Kredit und Kapital, 45. Jahrgang, Heft 2, Seiten 175–187
Abhandlungen

Risk Measurement with a Safety Belt: Pareto Meets Chebyshev

By Karl-Heinz Tödter, Frankfurt/M.*

I. Introduction

Even though *Bernstein* (1996) had declared the mastery of risks, we recently witnessed the materialization of huge environmental (Deepwater Horizon in 2010), nuclear (Fukushima in 2011) and financial (Lehman Brothers in 2008) disasters. Underestimation of risks, often paired with overconfidence in grounded procedures, can trigger severe crises and impose large losses on societies, raising the question: "If everyone agrees that extreme events occur rather more frequently than we might like, why don't we take more cognisance of the possibility of such events?"[1] In this article, a "worst-possible" distribution is defined that is resilient to departures from normality and safeguards against extreme outcomes. Specifically, combining the Pareto law with finite variance bounds of Chebyshev results in a density that captures the tail behaviour of any random variable with unknown distribution, provided its variance exists. In Section II the Pareto-Chebyshev distribution is outlined and in Section III it is applied to the measurement of forecast uncertainty for GDP growth forecasts for the U.S. and Germany. In Section IV conservative measures of value at risk and expected shortfall are applied to the DOW Jones and DAX stock market indices. Section V concludes.

II. Pareto Meets Chebyshev

In order to model complex phenomena, two families of distributions play a leading role, normal (Gaussian) distributions and power law (Pareto) distributions. The importance of the "thin tailed" normal distribu-

* The views expressed in this article do not necessarily comply with those of the Deutsche Bundesbank. I thank Günter Bamberg, Uwe Hassler, Malte Knüppel, Guido Schultefrankenfeld and an anonymous referee for their helpful comments.

[1] *Kemp* (2011), p. 7.

tion can mainly be attributed to the central limit theorem, which assumes that many small and independent effects contribute additively to each observation.[2] However, empirical distributions often exhibit power law behaviour with fat tails. Power laws are scale invariant and appear as a straight line in a log-log plot.[3] However, there is no clear boundary between thin and fat tails. Fat tails is a "technical term for the annoying fact that in reality the 1000-year flood comes about every ten years."[4] A range of densities exists with tails heavier than those of the normal with finite variance, such as the Laplace distribution, or with infinite variance, such as Lévy-type α-stable distributions. *Mandelbrot* (1963) provided early evidence that infinite variance stable models better fit certain financial variables than Gaussian models.

Consider the cumulative power law distribution of a *Pareto* (1897) random variable X, $G(x) = 1 - (x_{\min}/x)^a$, where $x_{min} \leq x$ is the positive minimum value of X and $a > 0$ is the tail index. The expectation (variance) does not exist if $a \leq 1$ $(a \leq 2)$. We seek a symmetric distribution with tails *at least as fat* (but not fatter) as those of any random variable with existing variance but otherwise unknown density $f(x)$.[5] Tails are defined here as the part of a density which is at least one standard deviation away from its mean. Consider the following density function:[6]

$$\varphi(x) = \begin{cases} \dfrac{\sigma^2}{\left|x-\mu\right|^3}, & \left|x-\mu\right| \geq \sigma \\ 0, & \left|x-\mu\right| < \sigma\,. \end{cases} \tag{1}$$

The density φ(x) is closely related to a symmetric Pareto distribution with tail index at the boundary of infinite variance $(a = 2)$.[7] The expec-

[2] The Gaussian distribution can also be justified as the most parsimonious choice, absent any information other than the mean and variance: it maximizes the information entropy among all distributions with known mean and variance; see *Jaynes* (2003).

[3] Extreme value theory (EVT) shows that power laws are the limiting behavior of large events for a wide class of probability distributions, providing a rationale for the widespread observation of power laws; see *Coles* (2004), *Alfarano/Lux* (2010).

[4] *Putnam/Wilford/Zecher* (2002), p. 207.

[5] Finite variance implies finite expectation of a distribution; see *Hogg/McKean/Craig* (2005), p. 69.

[6] In the context of stochastically ordered random variables, density (1) was given by *Hürlimann* (2008), p. 170.

