On time-scaling of risk and the square–root–of–time rule *

Jón Daníelsson and Jean-Pierre Zigrand

Department of Accounting and Finance and Financial Markets Group
London School of Economics
Houghton Street, London WC2A 2AE

July 26, 2005

Abstract

Many financial applications, such as risk analysis and derivatives pricing, depend on time scaling of risk. A common method for this purpose, though only correct when returns are iid normal, is the square–root–of–time rule where an estimated quantile of a return distribution is scaled to a lower frequency by the square-root of the time horizon. The aim of this paper is to examine time scaling of quantiles when returns follow a jump diffusion process. It is argued that a jump diffusion is well-suited for the modeling of systemic risk, which is the raison d'être of the Basel capital adequacy proposals. We demonstrate that the square–root–of–time rule leads to a systematic underestimation of risk, whereby the degree of underestimation worsens with the time horizon, the jump intensity and the confidence level. As a result, even if the square–root–of–time rule has widespread applications in the Basel Accords, it fails to address the objective of the Accords.

Journal of Economic Literature classification numbers: G18, G20, D81.
Keywords: Square-root-of time rule, time-scaling of risk, value-at-risk, systemic risk, risk regulation, jump diffusions.

*We have benefited from comments by Sudipto Bhattacharya, Casper de Vries, Antonio Mele and Andrew Patton, as well as by participants at the 2004 EFA meetings and by two anonymous referees; the remaining mistakes are of course our responsibility alone. Jean-Pierre Zigrand is the corresponding author: j.p.zigrand@lse.ac.uk, +44 020 79556201.

1 Introduction

The *square–root–of–time rule* is commonly assumed when financial risk is time aggregated whereby high frequency risk estimates are scaled to a lower frequency T by the multiplication of $\sqrt{T}$. One common application of the square–root–of–time rule is the time scaling of *volatilities*, such as in the Black–Scholes equation where the T–period volatility is given by $\sigma\sqrt{T}$. To take another example, a standard method for estimating *quantiles*, and in particular value–at–risk (VaR), is by estimating a one day VaR and multiplying it by $\sqrt{10}$, for the ten day regulatory requirement. Recall that VaR is the quantile that solves $\epsilon = \int_{-\infty}^{-\mathrm{VaR}} \widehat{f(x)}dx$, where $\widehat{f(x)}$ is the estimated probability density function of a financial institution's return and ϵ is the confidence level, say .01. Indeed, this is the method recommended by banking supervisors (see the Basel Committee on Banking Supervision, 1996), and is widely used throughout the financial industry. But VaR has not just found prominence via the external Basel regulations, it has effectively become a cornerstone of internal risk management systems in financial institutions following the success of the J.P. Morgan RiskMetrics system.

However, the time scaling of (conditional) *volatilities* requires returns to be conditionally homoscedastic and conditionally serially uncorrelated at all leads,[1] an assumption Engle (1982) argues is incorrect because of the presence of volatility clusters. This condition is only slightly weaker than requiring outright that returns are independently and identically distributed (iid). Also, see Müller et al. (1990) and Diebold et al. (1997) for detailed examples of how poor the approximation can be, and Groenendijk et al. (1998) for an analysis on the scaling of volatilities in large samples. This paper, however, is not concerned with the scaling of volatilities but with the scaling of quantiles. When applied to *quantiles*, much less is known about the square–root–of–time rule, other than that it obtains only in circumstances even more exceptional than the ones in which volatility scaling applies. If the square–root–of–time rule is required to be correct for all quantiles and horizons, it appears that it not only requires the iid property of zero-mean returns, but also normality of the returns. For instance, in continuous-time the time-scaling of quantiles is a consequence of self-similarity. The only Lévy process (i.e. process with independent and stationary increments) that is 1/2-similar is the Brownian Motion, i.e. the Gaussian Lévy process.[2] To our knowledge, no study has been undertaken in the academic literature that studies the quality

[1] Given any t, scaling of volatilities is equivalent to $\mathrm{Var}_t(r_{t+k}) = c_t$, for some constant c_t, all $k \geq 1$, and $\mathrm{Cov}_t(r_{t+k}, r_{t+l}) = 0$, all $k \geq 1$, $l \geq 1$, $k \neq l$.

[2] If we require square-root scaling only for extreme quantiles in the far tails, then returns that are i.i.d. and have tails which vary regularly at infinity (the so-called "fat tails"), i.e. $F(-x) = ax^{-\alpha}[1 + o(x^{-\alpha})]$ as $x \to \infty$, and $a > 0$, do scale like the square-root if $\alpha = 2$. This follows from Feller's proposition that $\mathbb{P}(\sum_{i=1}^{n} X_i < -x) \approx nax^{-\alpha}$, as $x \to \infty$, see e.g. Feller (1971, VIII.8). For instance a student t with 2 degrees of freedom would scale like the square-root in the tails. And conversely, for $\alpha \neq 2$ the quantiles do not scale like the square-root, not even in the far tails. In particular, no stable law other than the normal one can obey the square–root–of–time rule when

of approximation in the jump diffusion case.

The underlying distributional assumptions behind this extrapolation method to quantiles are quite stringent, and are violated in most, if not all, practical applications. If returns are not iid, for instance if the drift or the volatility is time-varying, or if returns are not normal (even if they are distributed according to a stable law), then time scaling of quantiles according to the square–root–of–time rule will no longer hold. In particular overall time-scaling fails for many popular processes, such as GARCH processes, stochastic volatility processes or jump processes.[3] Since the stylized facts for return distributions are well documented, the reason for the prevalence of the square–root–of–time rule must be a scarcity of robust alternative methods coupled with a lack of understanding of the shortcomings of the square–root–of–time rule.

Different biases arise from different data generating processes, dgps, except for the normal iid case with zero mean. In choosing between different dgps, one must consider the intended applications. In the pricing of barrier derivatives for instance, one may for instance want to use a GARCH, or a stable (non-normal) dgp, which would all result in violations of the square–root–of–time rule. The bias in using the square–root–of–time rule to scale quantiles in the fat-tailed (in particular the stable non-normal) iid case can be studied along the lines of Dacorogna et al. (2001). It is easy to see for instance that for an iid $\{X_i\}_{i=1}^n$ drawn from an $\alpha-$stable law with zero location parameter (recall that $\alpha \in (0, 2]$), the VaR for the partial sum $\sum_{i=1}^n X_i$ is $\text{VaR}(n) = n^{1/\alpha} F_\alpha^{-1}(\epsilon)$, where F_α is the distribution function. Here we see that only for the boundary case where $\alpha = 2$, the normal law, do we get the square–root–of–time rule for all ϵ and n. Any other stable distribution leads the square–root–of–time rule to *underestimate* the VaR: $\frac{\text{VaR}(n)}{n^{1/2}\text{VaR}(k)} = n^{1/\alpha-1/2} > 1$ iff $\alpha < 2$.

