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

Jón Daníelsson
Jean–Pierre Zigrand

Department of Accounting and Finance
Financial Markets Group
London School of Economics

March 2003

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 risk 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.

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[1] (VaR), is by estimating a one day VaR and multiplying it by $\sqrt{10}$. 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. For instance VaR and related measures are used to control, and set limits to, traders' positions.

The underlying distributional assumptions behind this method are quite stringent, and are violated in most, if not all, practical applications. Since the distributional stylized facts for returns 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. Below we formally analyze the biases introduced by the square–root–of–time rule applied to quantiles when asset returns are driven by a jump diffusion. 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.

The time scaling of volatilities for instance implicitly depends on returns being iid (identically and independently distributed), an assumption Engle (1982) argues is incorrect because of the presence of volatility clusters. When applied to quantiles, the square–root–of–time rule also depends on the normality of returns. It has been known at least since Mandelbrot (1963) and Fama (1965) that returns exhibit excess kurtosis, e.g. they are "fat tailed." In general, the presence of fat tails introduces an additional bias in applications of the square–root–of–time rule[2] to quantile forecasts, such as in

---

[1] VaR is the quantile that solves $\epsilon = \int_{-\infty}^{-\text{VaR}} \widehat{f(x)}dx$, where $\widehat{f(x)}$ is the estimated probability density function of a financial institution's return and ϵ is the confidence level, say .01. The reason we write -VaR in the above equation is that VaR represents a critical potential loss, and we find it more intuitive to think about a loss as of a positive number.

[2] If returns are conditionally normal but with volatility clusters, the square–root–of–time rule may either over or underestimate quantile forecasts, (see e.g. Drost and Nijman, 1993). If returns follow an unconditionally iid fat tailed distribution, then the square–root–of–time rule overestimates quantile forecasts, see e.g. Feller (1971, VIII.8) and Dacorogna et al. (1999) for more on this. If returns are fat tailed with volatility clusters the square–root–of–time rule may either over or underestimate the

the pricing of path dependent derivatives and in risk management. For example, in the pricing of barrier derivatives, a violation of iid normality implies that the knockout and knockin probabilities are different from what is assumed by standard pricing models. It is in risk management, however, that violations of iid normality may have the strongest impact. In this paper, we shall show that under reasonable assumptions, the square–root–of–time rule applied to VaR underestimates the true VaR, and can do so by a very substantial margin. This result is reminiscent of Ju and Pearson (1999), where the estimated VaR underpredicts the true VaR because of superior information possessed by the trader vis-à-vis the historically estimated VaR. If the trader knows that the estimated VaR lies below the true VaR, she takes on more risk than allowed by the (true) VaR risk limit. In our paper, underestimation arises from statistical reasons.

In order to measure the bias introduced by the square–root–of–time rule, a data generating process needs to be specified. The literature has seen two major approaches to the modelling of fat tails in return series, fat tailed distributions (e.g. the Student–t) and jump diffusion processes. In choosing between those two methodologies, one must consider the intended applications. The fat tailed distribution approach generates fat-tailed return series by putting more mass in the tails than the normal does. Each period's return therefore is a bit more likely to be in the tail, but intuitively such realizations in the tails occur "continuously in time." This is a common assumption in risk modelling, and often provides more accurate risk forecasts than an assumption of normality. Unfortunately, no consensus exists as to which fat tailed return distribution should be used, with the Student–t a common choice. Risk modelling with continuous fat tailed distributions is further hampered by the problem that multivariate representations of fat tailed distributions either do not exist, are not unique, or are computationally difficult. The bias in using the square–root–of–time to scale quantiles in this case is established by Dacorogna et al. (1999) who demonstrate in theory that if financial returns are iid fat tailed but do not follow the stable law, the square–root–of–time rule overestimates risk.

While fat-tailed distributions may be more suited for day–to–day internal risk management, jump diffusion models are better suited for the modelling of uncommon one–off events, e.g. systemic risk. Jump diffusions define a mostly continuous return process that is subjected to occasional discrete jumps, and the resulting distribution also exhibites fatter than normal tails. Such models capture rare but large shocks to the financial system over and above the normal day-to-day volatility. In particular, we feel that jump processes are better suited for the modelling of systemic risk than the continuous approach. Section 2.1 below provides a short survey of the relevant literature on jumps. Recently, several different methods for modeling systemic risk from a financial perspective have been proposed, e.g. Allen and Gale (2000), who model the domino effect of bank failures in markets with incomplete interactions, Diamond and Rajan

---

quantile forecast.