[7] The symmetric Pareto distribution is discussed in *Grabchak/Samorodnitsky* (2010).

tation is μ, its maxima $(1/\sigma)$ are located at $x = \mu \pm \sigma$, and the density approaches zero as x tends to $\pm\infty$. Thus, the parameter σ determines the range where the density $\varphi(x)$ is positive, it is *not* the variance implied by $\varphi(x)$. Actually, the variance does not exist: $\int_{-\infty}^{\mu-\sigma} (x-\mu)^2 \varphi(x) dx + \int_{\mu+\sigma}^{\infty} (x-\mu)^2 \varphi(x) dx = \infty$.[8] Infinite variance is required because if (1) is to capture extreme outcomes of any random variable with finite but arbitrarily large variance, it cannot have finite variance. The distribution function implied by (1) is

$$\Phi(x) = \frac{1}{2}\begin{cases} \dfrac{\sigma^2}{(x-\mu)^2}, & -\infty \leq x \leq \mu - \sigma \\ 1, & \mu - \sigma < x < \mu + \sigma \\ 2 - \dfrac{\sigma^2}{(x-\mu)^2} & \mu + \sigma \leq x \leq \infty\,. \end{cases} \tag{2}$$

The *tail probability* of realisations outside the interval $(\mu - k\sigma\,,\ \mu + k\sigma)$ is

$$P(|X-\mu| \geq k\sigma) = \int_{-\infty}^{\mu-k\sigma} \varphi(x)dx + \int_{\mu+k\sigma}^{\infty} \varphi(x)dx = \frac{1}{k^2}\,, \quad k \geq 1. \tag{3}$$

This exactly matches the bounds of the *Chebyshev* (1867) inequality.[9] It is often claimed that no distributional assumption is made when Chebyshev's inequality is applied,[10] but (3) shows that a symmetric Pareto distribution with tail index $a = 2$ underlies Chebyshev's inequality. We refer to (1) as Pareto-Chebyshev (PaCh(μ,σ)) density hereafter. Although the following results can be derived directly from *Chebyshev's* inequality, it is convenient to have an explicit density function to work with.[11] Figure 1 shows the PaCh(0,1)-density, together with the Laplace$(0,1/\sqrt{2})$- and the Gaussian N(0,1)-density.[12]

---

[8] The higher even moments do not exist either, while the higher uneven moments are undefined. Yet, the mean absolute deviation is finite, assuming the value 2σ.

[9] See *Lindgren* (1968), *Kazmier* (1979).

[10] See, for example, *Alexander/Baptista* (2001), p. 1162.

[11] The density can be used, for example, in Monte Carlo experiments to generate synthetic random variables with a PaCh(μ,σ) distribution from $x = \mu - \sigma/\sqrt{2U}$ $(U \leq 1/2)$ and $x = \mu + \sigma/\sqrt{2U}$ $(U > 1/2)$, where U is a standard uniform random variable.

[12] The Laplace distribution $h(x;\mu,\upsilon) = (2\upsilon)^{-1}\exp(-|x-\mu|/\upsilon)$ has mean μ, variance $2\upsilon^2$ and kurtosis 6.

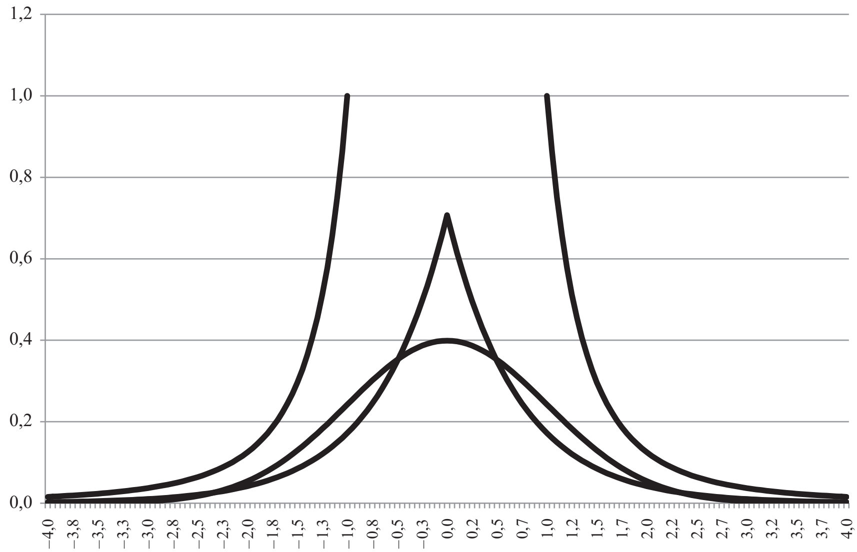

Figure 1: PaCh, Laplace and Gaussian Density

For the PaCh distribution the following bounds $k(\alpha)$ satisfying $P(\left|X-\mu\right| \geq k\sigma) = \alpha$ are obtained from (3) as a function of the tail probability α:[13]

