribution. The rejection is even stronger when the time-series properties of the returns, i.e., second-moment dynamics, are taken into account.

An outright rejection of the Laplace model may not be called for, however, given its attractive properties. More flexible but still simple models, building upon the Laplace and combining it with the Gaussian distribution, have been proposed. An empirical investigation of 25 German blue-chip stocks shows that models based on such Gauss-Laplace combinations offer valid descriptions of the conditional distribution both in- and out-of-sample.

A comparison with the symmetric hyperbolic distribution indicates that the new models are at least as good, and sometimes more suitable for modeling the conditional distribution of asset returns. For the 25 series under consideration in this paper, only two exhibited mildly substantial skewness, so that the symmetry restriction of the models considered was not a detriment. Of course, significant asymmetry in some asset return series is well-documented. Fortunately, both Gauss-Laplace models examined here can easily be extended to support asymmetry. In the Gauss-Laplace mixture (9), this can be achieved by allowing for different means in the two components. In the Gauss-Laplace sum (12), the Laplace variable  $L$  could have an asymmetric Laplace distribution, as given, e.g., in Kotz *et al.* (2001). This gives rise to the asymmetric Gauss-Laplace sum, the properties of which can be derived using the methods employed in the Appendix. Finally, a GARCH structure which couples a normal and Laplace component along the lines of the model proposed in Haas, Mittnik and Paolella (2004a,b) could be entertained. In addition to allowing for *time-varying* skewness and kurtosis, the Laplace component could give rise to a more parsimonious fit and potentially better forecasts. The usefulness of such extensions for density prediction is currently under investigation.

## Appendix

### Kurtosis of the GED distribution (4)

Differentiating (5) with respect to  $p$  produces

$$\begin{aligned} \frac{d\kappa^{\text{GED}}(p)}{dp} &= -\frac{5\Gamma'\left(\frac{5}{p}\right)\Gamma\left(\frac{1}{p}\right)\Gamma^2\left(\frac{3}{p}\right) + \Gamma'\left(\frac{1}{p}\right)\Gamma\left(\frac{5}{p}\right)\Gamma^2\left(\frac{3}{p}\right) - 6\Gamma\left(\frac{5}{p}\right)\Gamma\left(\frac{1}{p}\right)\Gamma\left(\frac{3}{p}\right)\Gamma'\left(\frac{3}{p}\right)}{p^2\Gamma^4\left(\frac{3}{p}\right)} \\ &= \frac{2\Gamma\left(\frac{5}{p}\right)\Gamma\left(\frac{1}{p}\right)}{p\Gamma^2\left(\frac{3}{p}\right)} \left[ \frac{3}{p}\psi\left(\frac{3}{p}\right) - \frac{1}{2}\frac{5}{p}\psi\left(\frac{5}{p}\right) - \frac{1}{2}\frac{1}{p}\psi\left(\frac{1}{p}\right) \right], \end{aligned}$$

where  $\psi(\cdot)$  is the digamma function, i.e.,  $\psi(x) = (\log \Gamma(x))' = \Gamma'(x)/\Gamma(x)$ . From Alzer (1997, Theorem 4), the function  $x\psi(x)$  is strictly convex on  $(0, \infty)$ . Hence, the term in brackets is negative, and kurtosis strictly decreases with increasing  $p$ .

### Properties of the Gauss–Laplace Sum (GLaS)

We begin with deriving the cumulative distribution function (cdf) of the GLaS defined by (12). Note that  $Z = X + Y$ , where the densities of  $X$  and  $Y$  are given by

$$f_X(x; \theta_s) = \frac{1}{\sqrt{2\pi}\theta_s} \exp\left\{-\frac{x^2}{2\theta_s^2}\right\} \quad \text{and} \quad f_Y(y; \theta_s) = \frac{1}{2(1-\theta_s)} \exp\left\{-\frac{|y|}{1-\theta_s}\right\},$$

respectively. Let the density and the cdf of the standard normal distribution be denoted by  $\phi(\cdot)$  and  $\Phi(\cdot)$ , respectively. To derive the cdf of  $Z$ ,  $F_Z(z)$ , we make use of the fact the cdf of the Laplace distribution is available in closed form, i.e.,

$$F_Y(y) = \frac{1}{2(1-\theta_s)} \int_{-\infty}^y \exp\left\{-\left|\frac{\xi}{1-\theta_s}\right|\right\} d\xi = \begin{cases} \frac{1}{2} \exp\left\{\frac{y}{1-\theta_s}\right\} & \text{for } y < 0 \\ 1 - \frac{1}{2} \exp\left\{-\frac{y}{1-\theta_s}\right\} & \text{for } y \geq 0. \end{cases} \quad (22)$$