(2001) who model systemic events induced via a liquidity crisis, and Danielsson and Zigrand (2001), who model systemic risk as arising from externality–induced excessive risk–taking. Despite being in early stages, what links these approaches is the notion that the financial system experiences a discrete transition from stable to crisis periods, suggesting that a jump process model, with only downside jumps, is well–suited as a statistical modelling methodology of systemic risk.

The fear of a systemic crisis is the driving force behind international capital adequacy regulations, especially after the Asian and Russian-LTCM crises. In fact, the Basel agreements are motivated by the notion that capital adequacy regulations are not meant to address day to day risks but rather large systemic events. To quote the General Manager of the BIS, Crockett (2000) "The quintessential micro–prudential dictum is that financial stability is ensured as long as each and every institution is sound (. . .) This statement may strive for too much, because the occasional failure of individual institutions is not the problem. Trying to avoid such outcomes risks providing excessive protection, with the result that market disciplinary and allocative mechanisms are weakened." This view suggests that supervisors, in designing capital adequacy regulations, focus primarily on systemic events, where the chosen regulatory instrument should not be sensitive to day–to–day risks. The balance suggested by the BIS is to use a particular quantile of the return distribution as the main control for risk, i.e. the 99% 10 day holding period VaR.[3] While deriving bank capital from this VaR level may reduce systemic risk, in practice it is usually not possible to estimate the 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.

Our main finding is that by applying the square–root–of–time rule to the time scaling of quantiles of return distributions, risk is underestimated at an increasing rate as the extrapolation horizon is extended, as the probability of a jump increases or as the confidence level is raised. In particular, as the scaling horizon increases, the bias introduced by the square–root–of–time rule grows at a rate faster than time. Since the

---

[3]A key objective of this paper is to examine time scaling of the regulatory 99% VaR. Our objective is not to explore the optimality of the chosen regulatory regime. For instance, Ahn et al. (1999) and Ju and Pearson (1999) have shown how traders can "game" the VaR, in the former paper via options strategies, and in the latter paper due to the superior information about VaR that a trader may have compared to the historically estimated VaR number. We also refer the reader to Artzner et al. (1999) for a discussion of the incoherence of the VaR measure, and to Basak and Shapiro (2001), Danielsson and Zigrand (2001) and to Danielsson et al. (2003) for a study of the, sometimes perverse, effects of VaR regulation in equilibrium models. Instead, here we take them as given, and explore the implications of a reasonable statistical model of systemic crises for the regulatory VaR measure.

drift only grows linearly with time, and the Brownian part is only of order $\sqrt{t}$, this result holds eventually for any jump diffusion process. We argue that these results hold for reasonable parameter values and visualize them by numerical simulations presented graphically towards the end of the paper. While an exact measurement of the bias can be obtained, we provide simple approximations that may provide a useful *rule of thumb* when the parameters are not known.

To augment the theoretical results, we also consider numerical simulations from a stochastic process calibrated to the S&P–500 return. Even with upward drifts, for reasonably long holding periods, we find that the square–root–of–time rule understimates risk. When we consider short regulatory holding periods around 10 days, we find that the square–root–of–time rule underestimates the true VaR when the market underperforms historical averages, and overestimates the VaR when the market exceeds the historical average. On average, while the square–root–of–time rule may be a poor approximation over many horizons, it performs best for horizons in the neighbourhood of 10 days. The reason is that for a 10 day horizon, the underestimation arising from the failure to address the systemic risk component is counterbalanced by the overestimation arising from the historically positive drift.