$$k(\alpha) = \frac{1}{\sqrt{\alpha}}. \tag{4}$$

These bounds can be used to compute 1-α confidence bands for random variables with unknown distribution, requiring knowledge of μ and σ only. Compared to Gaussian confidence intervals, the PaCh intervals are substantially wider and provide a well-defined "safety belt" that accounts for extreme outcomes. Table 1 shows the bounds $k(\alpha)$ for various choices of α. For comparison, bounds for the Laplace and the Gaussian distribution are provided as well, parameterized such that both have variance 1. Thus, under the PaCh density, a 99 % confidence interval around the mean is about 3.9 (3.1) as wide as in the Gaussian (Laplace) case. For $\alpha < 0.093$ the bounds $k(\alpha)$ of the Laplace distribution are greater than those of the Gaussian distribution, reflecting the heavier tails of the former.

[13] For the Laplace distribution, the bounds for the tail-probability $P(|x-\mu|) \geq k\sqrt{2v^2} = \alpha$ are $k(\alpha) = -\ln(\alpha)/\sqrt{2}$.

Table 1

Bounds (k) for Tail Probabilities α

α	0.50	0.25	0.20	0.10	0.05	0.02	0.01	0.001
Gauss (μ,σ)	0.67	1.15	1.28	1.64	1.96	2.33	2.58	3.29
Laplace (μ,υ)	0.49	0.98	1.14	1.63	2.12	2.77	3.26	4.88
PaCh (μ,σ)	1.41	2.00	2.24	3.16	4.47	7.07	10.00	31.62

$\mu = 0$, $\sigma^2 = 1$, $\upsilon^2 = 1/2$. The left (right) tail probabilities are $\alpha/2$.

III. Application to Forecast Uncertainty

Many central banks and other institutional forecasters publish their point estimates along with measures of forecast uncertainty.[14] In the wake of the global financial crisis, GDP fell markedly in many countries. In the U.S. and in Germany the slump was most severe in 2008Q4 and 2009Q1. Given the volatility of growth rates observed in the past, such events are extremely unlikely under the normal distribution, leading to an excessive number of violations of Gaussian forecast intervals.

Empirically, we consider forecasts of annual growth rates ($\hat{x}_t$) of real GDP for the U.S. and Germany. The forecasts are based on seasonally adjusted quarterly growth rates (q_t), they have a horizon of 4 quarters and are calculated not just for calendar years but after each quarter.[15] The annual growth forecast made in quarter t is computed as

$$\hat{x}_t = \frac{1}{4}(q_{t-6} + 2q_{t-5} + 3q_{t-4}) + \frac{1}{4}10\,\ddot{q}_{t-4}. \tag{5}$$

The first part on the right hand side of (5) is the carry-over effect, which is already known at the time of forecasting. The second part estimates the unknown quantity $(1/4)\,[4q_{t-3} + 3q_{t-2} + 2q_{t-1} + q_t]$, where $\ddot{q}_{t-4}$ in (5) is the arithmetic mean of quarterly observations up to quarter $t-4$. The estimation window for the means starts in 1992Q1 and successively widens until 2007Q4 (pre-crisis) and 2010Q4 (crisis included), re-

[14] See *Deutsche Bundesbank* (2010) for an overview.

[15] The annual growth rate in quarter t (x_t) can be approximated as a weighted sum of quarterly growth rates (q_t): $x_t = (1/4)(q_{t-6} + 2q_{t-5} + 3q_{t-4} + 4q_{t-3} + 3q_{t-2} + 2q_{t-1} + q_t)$; see *Tödter* (2011).