Our focus here is on systemic risk, however. While fat-tailed distributions may be more suited for day–to–day internal risk management, they may not be suited for the modelling of uncommon one–off events. The natural dgp for returns prone to systemic shocks (the raison d'être of Basel-II) is the jump diffusion, with its mostly continuous returns but with some rare, but very destructive, systemic events.[4] The validity of the

applied to arbitrary quantiles.

[3]While not much is known in theory, Guidolin and Timmermann (2004) provide some empirical evidence that suggests that some GARCH(1,1) processes may scale faster than the square–root–of–time rule over some horizons and for some VaR levels ϵ.

[4]While choosing Poisson jumps together with a Brownian motion leads to the most transparent underestimation results when the dgp is subjected to rare events, it is clear that an $\alpha-$ stable process with a further Poisson component would lead to a further underestimation compared to the jump-diffusion case studied here. In fact, our methodology applies if the Brownian motion was replaced by any other $\alpha-$stable process, but the results of our paper would be weaker since other than for $\alpha = 2$, the driving process itself would not obey the square–root–of–time rule. We assume a Brownian driving term not because it is necessarily easier, but because it biases our results in favour of the square–root–of–time rule and leads to a more transparent analysis of the failures of the square–root–of–time rule due to rare events.

jump diffusion approach is confirmed empirically by the substantially negatively skewed implied distribution of post-'87 out-of-the money put options, see Bates (2000), Pan (2002) and Carr and Wu (2003). Estimated jumps are on average negative and quite substantial, embodying the fear of further crashes and the rationale behind Basel-II.

In practice it is usually not possible to estimate the regulatory VaR since it requires perhaps a minimum of 300 observations of 10 day returns, suggesting that 12 years (250 trading days per year) are required for the estimation. As a result, the Basel Committee on Banking Supervision (1996) suggests that financial institutions estimate VaR at the daily frequency and scale it up to the 10 day frequency by $\sqrt{10}$. It is here that the question of the validity of the the square–root–of–time rule becomes especially pertinent. We demonstrate that the square–root–of–time rule leads to a systematic underestimation of risk, and can do so by a very substantial margin. This is so even if the dgp is iid and the square–root–of–time scaling of volatilities still obtains. The degree of underestimation worsens with the time horizon (at least up to some distant horizon for those cases for which a CLT applies), the jump intensity and the confidence level. Doing so, this paper also provides an interesting and perhaps unexpected insight into how Brownian and Poisson terms interact, and over which horizons one of the terms can dominate the other ones.

We find that even with upward drifts, for reasonably long holding periods, the square–root–of–time rule underestimates risk. The square–root–of–time rule performs best for horizons in the neighbourhood of 10 days, where the underestimation arising from the failure to address the systemic risk component is counterbalanced by the overestimation arising from the historically positive drift. This observation may provide a rationale for the choice of the scaling parameter 10.

2 The Economy

Following Merton (1971), we assume that wealth Y is governed by a jump diffusion process, where in the absence of a jump, the evolution of wealth follows a geometric Brownian motion. A Poisson shock occurs at some random time τ, at which time a fraction $(1-\delta) \in [0,1]$ of the portfolio value is wiped out. The recovery rate δ is constant and deterministic for simplicity. These dynamics can be written as (here Y_{t-} refers to the wealth at time t prior to any Poisson jumps that might occur at time t):

$$dY_t = \left(\mu + \frac{1}{2}\sigma^2\right) Y_{t-}dt + \sigma Y_{t-}dW_t - (1-\delta)Y_{t-}dq_t \tag{1}$$

Here $W = (W_t)_{t\in[0,T]}$ is a Brownian motion, μ is a constant and deterministic drift parameter and σ is a constant and deterministic diffusion parameter different from zero. The Poisson process driving the jumps is denoted by q, with constant and deterministic intensity λ. The units of λ^{-1} are average number of years between systemic events. q and W are stochastically independent processes.

Applying Itô's Lemma to the function $\ln(Y_t)$ and integrating from time 0 to t we get the expression for the wealth levels (by convention $0^0 = 1$ and $0 \cdot \infty = 0$),

$$Y_T = Y_t \exp(\mu(T-t) + \sigma(W_T - W_t))\delta^{q_T - q_t}$$

We fix the basis period k to a unit of time, e.g. one day or $k = 1/250$, and analyze the VaR over a horizon of ηk, e.g. $\eta k = 10$ days. The relevant return is then

$$X(t, t+\eta k) = \mu\eta k + \sigma(W_{t+\eta k} - W_t) + \ln(\delta)(q_{t+\eta k} - q_t) \tag{2}$$

Notice that the distribution of $X(t, t+\eta k)$ is independent of information available at time t: the underlying returns process is iid and volatilities scale with the square–root–of–time rule. This strengthens the message of this paper since the failure of the square–root–of–time rule when applied to quantiles cannot be due to the failure of volatilities to scale according to the square–root–of–time rule. Call the maximal number of shocks possible under the dgp as $I \in \{1, 2, \ldots, \infty\}$. It is understood that for $I < \infty$, $p_I(\eta k) := 1 - \sum_{i=0}^{I-1} p_i(\eta k)$. The $\mathrm{VaR}(\eta k)$ at the $\epsilon \in (0,1)$-level with a horizon $\eta k > 0$ can be deduced as follows, using Φ to represent the standard normal distribution function:

$$\epsilon = \mathbb{P}(X(t, t+\eta k) \leq -\mathrm{VaR}(\eta k)) = \sum_{i=0}^{I} \Phi\left(\frac{-\mathrm{VaR}(\eta k) - i\ln(\delta) - \mu\eta k}{\sigma\sqrt{\eta k}}\right) p_i(\eta k) \tag{3}$$

where $p_i(\eta k) := \mathbb{P}(q_{t+\eta k} - q_t = i) = \frac{(\lambda\eta k)^i}{i!} e^{-\lambda\eta k}$, $i < I$. With $\delta > 0$, a solution $\mathrm{VaR}(\eta k)$ to (3) always exists and is unique and continuously differentiable. If $\delta = 0$, a solution exists iff the probability of a crash is not too high, i.e. $\lambda\eta k < -\ln(1-\epsilon)$.