Using this, we can derive

$$\begin{aligned} F_Z(z) &= \text{P}(X + Y \leq z) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} \frac{1}{2(1-\theta_s)} \exp\left\{-\frac{|y|}{1-\theta_s}\right\} dy \frac{1}{\sqrt{2\pi}\theta_s} \exp\left\{-\frac{x^2}{2\theta_s^2}\right\} dx \\ &=: A + B, \end{aligned}$$

where, using (22) and completing the square,

$$\begin{aligned}
A &= \int_{-\infty}^z \int_{-\infty}^{z-x} \frac{1}{2(1-\theta_s)} \exp\left\{-\frac{|y|}{1-\theta_s}\right\} dy \frac{1}{\sqrt{2\pi\theta_s}} \exp\left\{-\frac{x^2}{2\theta_s^2}\right\} dx \\
&= \frac{1}{\sqrt{2\pi\theta_s}} \int_{-\infty}^z \left(1 - \frac{1}{2} \exp\left\{-\frac{z-x}{1-\theta_s}\right\}\right) \exp\left\{-\frac{x^2}{2\theta_s^2}\right\} dx \\
&= \Phi\left(\frac{z}{\theta_s}\right) - \frac{1}{2} \exp\left\{-\frac{z}{1-\theta_s}\right\} \exp\left\{\frac{1}{2}\left(\frac{\theta_s}{1-\theta_s}\right)^2\right\} \Phi\left(\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right)
\end{aligned}$$

and

$$\begin{aligned}
B &= \int_z^{\infty} \int_{-\infty}^{z-x} \frac{1}{2(1-\theta_s)} \exp\left\{-\frac{|y|}{1-\theta_s}\right\} dy \frac{1}{\sqrt{2\pi\theta_s}} \exp\left\{-\frac{x^2}{2\theta_s^2}\right\} dx \\
&= \frac{1}{2} \int_z^{\infty} \frac{1}{\sqrt{2\pi\theta_s}} \exp\left\{\frac{z-x}{1-\theta_s}\right\} \exp\left\{-\frac{x^2}{2\theta_s^2}\right\} dx \\
&= \frac{1}{2} \exp\left\{\frac{z}{1-\theta_s}\right\} \exp\left\{\frac{1}{2}\left(\frac{\theta_s}{1-\theta_s}\right)^2\right\} \Phi\left(-\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
F_Z(z) &= \Phi\left(\frac{z}{\theta_s}\right) + \frac{1}{2} \exp\left\{\frac{1}{2}\left(\frac{\theta_s}{1-\theta_s}\right)^2\right\} \\
&\quad \times \left[ \exp\left\{\frac{z}{1-\theta_s}\right\} \Phi\left(-\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) - \exp\left\{-\frac{z}{1-\theta_s}\right\} \Phi\left(\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) \right].
\end{aligned}$$

The density function of  $Z$  is obtained by differentiating its cdf, i.e.,

$$\begin{aligned}
f_Z(z) &= \frac{dF_Z(z)}{dz} = \frac{1}{\theta_s} \phi\left(\frac{z}{\theta_s}\right) + \frac{1}{2} \exp\left\{\frac{1}{2}\left(\frac{\theta_s}{1-\theta_s}\right)^2\right\} \\
&\quad \times \left[ \frac{1}{1-\theta_s} \exp\left\{\frac{z}{1-\theta_s}\right\} \Phi\left(-\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) - \frac{1}{\theta_s} \exp\left\{\frac{z}{1-\theta_s}\right\} \phi\left(-\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) \right. \\
&\quad \left. + \frac{1}{1-\theta_s} \exp\left\{-\frac{z}{1-\theta_s}\right\} \Phi\left(\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) - \frac{1}{\theta_s} \exp\left\{-\frac{z}{1-\theta_s}\right\} \phi\left(\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) \right] \\
&= \frac{1}{2(1-\theta_s)} \exp\left\{\frac{1}{2}\left(\frac{\theta_s}{1-\theta_s}\right)^2\right\} \\
&\quad \times \left[ \exp\left\{\frac{z}{1-\theta_s}\right\} \Phi\left(-\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) + \exp\left\{-\frac{z}{1-\theta_s}\right\} \Phi\left(\frac{z}{\theta_s} - \frac{\theta_s}{1-\theta_s}\right) \right]. \quad (23)
\end{aligned}$$