Table 2

Percentage of Observed Violations (p) of Forecast Confidence Bands for GDP Growth

Conf. band coverage	50 %			95 %		
Distribution	Gauss	Laplace	PaCh	Gauss	Laplace	PaCh
U.S., pre-crisis	65.0**	73.3***	33.3	13.3***	11.7**	0
U.S., with crisis	66.7***	75.0***	36.1	19.4***	18.1***	4.2
Germany, pre-crisis	55.0	65.0**	11.7	5.0	5.0	0
Germany, with crisis	56.9	68.1***	16.7	9.7*	9.7*	2.8

Based on $T = 60$ observations (pre-crisis) and $T = 72$ observations (with crisis).
(*, **, ***): significant at the $(10, 5, 1)$ % level.
Tests of the null hypothesis $p = p_o$ against the one sided alternative $p > p_o$ with $p_o = (0.5, 0.05)$ are based on the Binomial distribution $f(x) = (T!/x!(T-x)!)p_o^x(1-p_o)^{T-x}$ and $x = Tp$.

sulting in $T = 60$ and $T = 72$ forecasts, respectively. Associated confidence intervals are calculated as $\hat{x}_t \pm k(\alpha)\,\hat{\sigma}_{\hat{x}_t}$, where $\hat{\sigma}_{\hat{x}_t}$ is the estimated variance based on quarterly observations up to quarter $t-4$, again applying a widening window and taking into account the information of the carry-over effect. The bounds $k(\alpha)$ for $\alpha = [50,\ 95]$ %-confidence intervals are shown in Table 1. One should expect to observe $[50, 5]$ % violations for the Gaussian and the Laplace intervals and at most $[50, 5]$ % violations for the PaCh distribution. As Table 2 shows, the confidence bands for the U.S. forecasts and the German forecasts are violated significantly more often than they should. This is the case for both, the Gauss and the Laplace distribution. Thus, confidence intervals based on these distributions are too narrow, even when the crisis period is not included in the sample. In contrast, the number of violations of the PaCh confidence intervals remains well below the expected limits in all cases.

IV. Application to Risk Measurement

Value at risk (VaR) is a widely used risk measure in financial institutions which has also made its way into the Basel II capital-adequacy framework.[16] We follow the convention and present results for the right (or upper) tail of distributions. Let X be a variable which measures the loss of a portfolio of assets over a given horizon. Given a confidence level

[16] See *McNeil/Frey/Embrechts* (2005), p. 37.

$\alpha \in (0,1)$, VaR solves the equation $P(x \geq VaR(\alpha)) = 1 - \alpha$. Thus, VaR($\alpha$) is the α-quantile of the loss distribution, that is, the maximum loss not exceeded with probability α. For the PaCh density (1) the following simple expression is obtained:[17]

$$VaR(\alpha) = \mu + \frac{\sigma}{\sqrt{2(1-\alpha)}}. \tag{6}$$

This value can be interpreted as the worst loss that could occur within the class of loss distributions with finite variance.[18] For regulatory purposes the Basel Committee imposes a 99 percent confidence level over a 10 business day horizon for the VaR in the internal models approach. Then, the resulting VaR is multiplied by the safety factor λ to provide a minimum capital requirement (CR):

$$CR(\alpha) = c + \lambda \, VaR(\alpha). \tag{7}$$

If banks underestimate the VaR they may be penalized by the additive constant (c) or by an increase of the multiplicative factor λ from 3 to at most 4.[19]

A weakness of the VaR measure is that it does not take into account the magnitude of losses beyond the VaR.[20] This is accounted for by the expected shortfall (or conditional value at risk), which is defined as $ES(\alpha) = \int_{VaR(\alpha)}^{\infty} x\tilde{f}(x)dx$, where $\tilde{f}(x) = f(x) / \int_{VaR(\alpha)}^{\infty} f(x)dx \; = f(x)/(1-\alpha)$ is the re-normalized density $f(x)$. For the PaCh density (1), expected shortfall is:

[17] Variants of the density (1) have been used for estimating VaR and Expected Shortfall by *Hürlimann* (2002). *Alexander/Baptista* (2002), *Putnam/Wilford/Zecher* (2002), *ElGhaoui/Oks/Oustry* (2003) have used *Chebyshev's* inequality in estimating VaR.