In order to show whether the square–root–of–time rule over– or underestimates the true VaR, and assuming we know the true data-generating process, we need to compare the proposed approximate VaR number $\sqrt{\eta}\mathrm{VaR}(k)$ with the true VaR number $\mathrm{VaR}(\eta k)$. In order to do so, define the (relative) underestimation (or correction) function $f_r(\eta) := \frac{\mathrm{VaR}(\eta k)}{\sqrt{\eta}\mathrm{VaR}(k)}$.

3 Main Results

It is well-known that the square–root–of–time rule is invalid over longer horizons, even in the absence of jumps, fat tails or time-varying volatility, for the drift matters over longer horizons and the VaR would have to be scaled by time rather than by the square–root–of–time. Without Poisson jumps, $\lambda = 0$, we get the standard VaR formula used in the financial industry

$$\mathrm{VaR}(\eta k) = -\sigma\sqrt{\eta k}\Phi^{-1}(\epsilon) - \mu\eta k \tag{4}$$

Absent the drift term $-\mu\eta k$, it embodies the square–root–of–time rule, $f_r(\eta) = 1$ for all $\eta \geq 0$. With the drift term, we call it the *mean-corrected square–root–of–time rule.* We see that $f_r(\eta) < 1$ in the realistic case where $\mu > 0$ and $\eta > 1$: the *approximative* critical loss $\sqrt{\eta}\mathrm{VaR}(k)$ is *larger* than the *actual* critical loss $\mathrm{VaR}(\eta k)$.

Since there is no explicit solution for VaR except in the borderline cases $\delta = 0, \delta = 1, \lambda = 0, \lambda = \infty$, $k = 0$ or $k = \infty$, and since the relation defined by (3) involves essentially non–algebraic implicit functions, definite results can only be shown under slightly more restrictive assumptions, as presented in the next propositions. The reason why theoretical results are important is that the problem of computing VaR numerically is prone to large numerical errors in the jump diffusion context. VaR is the zero of a rather special nonlinear function. Due to the low probability of a systemic crash, the function solved by VaR is nearly totally flat at zero over a large subset of the search space, requiring the programmer to devise very careful algorithms dealing with the problem of machine or floating point precision to capture the unique zero exactly. Any other candidate for the zero may lead to significantly erroneous VaR estimates.

Proposition 1 says that the VaR over any finite horizon per unit of square-root-of-time is larger than the square-root-scaled instantaneous basis VaR. If the basis is arbitrary, the scaling rule is derived in two benchmark cases. Please note that all proofs are relegated to the appendix.

Proposition 1 *Let the number of possible Poisson jumps satisfy $I \geq 1$. Then we have the following results:*

(i)

$$\frac{VaR(\eta k)}{\sqrt{\eta k}} + \mu\sqrt{\eta k} > \lim_{x\to 0}\left[\frac{VaR(x)}{\sqrt{x}} + \mu x\right], \qquad \text{all } \eta k > 0 \tag{5}$$

(ii) Assume that $\delta > 0$, then as $\eta k \to \infty$,

$$\frac{VaR(\eta k)}{\sqrt{\eta k}} \approx -\sqrt{\sigma^2 + \lambda(\ln\delta)^2}\Phi^{-1}(\epsilon) - \sqrt{\eta k}(\mu + \lambda\ln\delta) \tag{6}$$

(iii) Assume $\delta = 0$, then

$$\frac{VaR(\eta k)}{\sqrt{\eta k}} = \begin{cases} -\sigma\Phi^{-1}\left(1-(1-\epsilon)e^{\lambda\eta k}\right) - \mu\sqrt{\eta k} & ;\ \text{if } \lambda\eta k < -\ln(1-\epsilon) \\ +\infty & ;\ \text{otherwise} \end{cases} \tag{7}$$

where $\Phi^{-1} : (0,1) \to \mathbb{R}$ is the inverse function of Φ. If $\delta = 0$ and also $\mu = 0$, then $f_r(\eta) > 1$ for $\eta > 1$ and $f_r(\eta)$ increases in η.

Result (5) with $\mu = 0$ shows that the square–root–of–time rule strictly underestimates the risk when using the base VaR or the VaR over any other sufficiently short base horizon, i.e. $f_r(\eta) > 1$:

$$\mathrm{VaR}(\eta k) > \sqrt{\eta}\mathrm{VaR}(k) \quad \text{for } k \geq 0 \text{ sufficiently small}, \eta > 1$$

With $\mu \neq 0$, the mean-corrected square–root–of–time rule underestimates the true VaR. This result also implies that there must be some horizon, say η^h, over which the underestimation worsens with η, $\eta \leq \eta^h$.

Result (5) is a complete characterization of the problem with an infinitesimal basis period. In practice, however, the basis time period is a discrete amount of time, usually a day, and while likely, we are not guaranteed in theory that $k = 1/250$ is sufficiently small in the sense of the previous proposition. In practice, however, we are comfortable that (5) extends to an arbitrary k. First, and for any given base period k, we do have closed form results if δ is either 0 or 1. The working paper version of this article (Danielsson and Zigrand (2005)) provides a proposition that uniformly extends result (5) to neighbourhoods of δ around 0 and 1 for any given base period k, as well as to neighbourhoods of $k = 0$ for any given $\delta \in [0, 1]$. Furthermore, extensive numerical simulations performed by the authors suggest that these results are valid for intermediate values of $\delta \in [0, 1]$ as well, as can be seen in Figures 2 and 3. Looking at those two figures, it becomes clear that the relative risk errors $f_r(\eta) = \mathrm{VaR}(\eta k)/(\sqrt{\eta}\mathrm{VaR}(k))$ arising from the square–root–of–time rule are roughly the same for any potential losses above 25%, and that is true for all levels of λ. This is why we can claim that the predictions of our model carry over to more realistic environments with both partial crashes and an arbitrary basis period.

It turns out that two distinct main cases need to be analyzed: $\delta > 0$ and $\delta = 0$. If $\delta > 0$, result (6) characterizes the scaling law as the square–root–of–time rule (or the mean-corrected version thereof) in the limit, regardless of the basis, for in this case the CLT applies. While in practice one is more interested in finite η, this result is important as it shows that for $\delta > 0$ (the same will not be true for $\delta = 0$), the initial worsening of the underestimation as the horizon η is extended eventually will give way to a bettering of the underestimation (but always to an underestimation).