From a policy perspective, our results illustrate the perversity of the square–root–of–time rule. The supervisors imposed a VaR based capital requirement presumably to minimize systemic risk, whereas the square–root–of–time rule all but eliminates this aspect from the capital requirement by relying only on sources of risk that scale up with the square-root of time e.g. the Brownian risk). The only downside risk that is taken into consideration is thus the non–systemic one, which is the regular volatility of wealth. As such, the square–root–of–time rule may be viewed as failing in its intended task. Furthermore, efforts to extend the square–root–of–time rule to longer horizons, quarterly and beyond, must be considered misguided, and a proper correction factor needs to be applied carefully.

The structure of the paper is as follows. We discuss the benchmark model in section 2, while presenting the impa iows-39958.5(d0.3(f)-5ll)-395.1(but)-b]TJ 0 -3080.2(n)-3ctor080.2(n)-

2 The Benchmark Economy

2.1 Jump Models and Systemic Risk

While it has been known at least since Mandelbrot (1963) and Fama (1965) that financial returns are fat tailed in the sense that the kurtosis of returns exceeds 3, no consensus exists as to how to model the fat tails. The actual choice of a modelling strategy depends on the underlying questions and on the theoretical models. For the question of systemic risk in particular, the rare event when sizeable fraction of wealth is wiped out, a jump model provides a useful approximation. Indeed, the reduced form of recent models of systemic financial risk (see e.g. Allen and Gale, 2000; Diamond and Rajan, 2001; Danielsson and Zigrand, 2001) is in essence a jump model.

Jump models have been widely applied in both theoretical and empirical finance. Our model follows most closely the intensity-based bankruptcy model of Jarrow and Turnbull (1995). They assume that the intensity of the Poisson process which jumps from the no-default to the default state is an exogenous constant. In our context there is, given our assumptions, in essence little formal difference between the conventions of recovery-of-treasury (the creditor receives upon default a fraction of an identical but default-free bond) imposed by Jarrow and Turnbull (1995) and the "recovery-of-market-value" (a fixed fractional recovery rate of the market value) assumed by Duffie and Singleton (1999). Empirically, many authors have estimated jump models of asset returns (e.g. Ball and Torous, 1983; Akgiray and Booth, 1988; Jorion, 1988; Johnson and Schneeweis, 1994), interest rates (e.g. Attari, 1999; Piazzesi, 2001), or of the various derivatives with payoffs dependent on such underlying jump processes (e.g. Merton, 1976; Ball and Torous, 1985; Carr and Wu, 2003). Most research in option pricing follows Merton by assuming that the jump component is non–system(at)ic, but more recently several authors have argued that in order to better understand the data, and in particular option prices, the system(at)ic aspect of jumps needs to be considered (see e.g. Jarrow and Rosenfeld, 1984; Naik and Lee, 1990; Schwert, 1990; Kim et al., 1994; Bakshi et al., 1997). In particular the substantially negatively skewed implied distribution of post-'87 out-of-the money put options would call for a jump-diffusion model, possibly in conjunction with stochastic volatility. Bates (2000) and Pan (2002) for instance argue that a jump-diffusion model with stochastic volatility and asymmetric jumps is a more compatible with observed option prices post-'87 than a pure stochastic volatility model. Estimated jumps are on average negative and quite substantial, embodying the fear of further crashes. In view of the existing literature, the wealth process in this paper can be viewed either as a diversified market portfolio subjected to global systemic risk, or as a more specialized portfolio facing particular risks, such as correlated default.

Our interest here is to consider the modelling of lower quantiles of returns distributions,

and how assumptions about the time scaling of such quantiles affect the reliability of the estimates. In particular, we focus on systemic risk where a rare but adverse shock hits asset prices. For this purpose, a diffusion model with downward Poisson jumps provides a useful approximation. Initially, we assume for simplicity that all wealth is wiped out which allows us to derive our results in closed form. Subsequently we allow for a partial wipe out in 5 and we show that the results are materially unaffected. Still, we do not aim to build a jump-diffusion model with the realistic market microstructure of internal day-to-day risk management, but we rather try to focus only on the catastrophic systemic events that underlie much of regulation. More realistic VaR environments with jumps can be found in Duffie and Pan (2002) for instance.

2.2 Benchmark Jump Diffusion Model

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 the portfolio value is wiped out and wealth thereafter is identically equal to zero by self-financingness and by the absence of arbitrage. These dynamics can be formally written as:

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

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. q is the Poisson process driving the jumps, 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.