[18] Hence, (6) captures uncertainty about the loss distribution. See *Simonian/Davis* (2010) for a robust VaR measure that takes into account model misspecification.

[19] See *Jorion* (2001), p. 119. The regular scaling factor $\lambda = 3$ is about the same $(7.07/2.33 = 3.03)$ that results from application of the PaCh density (see Table 1) for measuring risk. It was noted by *Stahl* (1997) that a safety factor of 3 could be justified by *Chebyshev's* inequality. However, this is true only for the $\alpha = 99\,\%$ level.

[20] Moreover, the value at risk is not a coherent risk measure as it lacks subadditivity, i.e. the risk of two different portfolios can exceed the sum of the individual risks. On coherent risk measures, see *McNeil/Frey/Embrechts* (2005), p. 241 and *Jorion* (2001), p. 115.

Table 3

Value at Risk and Expected Shortfall as Functions of Confidence Level α

Distribution	Value at Risk	Expected Shortfall
PaCh	$VaR_{PaCh} = \mu + \sigma/\sqrt{2(1-\alpha)}$	$ES_{PaCh} = 2\,VaR_{PaCh} - \mu$
Laplace	$VaR_{Lapl} = \mu - \upsilon \ln(2(1-\alpha))$	$ES_{Lapl} = VaR_{Lapl} + \upsilon$
Gauss	$VaR_N = \mu + \sigma\Phi_N^{-1}(\alpha)$	$ES_N = \mu + \sigma\,\varphi_N\left[\Phi_N^{-1}(\alpha)\right]/(1-\alpha)$

$\upsilon^2 = \sigma^2/2$, Φ_N (φ_N) denotes the standard normal distribution (density) function and $\Phi_N^{-1}(\alpha)$ is the α-quantile of Φ_N.

$$(8) \qquad ES(\alpha) = \mu + \frac{2\sigma}{\sqrt{2(1-\alpha)}} = 2\,VaR(\alpha) - \mu.$$

Table 3 provides formulas for both risk measures, VaR and ES, for the PaCh, the Laplace and the Gaussian density.

We computed VaRs for daily data of two stock market indices, the U.S. DOW Jones Industrial Average Index and the German DAX Index. Let z_t denote the index values, then daily losses (negative returns) are defined as $r_t = -(z_t - z_{t-1})/z_{t-1}$. To account for time-varying volatility, mean losses ($\ddot{r}_t$) and variances (s_t^2) are obtained by averaging over a moving window of the preceding 100 trading days, i.e. $\ddot{r}_t = (1/100)\sum_{\tau=t-100}^{t-1} r_\tau$ and $s_t^2 = (1/100)\sum_{\tau=t-100}^{t-1}(r_\tau - \ddot{r}_t)^2$. Value at risk for a one day horizon for the PaCh density was then obtained from equation (6) with $\ddot{r}_t$ (s_t) estimating μ (σ) and for $\alpha = [90, 95, 99]\,\%$. In a similar way value at risk was computed using quantiles from the Gaussian and the Laplace distribution.[21] Finally, over 3 two-year-periods (2005–06, 2007–08, 2009–10), each including about $T = 520$ trading days, we counted how often the realized daily losses (r_t) exceeded the corresponding value at risk ($VaR_t(\alpha)$).

Table 4 shows the observed percentages of VaR-exceedances. Not surprisingly, most VaR-exceedances occurred during the crisis years 2007–08. For both, the Gaussian and the Laplace distribution, the number of violations is significantly and substantially larger than expected. This is the case at all three confidence levels and for both stock market indices,

[21] In the latter case, the parameter υ^2 was estimated by $s_t^2/2$.