In case the disaster is severe, $\delta = 0$, we are able to find a closed-form solution in equation (7). If the drift is zero, then the degree of underestimation is positive. In fact, it worsens with the length of the extrapolation horizon, violating the predictions of the CLT.[5] It is worthwhile to see the drivers of this result. Losses worse than VaR can occur for two reasons: either the return is driven down by a string of bad news as modelled by the Brownian motion, or a systemic Poisson jump occurs. The crux lies in how the Brownian and the Poisson terms interact. If $\lambda = 0$, then the Brownian term is seen to be of order $\sqrt{\eta}$, as expected. If $\lambda > 0$, the Poisson risk introduces via

[5]The version with $\delta = 0$ is not in the Gaussian domain of attraction.

$(1-\epsilon)\big(1-e^{\lambda\eta k}\big)$ an additional term nonlinear in η. In some sense, the effect of the Poisson risk on VaR is to scale down the confidence level ϵ in an exponential fashion as a function of η. This makes the Φ^{-1} term depend on η: for small values of η, $F(\eta) := -\Phi^{-1}\left(\epsilon + (1-\epsilon)\big(1-e^{\lambda\eta k}\big)\right)$ is rather flat, while F increases exponentially as η grows. The fact that the Poisson term dominates the drift term over the longer horizons may be a bit surprising. Recall that it is well-known that the Brownian term on its own is of order $\sqrt{\eta}$, the drift is of order η and the Poisson term has both a variation and a quadratic variation equal to the counting measure. Our results show that the common additive linear intuition one might have from the linear additive jump diffusion (1) that neither the Poisson term nor the Brownian term is therefore ever going to dominate over any longer horizon due to the order of their variations compared to the drift is false. The two lower orders interact nonlinearly.

The key question must be how relevant the downward biais is in practice. We perform numerical simulations where the drift and volatility are calibrated to historical annualized values of the S&P–500 index, 5.48% and 15.84%, respectively. In the first set of numerical simulations we follow standard practice in the risk management industry and assume the drift is zero. Tables 1 and 2 show the VaR from using the square–root–of–time rule ($\sqrt{\eta}\mathrm{VaR}(k)$) as well as the correct VaR ($\mathrm{VaR}(\eta k)$), while Figure 1 shows the relative error in the square–root–of–time rule for a range of holding periods (η) and crash frequencies ($1/\lambda$). As expected, the lower the crash frequency and the shorter the holding periods, the smaller the bias is. It is, however, at the other range where the results become more interesting. Longer holding periods or higher crash probabilities λ imply that the square–root–of–time rule becomes increasingly inaccurate because a systemic crash becomes increasingly likely. As can be seen in the table, the errors committed can be quite substantial even for such short horizons as 10 or 20 days.

4 More on Scaling with a Positive Drift

In most applications of the square–root–of–time rule to risk management, it is assumed that the drift is zero. This practice has also been advocated, among others, by Jorion (2001). There are several reasons for this. First, since most risk models operate at high frequencies, assuming the drift is zero is relatively innocuous since at those measurement horizons the drift is an order of magnitude smaller than the volatility. The main problem is however that there is no obvious way to obtain an accurate estimate of the drift (see e.g. Merton, 1981).

Suppose however that the drift μ is strictly positive and known and that the regulatory rules require use of the VaR applied to raw returns. We assume throughout this section that $\delta = 0$. Since some practitioners recognize (Blake et al., 2000) that the presence of a strictly positive drift term leads to an overestimation of risk, they presumably take this overestimation into consideration when allocating risk and capital. Jumps can therefore again lead to undercapitalization. The flavour of the following theorem can be summarized by saying that as long as the drift is not too large, the square–root–of–time rule still underestimates the true VaR.

Proposition 2 *Assume that $I = 1$, $\lambda > 0$ and that $\delta = 0$.*

(i) For a given horizon η, the square–root–of–time rule underestimates the value-at-risk iff μ is smaller than some critical $\bar{\mu} := \bar{\mu}(\eta, k, \lambda, \sigma, \epsilon)$, equivalently iff $\eta > \eta'(\mu, \lambda, k, \epsilon)$. We have $\bar{\mu} > 0$ and $\frac{\partial \bar{\mu}}{\partial \eta} > 0$ iff $\eta > 1$.

(ii) Assume that $\epsilon \leq 1 - \frac{1}{2}e^{-\lambda k}$. Within the interval of underestimation $\{\eta : \eta > \eta'\}$, $f_r(\eta)$ is strictly increasing in η and in λ.

Proposition 2 says that as long as the drift term is sufficiently small,[6] the square–root–of–time rule underestimates the VaR, and this error is worse the larger the likelihood of a systemic event and the larger the horizon of extrapolation.

To acquire a deeper intuitive understanding of the interrelation between normal market risk, as modelled by the Brownian part and the drift part, and the systemic risk, as modelled by the Poisson part, consider the simple case where $\delta = 0$. Refer again to (7). By what we said above, for small η, the Poisson term is dominated by the drift term, while for larger η the Poisson term (in interaction with the Brownian term) dominates the drift term despite the fact that the drift term is linear and of order η. In particular, and despite the fact that one does encounter this argument from time to time, it is not true that the VaR over long periods can necessarily be assumed to be the quantile of a normal distribution, not even if it is mean-corrected.

In order to assess the bias induced by the square–root–of–time when the drift is positive, we perform numerical simulations where the drift and volatility are calibrated to historical annualized values of the S&P–500 index, (from beginning until 2002) i.e., 5.48% and 15.84%, respectively. With $\lambda = 1/25$ and $\epsilon = .01$, we can see from Figure 4 that $\bar{\mu} \geq .0548$ roughly from $\eta = 3$ onwards. In other words, for an extrapolation horizon of more than three days, the square–root–of–time rule not only underestimates the true VaR, but the errors committed increase with η.