Next, we derive the moments of  $Z$ . Using the binomial theorem,  $(a+b)^n = \sum_{i=0}^n \binom{n}{i} a^i b^{n-i}$ , we obtain

$$\mathbb{E}(Z^m) = \mathbb{E}[(\theta_s N + (1-\theta_s)L)^m] = \sum_{i=0}^m \binom{m}{i} \theta_s^i (1-\theta_s)^{m-i} \mathbb{E}(N^i) \mathbb{E}(L^{m-i}).$$

If  $m$  is odd, either  $i$  is odd or  $m - i$  is odd, and so, as all odd moments of  $N$  and  $L$  are zero, the odd moments of  $Z$  are all zero. For the even moments, we use the fact that  $E(N^{2i}) = 2^{-i}(2i)!/i!$  and  $E(L^{2i}) = (2i)!$ . This yields

$$\begin{aligned}
E(Z^m) &= E[(\theta_s N + (1 - \theta_s)L)^m] = \sum_{i=0}^m \binom{m}{i} \theta_s^i (1 - \theta_s)^{m-i} E(N^i) E(L^{m-i}) \\
&= \sum_{i=0}^{m/2} \binom{m}{2i} \theta_s^{2i} (1 - \theta_s)^{m-2i} E(N^{2i}) E(L^{m-2i}) \\
&= (1 - \theta_s)^m \sum_{i=0}^{m/2} \left( \frac{\theta_s}{1 - \theta_s} \right)^{2i} \frac{m!}{(2i)!(m-2i)!} \frac{(2i)!}{2^i i!} (m-2i)! \\
&= (1 - \theta_s)^m m! \sum_{i=0}^{m/2} \frac{1}{i!} \left( \frac{\theta_s^2}{2(1 - \theta_s)^2} \right)^i \\
&= (1 - \theta_s)^m m! e_{m/2} \left( \frac{\theta_s^2}{2(1 - \theta_s)^2} \right),
\end{aligned}$$

as claimed in (13).

Using the relation (Abramowitz and Stegun, 1974, p. 262)

$$e_n(x) = e^x [1 - \Gamma^{\text{inc}}(n+1, x)],$$

where

$$\Gamma^{\text{inc}}(n, x) = [\Gamma(n)]^{-1} \int_0^x t^{n-1} e^{-t} dt \quad (24)$$

is the incomplete gamma function, it is straightforward to show that

$$E(Z^m) = (1 - \theta_s)^m \Gamma(m+1) \exp \left\{ \frac{\theta_s^2}{2(1 - \theta_s)^2} \right\} \left[ 1 - \Gamma^{\text{inc}} \left( \frac{m}{2} + 1; \frac{\theta_s^2}{2(1 - \theta_s)^2} \right) \right], \quad (25)$$

the computation of which, given the availability of  $\Gamma^{\text{inc}}(n, x)$  in most numerical software packages, represents no practical problem. For computation purposes, it may be even more useful to use the fact that  $E(Z^m)$  satisfies the difference equation

$$E(Z^m) = (1 - \theta_s)^2 m(m-1) E(Z^{m-2}) + \theta_s^m \frac{\Gamma(m+1)}{2^{m/2} \Gamma(m/2+1)}, \quad m \geq 4 \text{ and even}, \quad (26)$$

which is easily affirmed through evaluation of  $E(Z^m) - E(Z^{m-2})$ . Relation (26) is particularly appealing because the first term on the right hand side, apart from the factor  $(1 - \theta_s)^2$  measuring the weight of the Laplace, defines the updating scheme for the moments of the Laplace distribution, while the second term is exactly the  $m^{\text{th}}$  moment of the Gaussian distribution, weighted by  $\theta_s^m$ .