Table 4

Percentage (p) of VaR-Exceedances

α		90 %			95 %			99 %		
	T	Gauss	Lapl	PaCh	Gauss	Lapl	PaCh	Gauss	Lapl	PaCh
DOW, 2005–06	*520*	10.4	12.7 **	1.7	5.8	5.8	0.4	1.3	1.2	0
DOW, 2007–08	*523*	14.3 ***	17.2 ***	4.6	10.3 ***	10.5 ***	1.5	4.4 ***	1.9 **	0.2
DOW, 2009–10	*522*	6.5	8.6	1.9	3.3	3.3	0.4	1.9 **	0.8	0
DAX, 2005–06	*520*	11.3	14.0 ***	3.1	6.5 *	6.5 *	0.6	2.7 ***	1.3	0
DAX, 2007–08	*523*	12.8 **	15.5 ***	4.0	8.4 ***	8.8 ***	1.3	3.4 ***	1.9 **	0.2
DAX, 2009–10	*522*	8.2	10.0	1.3	4.0	4.2	0.2	1.3	0.4	0

Source: Datastream. (*, **, ***): significant at the (10, 5, 1) % level.
Tests of the null hypothesis $p = p_o$ against the one-sided alternative $p > p_o$ with $p_o = (0.1, 0.05, 0.01)$ are based on the normal approximation to the Binomial distribution (which is acceptable since $Tp > 5$), using the standard normal test statistic $(p - p_o)/\sqrt{p_o(1 - p_o)/T}$.

the DOW and the DAX. In contrast, the "safety belt" of the PaCh density holds out: In the non-crisis periods no VaR-violations occurred and in the crisis period only one VaR-violation (0.2 %) was recorded at $\alpha = 99\,\%$.

VaR-violations say nothing about the losses incurred in these cases. For a portfolio of $K = 1$ Bill € (or $, for that matter), the realized loss on trading day t is defined as $L_t = K \cdot r_t$, and $L = \sum_{t=1}^{T} L_t$ is the total loss (or profit, if negative) over the whole period of T trading days. Realized shortfall (RS) is the sum of losses that occurred on days with tail events, i.e. on days with VaR-exceedances: $RS = \sum_{t=1}^{T} L_t J_t$, where J_t is an indicator variable which assumes the value 1 if a VaR-violation occurred $\left(L_t \geq VaR_t(\alpha)\right)$ and 0 otherwise. The difference L-RS is the loss that would have been obtained if the tail events had not happened.

The results for the DOW and DAX stock market indices are shown in Table 5 for the $\alpha = 99\,\%$ level. As can be seen, tail events account for a large part of total losses in the Gaussian and, to a lesser extent, in the Laplace case. For example, for the DAX in 2009–10 total loss was –420 Mill € of which –625 Mill € were realized on days without VaR-exceedances and 205 Mill € occurred through tail events. In the Laplace case tail events account for losses of 60 Mill € and for the PaCh density no tail event was recorded.

Table 5

Realized and Expected Shortfall for DOW and DAX (Mill €)

α = 99%		Gauss			Laplace			PaCh		
			ES	*UES*		ES	*UES*		ES	*UES*
DOW, 2005–06	L	**–155**			**–155**			**–155**		
	RS	120	107	*13*	107	122	*–15*	0	0	*0*
	L-RS	–275			–262			–155		
DOW, 2007–08	L	**268**			**268**			**268**		
	RS	760	645	*115*	431	400	*31*	33	57	*–24*
	L-RS	–492			–163			235		
DOW, 2009–10	L	**–319**			**–319**			**–319**		
	RS	257	236	*21*	126	130	*–4*	0	0	*0*
	L-RS	–576			–445			–319		
DAX, 2005–06	L	**–458**			**–458**			**–458**		
	RS	309	283	*26*	170	178	*–8*	0	0	*0*
	L-RS	–767			–628			–458		
DAX, 2007–08	L	**231**			**231**			**231**		
	RS	773	632	*141*	507	447	*60*	72	127	*–55*
	L-RS	–542			–276			159		
DAX, 2009-10	L	**–420**			**–420**			**–420**		
	RS	205	200	*5*	60	67	*–7*	0	0	*0*
	L-RS	–625			–480			–420		

L is total loss (or profit, if negative), RS (ES) denotes realized (expected) shortfall, UES (= RS – ES) is unexpected shortfall.

Total expected shortfall (ES) is obtained by applying the formulas in Table 3 and summing over the trading days with VaR-exceedances. The difference (UES = RS – ES) is the unexpected shortfall. For both, the DOW and the DAX, RS exceeds ES if the Gaussian distribution is applied to calculate VaR in all three sub-periods. In the crisis period (2007–08) realized shortfall exceeds expected shortfall by 115 (141) Mill € for the DOW (DAX). For the Laplace distribution, in the non-crisis periods ES is only moderately underestimated, while in the crisis period RS exceeds ES by 31 (60) Mill €. For the PaCh distribution no VaR-violations occurred during the non-crisis periods such that RS, ES and UES are all zero. During the two crisis years only one VaR-violation occurred under the PaCh distribution. However, the associated RS was substantially smaller than the ES, for both, the DOW and the DAX.