If the drift is positive the bias is less severe than without a drift. This can be seen in Figure 5, which is identical to Figure 1, with the exception of the drift which is set to the average S&P drift. The ratio of the true to the time-scaled VaR is still above 1 almost everywhere, but a bit more subdued. For $\lambda^{-1} \geq 25$ years, since $f_r(\eta = 1) = 1$,

[6] The expression for $\bar{\mu}$ can be found in the proof as equation (9), and the function is also plotted in Figure 4 in the Appendix. The condition $\epsilon \leq 1 - \frac{1}{2}e^{-\lambda k}$ is implied by $\epsilon \leq \frac{1}{2}$. Since VaR is about rare tail events, and that in practice $\epsilon \leq .1$, this condition is not binding. The shape of the function is qualitatively unaffected by the choice of parameters k, λ, σ and ϵ (as can be verified from the exact functional form provided in (9) in the Appendix).

raising the horizon slightly first leads to overestimation (in view of the fact that the drift term dominates the Poisson term for small η) and reaches a minimum for holding periods of around 10 days. As the holding periods increase, the ratio exceeds 1 at an increasing rate. For Basel II regulatons, the relevant neighbourhood is indeed $\eta = 10$. Figure 6 shows a slice of Figure 5 where the holding period $\eta = 10$. It is interesting to note that with historical (μ, σ), the 10 day horizon extrapolation is least prone to underestimation, at least if crashes are rather unlikely. The reason is that the historical drift term leads to an overestimation bias that compensates the underestimation bias arising from the jump term. In view of the importance of the 10 day VaR number for risk-regulation, the cancellation of both biases at that horizon is a happy coincidence.

5 Conclusion

Regulatory recommendations and common derivatives pricing models implicitly assume iid normal returns, implying that the square–root–of–time can be used to scale volatilities and risk.

We consider the implications of time scaling quantiles of return distributions by the square–root–of–time when the underlying stochastic process is a jump diffusion. Our results indicate that an application of the square–root–of–time rule to the forecast of quantile-based risk estimates (such as Value–at–Risk) when the underlying data follows a jump–diffusion process is bound to provide downward biased risk estimates. Furthermore, the bias increases at an increasing rate with longer holding periods (at least up to some remote horizon), larger jump intensities or lower quantile probabilities. The reason is that the scaling by the square root of time does not sufficiently scale the jump risk which interacts nonlinearly with the Brownian term. An exception may be at the 10 day horizon where the underestimation arising from the systemic component has historically been counterbalanced by the overestimation induced by the drift term.

A Appendix: Proofs

Proof of Proposition 1 (i) The equation determining VaR is given by:

$$\epsilon = \sum_{i=0}^{I} p_i(\eta k)\Phi\left(-\frac{\nu}{\sigma} - i\frac{\ln\delta}{\sigma\sqrt{\eta k}} - \frac{\mu\sqrt{\eta k}}{\sigma}\right) \tag{8}$$

with $\nu := \mathrm{VaR}/\sqrt{\eta k}$. We argue that we need $\Phi(-\frac{\nu}{\sigma} - \frac{\mu\sqrt{\eta k}}{\sigma}) < \epsilon$. Indeed, assume to the contrary that $\Phi(-\frac{\nu}{\sigma} - \frac{\mu\sqrt{\eta k}}{\sigma}) \geq \epsilon$. Then $\Phi(-\frac{\nu}{\sigma} - i\frac{\ln\delta}{\sigma\sqrt{\eta k}} - \frac{\mu\sqrt{\eta k}}{\sigma}) > \epsilon$ for all $i > 0$. But then the RHS of (8) is strictly larger than ϵ, a contradiction to ν being a solution.

For $\eta k \to 0$, the lower bound is in fact achieved. Since $\lim_{\eta k \to 0} p_i(\eta k) = 0$ for all $i > 0$ and $\lim_{\eta k \to 0} p_0(\eta k) = 1$, equation (8) converges to $\epsilon = \Phi(\lim_{\eta k \to 0} -\frac{\nu}{\sigma} - \frac{\mu\sqrt{\eta k}}{\sigma})$. Hence $\Phi(-\frac{\nu(\eta k)}{\sigma} - \frac{\mu\sqrt{\eta k}}{\sigma}) < \epsilon = \Phi(\lim_{\eta k \to 0} -\frac{\nu(\eta k)}{\sigma} - \frac{\mu\sqrt{\eta k}}{\sigma})$, as required.

As to the limiting result (6), it suffices to notice that period returns satisfy $E[X_j^2] < \infty$ and that the CLT applies as a result. Finally, result (7) simply follows by inverting Φ. ■

Proof of Proposition 2

We assume that parameters are in their respective domains, $\eta > 1$ and $0 < \lambda\eta k < -\ln(1-\epsilon)$. As to (i), by definition, $f_r(\eta) \geq 1$ iff

$$\mu \leq \bar{\mu}(\eta, k, \lambda, \sigma, \epsilon) := \frac{\sigma\left[\Phi^{-1}(1-(1-\epsilon)e^{\lambda k}) - \Phi^{-1}(1-(1-\epsilon)e^{\lambda\eta k})\right]}{\sqrt{k}(\sqrt{\eta}-1)} \tag{9}$$

As to (ii), denoting the standard normal density by ϕ, we see that $\frac{\partial f_r(\eta)}{\partial\eta} > 0$ iff $\Leftrightarrow \mu < \frac{2\sigma\sqrt{\eta k}(1-\epsilon)\lambda e^{\lambda\eta k}}{\phi(\Phi^{-1}(1-(1-\epsilon)\lambda e^{\lambda\eta k}))} := \underline{\mu}$. It is intuitively clear that as η grows, f_r first needs to rise before it can rise above 1, i.e. that $\bar{\mu} < \underline{\mu}$.[7] In order to prove that $\frac{\partial\bar{\mu}}{\partial\eta} > 0$ for $\eta > 1$, it turns out that $\frac{\partial\bar{\mu}}{\partial\eta} > 0$ iff $\bar{\mu} < \underline{\mu}$.

As to the last assertion, $\frac{\partial(\mathrm{VaR}(\eta k)-\sqrt{\eta}\mathrm{VaR}(k))}{\partial\lambda} > 0$ iff $2\ln\eta + 2\lambda k(\eta-1) - [\Phi^{-1}(1-(1-\epsilon)e^{\lambda k})]^2 + [\Phi^{-1}(1-(1-\epsilon)e^{\lambda\eta k})]^2 > 0$. The first two terms are positive, while the last two terms taken together are positive if $\epsilon \leq 1 - \frac{1}{2}e^{-\lambda k}$. If that is the case, then both expressions in the second term square negative numbers, the second of which being larger in absolute value given that the function Φ^{-1} is monotonically increasing on its domain $(0,1)$ and negative for values below $\frac{1}{2}$. ■

B Appendix: Tables and Figures

[7]Indeed, $\underline{\mu} > \bar{\mu}$ iff $\Phi^{-1}(1-(1-\epsilon)e^{\lambda k}) - \Phi^{-1}(1-(1-\epsilon)e^{\lambda\eta k}) < \frac{2k\sqrt{\eta}(\sqrt{\eta}-1)(1-\epsilon)\lambda e^{\lambda\eta k}}{\phi(\Phi^{-1}(1-(1-\epsilon)\lambda e^{\lambda\eta k}))}$. Now by assumption, $1-(1-\epsilon)\lambda e^{\lambda k} < \frac{1}{2}$ so that Φ^{-1} is increasing and concave over the posited range. By concavity we get $\Phi^{-1}(1-(1-\epsilon)e^{\lambda k}) - \Phi^{-1}(1-(1-\epsilon)e^{\lambda\eta k}) < \frac{2k(1-\epsilon)\lambda e^{\lambda\eta k}}{\phi(\Phi^{-1}(1-(1-\epsilon)\lambda e^{\lambda\eta k}))}(\eta-1)$. The proof is complete if $\eta - 1 < 2\sqrt{\eta}(\sqrt{\eta}-1)$, which always holds for $\eta > 1$.