Interestingly, if the proportion of normality,  $\theta_s$ , is small,  $[0.5\theta_s^2/(1-\theta_s)^2]^i$  approaches zero very fast as  $i$  increases, implying that the exponential sum function in (13) can be approximated by  $\exp\{0.5\theta_s^2/(1-\theta_s)^2\}$ . As such, a useful approximation to  $E(Z^m)$  for even  $m$  and small  $\theta_s$  is given by

$$E(Z^m) \simeq (1-\theta_s)^m m! \exp\left\{\frac{1}{2}\left(\frac{\theta_s}{1-\theta_s}\right)^2\right\}. \quad (27)$$

Table 7, which compares expressions (13) and (27) for selected values of  $\theta_s$  and  $m$ , shows that the accuracy of (27) is quite acceptable for small values of  $\theta_s$ .

Table 7: Moments of the Gauss–Laplace Sum Computed Exactly and by Approximation<sup>a</sup>

$\theta_s$	E( $Z^2$ )		E( $Z^4$ )		E( $Z^6$ )	
	Exact	Appr.	Exact	Appr.	Exact	Appr.
0.1	1.630	1.630	15.84	15.84	385.0	385.0
0.3	1.070	1.074	6.316	6.317	92.85	92.85
0.5	0.750	0.824	2.438	2.473	18.52	18.55
0.7	0.670	2.739	1.444	2.958	5.663	7.986

<sup>a</sup>“Exact” refers to the respective moments calculated using (13) and “Appr.” refers to use of approximation (27).

## References

- Abramowitz, M. and Stegun, I. A. (1974). *Handbook of Mathematical Functions*. Dover Publications, New York.
- Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In Petrov, B. and Csaki, F., editors, *2nd International Symposium on Information Theory*, pages 267–281, Akademiai Kiado, Budapest.
- Alzer, H. (1997). On Some Inequalities for the Gamma and Psi Functions. *Mathematics of Computation*, 66:373–389.
- Andrews, D. F. and Mallows, C. L. (1974). Scale Mixtures of Normal Distributions. *Journal of the Royal Statistical Society, Series B*, 36(1):99–102.
- Bai, X., Russell, J. R., and Tiao, G. C. (2003). Kurtosis of GARCH and Stochastic Volatility Models with Non-normal Innovations. *Journal of Econometrics*, 114:349–360.
- Berkowitz, J. (2001). Testing Density Forecasts, with Applications to Risk Management. *Journal of Business and Economic Statistics*, 19:465–474.
- Bibby, B. M. and Sørensen, M. (2003). Hyperbolic Processes in Finance. In Rachev, S. T., editor, *Handbook of Heavy Tailed Distributions in Finance*. Elsevier Science, Amsterdam.
- Clark, P. K. (1973). A Subordinated Stochastic Process Model With Finite Variance for Speculative Prices. *Econometrica*, 41:135–155.
- Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998). Evaluating Density Forecasts with Applications to Financial Risk Management. *International Economic Review*, 39:863–883.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1006.
- Epps, T. W. and Epps, M. L. (1976). The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis. *Econometrica*, 44:305–321.
- Fernandez, C., Osiewalski, J., and Steel, M. F. (1995). Modelling and Inference With V-Spherical Distributions. *Journal of the American Statistical Association*, 90:1331–1340.
- Gouriéroux, C. (1997). *ARCH Models and Financial Applications*. Springer-Verlag, New York.
- Granger, C. W. J. and Ding, Z. (1995). Some Properties of Absolute Return, An Alternative Measure of Risk. *Annales D'économie et de Statistique*, 40:67–91.
- Haas, M., Mittnik, S., and Paoletta, M. S. (2004a). Mixed Normal Conditional Heteroskedasticity. *Journal of Financial Econometrics*, 2(2):211–250.
- Haas, M., Mittnik, S., and Paoletta, M. S. (2004b). A New Approach to Markov Switching GARCH Models. *Journal of Financial Econometrics*, 2(4):493–530.

- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions, Vol. 2*. John Wiley & Sons, New York.
- Jones, P. N. and McLachlan, G. J. (1990). Laplace–Normal Mixtures Fitted to Wind Shear Data. *Journal of Applied Statistics*, 17:271–276.
- Kanji, G. K. (1985). A Mixture Model for Wind Shear Data. *Journal of Applied Statistics*, 12:49–58.
- Kass, R. E. and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90:773–795.
- Knight, J. and Satchell, S. (eds.) (2001). *Return Distribution in Finance*. Butterworth Heinemann, Oxford.
- Kotz, S., Podgorski, K., and Kozubowski, T. (2001). *The Laplace Distribution and Generalizations: A Revisit with Application to Communication, Economics, Engineering and Finance*. Birkhäuser, Boston.
- Kuechler, U., Neumann, K., Soerensen, M., and Streller, A. (1999). Stock Returns and Hyperbolic Distributions. *Mathematical and Computer Modelling*, 29:1–15.
- Liesenfeld, R. (2001). A Generalized Bivariate Mixture Model for Stock Price Volatility and Trading Volume. *Journal of Econometrics*, 104:141–178.
- Linden, M. (2001). A Model for Stock Return Distribution. *International Journal of Finance and Economics*, 6:159–169.
- McDonald, J. B. (1996). Probability Distributions for Financial Models. In Maddala, G. S. and Rao, C. R., editors, *Handbook of Statistics: Statistical Methods in Finance*, volume 14, pages 427–461. Elsevier Science, Amsterdam.
- Mittnik, S. and Paoella, M. S. (2000). Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates. *Journal of Forecasting*, 19:313–333.
- Mittnik, S., Paoella, M. S., and Rachev, S. T. (1998). Unconditional and Conditional Distributional Models for the Nikkei Index. *Asia-Pacific Financial Markets*, 5(2):99–128.
- Mittnik, S. and Rachev, S. (1993). Modeling Asset Returns With Alternative Stable Models. *Econometric Reviews*, 12:261–330.
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59:347–370.
- Palm, F. C. (1997). GARCH Models of Volatility. In Maddala, G. and Rao, C., editors, *Handbook of Statistics*, volume 14. Elsevier Science.
- Palm, F. C. and Vlaar, P. J. G. (1997). Simple Diagnostic Procedures for Modeling Financial Time Series. *Allgemeines Statistisches Archiv*, 81:85–101.
- Rachev, S. T. (2003). (editor) *Handbook of Heavy Tailed Distributions in Finance*. North-Holland.

- Reimann, S. (2005). On the Distribution of Stock–Market Returns: Implications of Evolutionary Finance. Working paper, Institute for Empirical Research in Economics, University of Zurich.
- Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *Annals of Mathematical Statistics*, 23:470–472.
- Schoutens, W. (2003). *Levy Processes in Finance : Pricing Financial Derivatives*. John Wiley & Sons, New York.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6:461–464.
- Stuart, A., Ord, J. K., and Arnold, S. (1999). *Kendall’s Advanced Theory of Statistics*, volume 2A. Edward Arnold, London.
- Theissen, E. (2003). Organized Equity Markets in Germany. CFS Working Paper 2003/17, Center for Financial Studies.

## CFS Working Paper Series:

No.	Author(s)	Title
2005/02	<b>Torben G. Andersen</b> <b>Tim Bollerslev</b> <b>Peter F. Christoffersen</b> <b>Francis X. Diebold</b>	Practical Volatility and Correlation Modeling for Financial Market Risk Management
2005/03	<b>Francis X. Diebold</b> <b>Monika Piazzesi</b> <b>Glenn D. Rudebusch</b>	Modeling Bond Yields in Finance and Macroeconomics
2005/04	<b>Torben G. Andersen</b> <b>Tim Bollerslev</b> <b>Francis X. Diebold</b> <b>Jin (Ginger) Wu</b>	A Framework for Exploring the Macroeconomic Determinants of Systematic Risk
2005/05	<b>Jan Pieter Krahen</b>	Der Handel von Kreditrisiken: Eine neue Dimension des Kapitalmarktes
2005/06	<b>Günter Franke</b> <b>Jan Pieter Krahen</b>	Default Risk Sharing Between Banks and Markets: The Contribution of Collateralized Debt Obligations
2005/07	<b>Dirk Krueger</b> <b>Harald Uhlig</b>	Competitive Risk Sharing Contracts with One-Sided Commitment
2005/08	<b>Torben G. Andersen</b> <b>Tim Bollerslev</b> <b>Peter F. Christoffersen</b> <b>Francis X. Diebold</b>	Volatility Forecasting
2005/09	<b>Markus Haas</b> <b>Stefan Mittnik</b> <b>Bruce Mizrach</b>	Assessing Central Bank Credibility During the ERM Crises: Comparing Option and Spot Market-Based Forecasts
2005/10	<b>Juan Carlo Conesa</b> <b>Dirk Krueger</b>	On the Optimal Progressivity of the Income Tax Code
2005/11	<b>Markus Haas</b> <b>Stefan Mittnik</b> <b>Marc S. Paolella</b>	Modeling and Predicting Market Risk With Laplace-Gaussian Mixture Distributions