As the historical results for the DOW and DAX stock market indices in Table 5 indicate, the Gaussian distribution underestimates losses beyond VaR even in normal times. In these periods the Laplace distribution performs better. However, in times of crisis both seriously underestimate expected shortfall. Due to the occurrence of extreme events, "that should not have happened for a thousand years",[22] the heavy-tailed PaCh distribution provides a safer basis of measuring risk.

What are the costs of financial conservatism? Consider, for example, the DAX index in the period 2009–10. For a capital of $K = 1$ Bill €, the average VaR at the $\alpha = 99\,\%$ level over a one day horizon was (38, 45, 117) Mill € under the Gaussian, Laplace and PaCh distribution, respectively. Applying the scaling factor $\lambda = 3$ yields a daily average capital requirement (CR) of (114, 136, 351) Mill € in that period. If the marginal cost of capital for a bank is set at 2 % p.a., which is approximately the average Euro area overnight index average rate (EONIA) in the period 2005–2010, the cost of capital requirement comes to (4.6, 5.5, 14.1) Mill € over the two years. Thus, the additional cost of using the conservative risk measure is relatively small compared to the benefits of avoiding large unexpected losses. Moreover, it could be argued that regulators should not apply the same scaling factor irrespective of the method banks calculate their VaR. If $\lambda = 3$ is applied in the Gaussian case at the 99 % level, $\lambda = 1$ would be justified when the heavy-tailed PaCh distribution is used to measure risk.

V. Conclusions

Risk measurement based on thin-tailed distributions as the Gaussian or the Laplace is likely to understate the probability of extreme events. The Pareto-Chebyshev (PaCh) density used in this article yields risk measures with a "safety belt". It captures fat tails of any random variables with unknown distribution and provides a prudent and simple approach to risk measurement. Risk managers do not know in advance when extreme events will strike such that timely switching risk manage-

[22] *Taleb* (2010) uses the term "Black Swans" as a metaphor for hard-to-predict rare events with a high impact, often being rationalized with hindsight.

ment from normal to "crisis mode" is not a viable option. Prudence requires to "strap the safety belt" all the time.

An empirical application to the measurement of uncertainty in annual growth forecasts of real GDP in the U.S. and Germany shows that forecast intervals based on the PaCh density capture the extreme events during the global financial crisis, while the narrower Gaussian and Laplace intervals are violated too often. For risk measurement using value at risk (VaR) and expected shortfall (ES), the PaCh density yields simple analytical formulas. Application to daily losses of the DOW Jones and DAX stock market indices shows that VaR based on the PaCh density well captures the extreme volatility during the 2007–08 crisis period, again in contrast to the corresponding Gaussian and Laplace measures. Moreover, tail events, i. e. realized losses in periods of VaR-violations, exceed ES by large amounts if based on the Gaussian and Laplace distributions, in contrast to measurement based on the PaCh distribution. In sum, the PaCh distribution is a robust tool that safeguards against extreme outcomes and complements the traditional risk measures used by forecasters and risk managers.