Table 1: 1% VaR on a $1000 portfolio with S&P–500 volatility, varying λ^{-1}

10 and 20 day VaR, $k = 1/250, \epsilon = 0.01, \sigma = .1584, \mu = 0$. Ratio is $f_r(\eta) = \text{VaR}(\eta)/(\sqrt{\eta}\text{VaR}(1))$. $1/\lambda$ indicates the expected time of crash in years. $1 - \delta = 100\%$ indicates the fraction of wealth wiped out in crash.

Measure	Expected crash time in years $1/\lambda$				
	10	20	30	40	50
VaR(10)	79.5	76.3	75.4	74.9	74.7
$\sqrt{10}$VaR(1)	74.2	73.9	73.9	73.8	73.8
Ratio	1.07	1.03	1.02	1.02	1.01
VaR(20)	128.6	112.5	109.3	107.9	107.1
$\sqrt{20}$VaR(1)	104.9	104.6	104.4	104.4	104.4
Ratio	1.23	1.08	1.05	1.03	1.03

Table 2: 1% VaR on a $1000 portfolio with S&P–500 volatility, varying η

VaR on a $1000 portfolio of S&P–500 Volatility. $k = 1/250, \epsilon = 0.01, \sigma = .1584$. $1/\lambda = 25$ indicates the expected time of crash in years, and $1 - \delta = 25\%$ indicates the fraction of wealth wiped out in crash. Ratio is $f_r(\eta) = \text{VaR}(\eta)/(\sqrt{\eta}\text{VaR}(1))$.

Measure	Holding period in days η					
	10	20	30	40	50	60
VaR(η)	75.7	110.5	140.5	170.0	203.4	257.6
$\sqrt{\eta}$VaR(1)	73.9	104.5	128.0	147.8	165.2	181.0
VaR(η)/($\sqrt{\eta}Var(1)$)	1.02	1.06	1.10	1.15	1.23	1.42

Figure 1: Relative error in the square–root–of–time rule, S&P-500 annual volatility, zero mean, full wealth wipeout, and 1% VaR

$\mu = 0, \sigma = .1584, \delta = 0, k = 1/250, \epsilon = 0.01$. Ratio is $\mathrm{VaR}(\eta)/(\sqrt{\eta}\mathrm{VaR}(1))$. $1/\lambda$ indicates the expected time of crash in years.

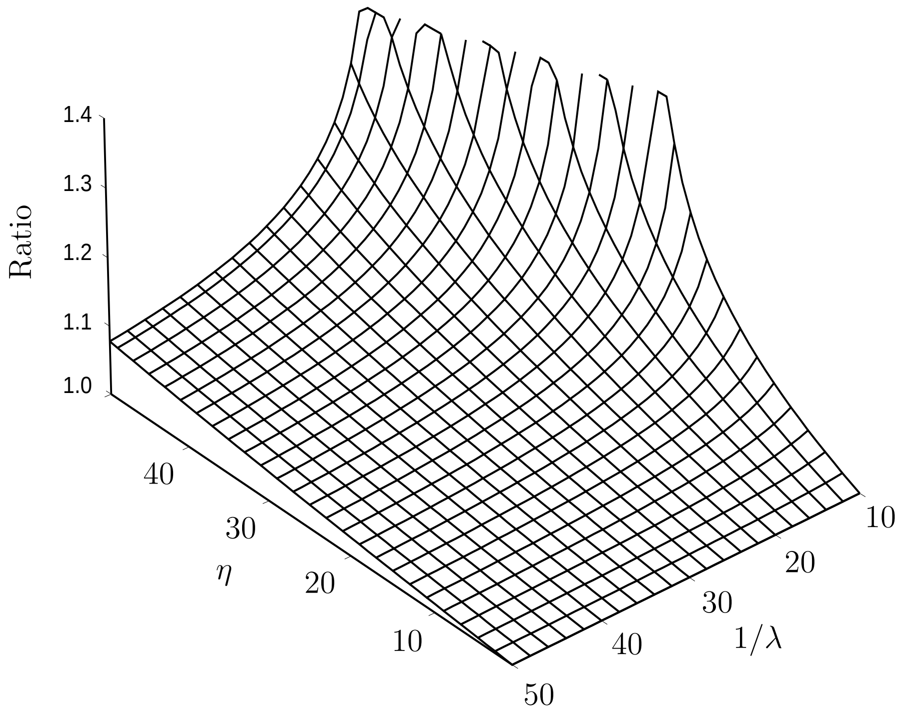


Figure 2: Relative error in the square–root–of–time rule, S&P-500 annual volatility, zero mean, partial wealth wipeout, and 1% 10 day VaR

Model parameters are: $\mu = 0.0$, $\sigma = .1584$, $\epsilon = 0.01$, $\eta = 10$. Ratio is $\mathrm{VaR}(\eta)/(\sqrt{\eta}\mathrm{VaR}(1))$. $1/\lambda$ indicates the expected time of crash, $1 - \delta$ is the fraction that is wiped out in the crash.

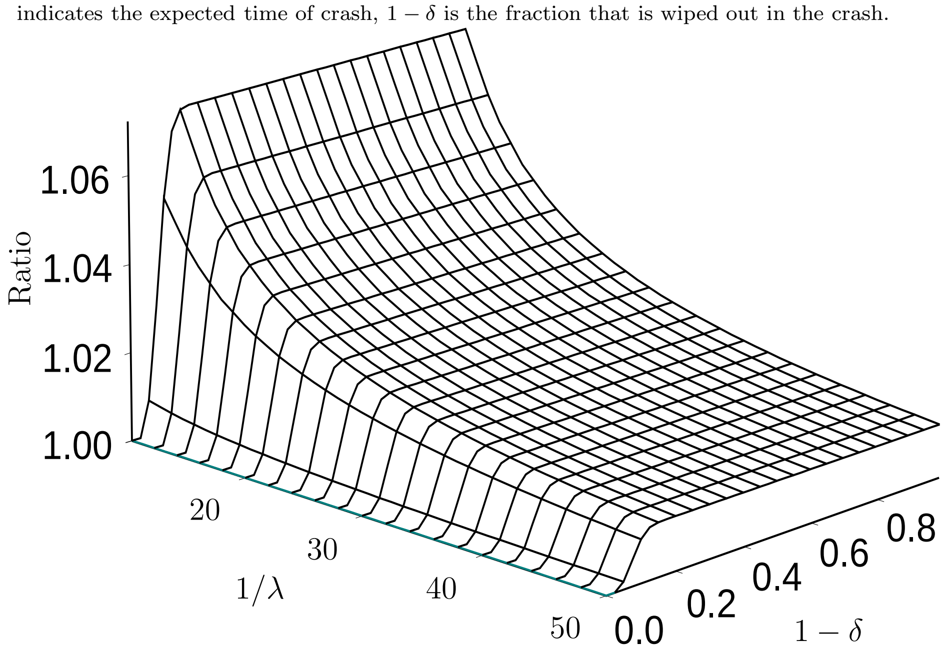


Figure 3: Relative error in the square–root–of–time rule, S&P-500 annual volatility, zero mean, partial wealth wipeout, and 1% VaR

Model parameters are: $\mu = 0.0, \sigma = .1584, k = 1/250, \epsilon = 0.01, 1/\lambda = 25$. Ratio is $\mathrm{VaR}(\eta)/(\sqrt{\eta}\mathrm{VaR}(1))$.

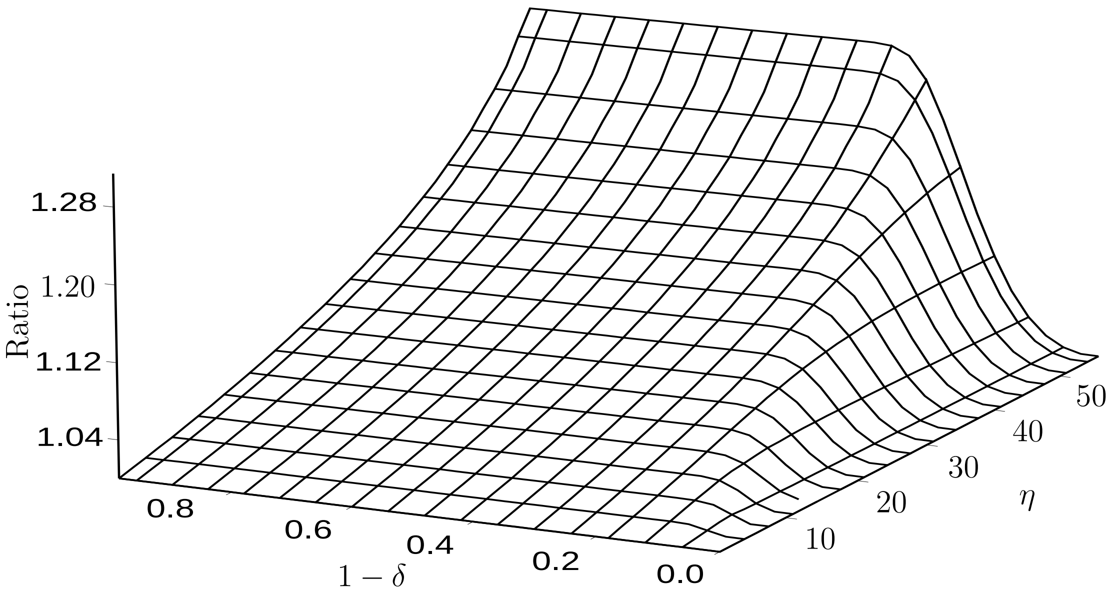


Figure 4: The critical drift function $\bar{\mu}$ as a function of η

$\sigma = .1584, \delta = 0, k = 1/250, \epsilon = 1/250, \lambda = 1/55$.

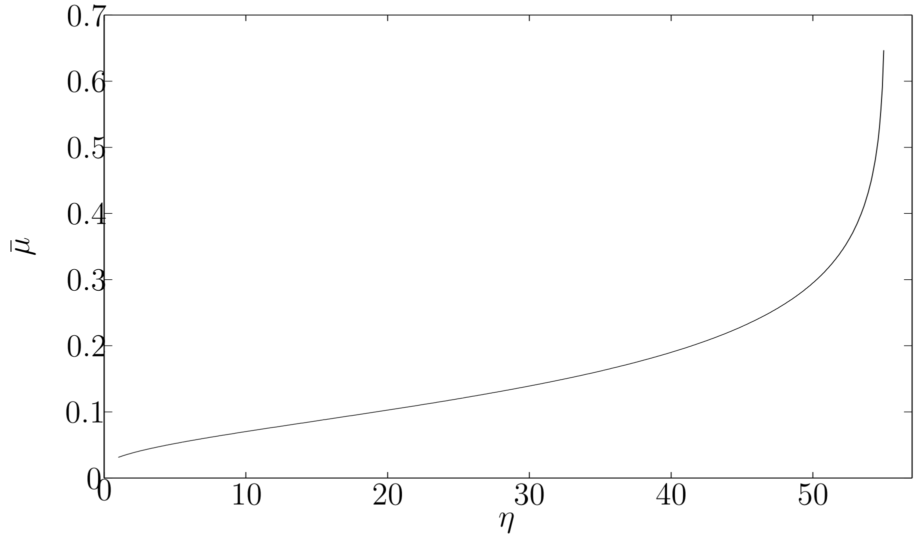


Figure 5: Relative error in the square–root–of–time rule, S&P-500 annual volatility and mean, full wealth wipeout, and 1% VaR

$\mu = 0.0548$, $\sigma = .1584$, $\delta = 0$, $k = 1/250$, $\epsilon = 0.01$. Ratio is $\mathrm{VaR}(\eta)/(\sqrt{\eta}\mathrm{VaR}(1))$. $1/\lambda$ indicates the expected time of crash in years.