References

Alexander, G. J./*Baptista,* A. M. (2002): Economic Implications of Using a Mean-VaR Model Portfolio Selection: A Comparison with Mean-Variance Analysis, in: Journal of Economic Dynamics & Control, 26, 1159–1193. – *Alfarano,* S./*Lux,* T. (2010): Extreme Value Theory as a Theoretical Background for Power Law Behavior, Kiel Working Papers No. 1648. – *Bernstein,* P. L. (1996): Against the Gods: The Remarkable Story of Risk, John Wiley & Sons, Inc., New York. – *Chebyshev,* P. L. (1867): Des Valeurs Moyennes, in: Journal de Mathématiques Pures et Appliquées, 2(12), 177–184. – *Coles,* S. (2004): An Introduction to Statistical Modeling of Extreme Values, Springer, London. – *Deutsche Bundesbank* (2010): Uncertainty of Macroeconomic Forecasts, in: Monthly Report, June, 29–46. – *El Ghaoui,* L./*Oks,* M./*Oustry,* F. (2003): Worst-Case Value-at-Risk and Robust Portfolio Optimization: A Conic Programming Approach, in: Operations Research, 51(4), 543–556. – *Grabchak,* M./*Samorodnitsky,* G. (2010): Do Financial Returns have Finite or Infinite Variance? A Paradox and an Explanation, in: Quantitative Finance, 10, 883–893. – *Hogg,* R. V./*McKean,* J. W./*Craig,* T. (2005): Introduction to Mathematical Statistics, Pearson Prentice Hall, New Jersey, USA. – *Hürlimann,* W. (2002): Analytical Bounds for Two Value-at-Risk Functionals, in: Astin Bulletin, 32(2), 235–265. – *Hürlimann,* W. (2008): Extremal Moment Methods and Stochastic Order – Application in Actuarial Science, *Chapters iv, v and vi, Boletin de la Asociación Venetolana, XV,* 153–301. – *Jaynes,* E. T. (2003): Probability Theory: The Logic of Science, edited by G. Larry Bretthorst, Cambridge University Press, Cambridge. – *Jorion,* P. (2001): Value at Risk: The New Benchmark for Managing Financial Risk, 2nd.ed., McGraw-Hill, New York. – *Kazmier,* L. J. (1979): Basic Statistics for Business and Economics, McGraw-Hill, New York – *Kemp,* M. H. D. (2011): Extreme Events: Robust Portfolio Construction in the Presence of Fat Tails, John Wiley & Sons, Ltd., Chichester, UK. – *Lindgren,* B. W. (1968): Statistical Theory, 3rd ed., Macmillan Publishing Co., New York and London. – *Mandelbrot,* B. (1963): The Variation of Certain Speculative Prices, in: Journal of Business, 26, 394–419. – *McNeil,* A. J./*Frey,* R./*Embrechts,* P. (2005): Quantitative Risk Management; Concepts, Techniques and Tools, Princeton University Press, Princeton and Oxford. – *Pareto,* V. (1897): Cours d'economie politique, Vol. II, F. Rouge, Lausanne. – *Putnam,* B. H./*Wilford,* D. S./*Zecher,* P. D. (2002): A Short Note on the Concept of Risk Management and VaR for Asset Management Firms, in: Review of Financial Economics, 11 (205–212). – *Simonian,* J./*Davis,* J. (2010): Robust Value-at-Risk: An Information-theoretic approach, in: Applied Economics Letters, 17, 1551–1553. – *Stahl,* G. (1997): Three Cheers, Risk, 10, 67–69. – *Taleb,* N. N. (2010): The Black Swan, 2nd. ed., Penguin. – *Tödter,* K.-H. (2011): The Carry-Over Effect and its Value in Forecasting Annual Growth Rates, in: Review of Economics, 62(2), 160–171.

Summary

Risk Measurement with a Safety Belt: Pareto Meets Chebyshev

Risk measures based on the Gaussian distribution are prone to understate the probability of extreme events. To capture fat tails and extreme events, we combine the Pareto law with finite variance bounds of Chebyshev. This density encompasses the tail behaviour of a wide range of random variables with unknown distribution. It provides a well-defined conservative measure of risks. Applications to measurement of forecast uncertainty and to value at risk and expected shortfall illustrate the approach empirically. (JEL D81, C53, G10)

Zusammenfassung

Risikomessung mit Sicherheitsgurt: Pareto trifft Tschebyschow

Auf der Normalverteilung beruhende Risikomaße neigen dazu, die Wahrscheinlichkeit von extremen Ereignissen zu unterschätzen. Um extreme Ereignisse und dicke Enden von Verteilungen zu berücksichtigen, kombinieren wir die Pareto-Verteilung mit der Tschebyschow-Schranke für Zufallsvariablen mit endlicher Varianz. Diese Dichtefunktion umschließt das Verhalten einer großen Klasse von Zufallsvariablen und erlaubt eine wohldefinierte konservative Messung von Risiken. Anwendungen auf die Messung von Prognoseunsicherheit sowie den "Wert im Risiko" und den erwarteten Verlust im Risikofall illustrieren den Ansatz empirisch.