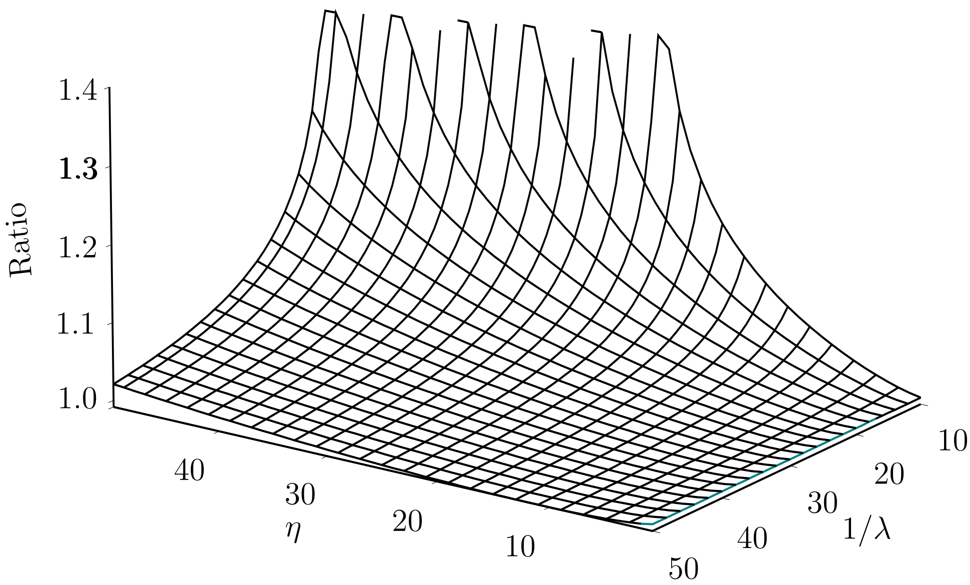
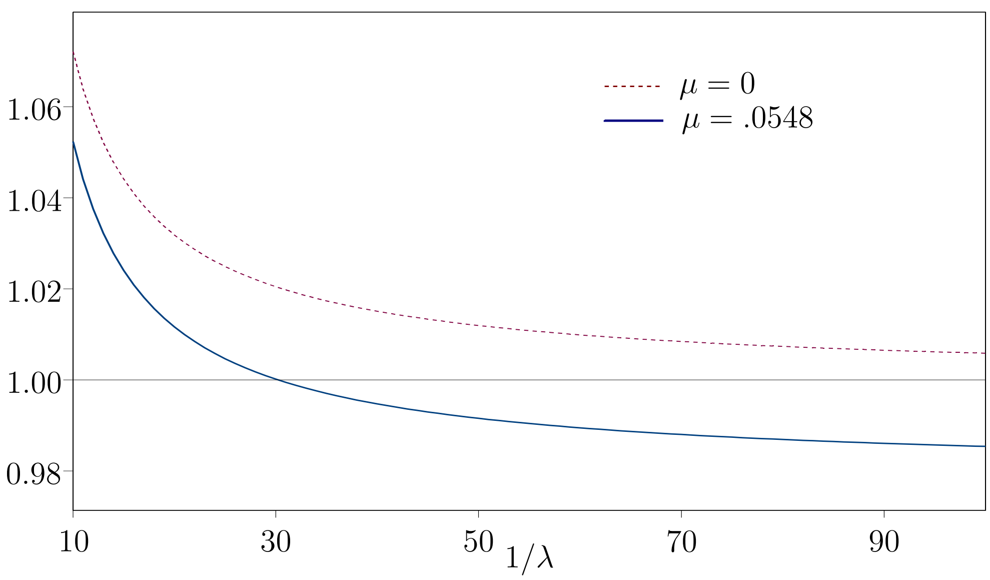


Figure 6: Relative error in the square–root–of–time rule, S&P-500 annual volatility and mean, full wealth wipeout, and 1% 10 day VaR

$\mu = 0.0548, \sigma = .1584, \delta = 0, k = 1/250, \epsilon = 0.01, \eta = 10$. $1/\lambda$ indicates the expected time of crash in years.



References

Basel Committee on Banking Supervision (1996). *Overview of the Amendment to the Capital Accord to Incorporate Market Risk*.

Basel Committee on Banking Supervision (2002). Quantitative impact study 3. technical guidance. Technical report.

Bates, D. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94:181–231.

Blake, D., Cairns, A., and Dowd, K. (2000). Extrapolating VaR by the square-root rule. *Financial Engineering News*, pages 3,7.

Carr, P. and Wu, L. (2003). What type of process underlies options? A simple robust test. *Journal of Finance, forthcoming*.

Dacorogna, M., Müller, U., Pictet, O., and de Vries, C. (2001). Extremal forex returns in extremely large datasets. *Extremes*, 4:105–127.

Danielsson, J. and Zigrand, J.-P. (2005). On time-scaling of risk and the square-root-of-time rule. Mimeo, `www.riskresearch.org`.

Diebold, F., Hickman, A., Inoue, A., and Schuermann, T. (1997). Converting 1-day volatility to h-day volatility: Scaling by $\sqrt{h}$ is worse than you think. Discussion paper series, no. 97-34, Wharton.

Duffie, D. and Pan, J. (2002). Analytical value-at-risk with jumps and credit risk. *Finance and Stochastics,* (forthcoming).

Eberlein, E., Keller, U., and Prause, K. (1998). New insights into smile, mispricing and value at risk. *Journal of Business*, 71:371–406.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica*, 50:987–1007.

Feller, W. (1971). *An Introduction to Probability Theory and Applications, Second Edition. Volume II.* John Wiley and Sons.

Groenendijk, P., Lucas, A., and de Vries, C. (1998). A hybrid joint moment ratio test for financial times series. Preprint, Erasmus University.

Guidolin, M. and Timmermann, A. (2004). Term structure risk under alternative econometric specifications. Working Paper.

Jarrow, R. and Turnbull, S. (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, 50(1):53–85.

Jorion, P. (2001). *Value at Risk, Second Edition.* McGraw Hill.

Ju, X. and Pearson, N. D. (1999). Using value-at-risk to control risk taking: How wrong can you be? *Journal of Risk*, 1(2):5–36.

Merton, R. (1971). Optimal consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3:373–413.

Merton, R. (1981). On estimating the expected return on the market: an exploratory investigation. *Journal of Financial Economics*, 8:323–361.

Müller, U., Dacorogna, M., Olsen, R., Pictet, O., Schwartz, M., and Morgenegg, C. (1990). Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking and Finance*, 14:1189–1208.

Pan, J. (2002). The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3–50.