This way of modelling wealth is a reduced form for the evolution of the value of a self–financing portfolio. For instance, we can view $dY_t = \theta_t dS_t - (S_t\theta_t)dq_t$, where θ is the vector portfolio process, and S is the vector price process. One could, at the expense of further complications, assume that the value of only a subset of assets held, for instance one particular desk or division of the financial institution under scrutiny, or one particular asset class held (such as defaultable bonds or OTC derivatives contracts), is subjected to such jump risk.

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:

$$Y_t = Y_0 \exp\left(\mu t + \sigma(W_t - W_0)\right) \mathbf{1}_{\{q_t=0\}}$$

where the indicator function $\mathbf{1}_{\{q_t=0\}} = 1$ if $q_t = 0$ and zero otherwise. The continuously compounded return process $X(t,T) \equiv \ln(Y_T/Y_t)$ follows directly:

$$X(t,T) = \mu(T-t) + \sigma(W_T - W_t) + \ln\big(\mathbf{1}_{\{q_T - q_t = 0\}}\big) \tag{2}$$

Conditionally on no jumps, $X(t,T)/(T-t)$ is iid normal, while conditionally on a jump, the return equals minus infinity, reflecting the total loss. Even though μ is not the drift of the arithmetic instantaneous return (which is $\mu + \frac{1}{2}\sigma^2$), it is the drift of the continuously compounded returns process X, and we shall therefore refer to μ as "drift."

We fix 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

$$X(t,t+\eta k) = \mu\eta k + \sigma(W_{t+\eta k} - W_t) + \ln \mathbf{1}_{\{q_{t+k} - q_t = 0\}}$$

The VaR(η) 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:

$$\begin{aligned}
\epsilon &= \mathbb{P}(X(t,t+\eta k) \le -\mathrm{VaR}(\eta)) \\
&= \mathbb{P}(\mu\eta k + \sigma(W_{t+\eta k} - W_t) + \ln \mathbf{1}_{\{q_{t+k} - q_t = 0\}} \le -\mathrm{VaR}(\eta)) \\
&= \mathbb{P}(\mu\eta k + \sigma(W_{t+\eta k} - W_t) \le -\mathrm{VaR}(\eta) \| q_{t+\eta k} - q_t = 0)\mathbb{P}(q_{t+\eta k} - q_t = 0) \\
&\quad + 1 \cdot \mathbb{P}(q_{t+\eta k} - q_t > 0) \\
&= \mathbb{P}\left(\frac{W_{t+\eta k} - W_t}{\sqrt{\eta k}} \le \frac{-\mathrm{VaR}(\eta) - \mu\eta k}{\sigma\sqrt{\eta k}} \Bigg\| q_{t+\eta k} - q_t = 0\right)e^{-\lambda k\eta} + (1 - e^{-\lambda k\eta}) \\
&= \Phi\left(\frac{-\mathrm{VaR}(\eta) - \mu\eta k}{\sigma\sqrt{\eta k}}\right)e^{-\lambda k\eta} + (1 - e^{-\lambda k\eta}) \tag{3}
\end{aligned}$$

We wish to solve the last equation for the unknown VaR function. For $\lambda k\eta < -\ln(1-\epsilon)$, it is immediate that,[4]

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

where $\Phi^{-1} : (0,1) \to \mathbb{R}$ is the inverse function of Φ, and therefore,

---

[4]From (3), there is a solution for VaR (as given in (4)) iff the probability of a crash is not too high, i.e. $\lambda k\eta < -\ln(1-\epsilon)$. With a λ higher than that, the probability of a Poisson crash is itself higher than ϵ, so that the joint probability of both the Poisson and the Brownian parts together being ϵ is impossible: in (3) the right-hand side would be larger than ϵ for any VaR(η), and therefore does not admit a solution. We shall therefore define VaR$(\eta) = +\infty$ for $\lambda k\eta \ge -\ln(1-\epsilon)$. But mostly we shall assume that $\lambda k\eta < -\ln(1-\epsilon)$ in what follows, which is consistent with our interpretation of the Poisson jump representing a rare systemic collapse. For future reference, we define $\bar{\eta} \equiv \bar{\eta}(\epsilon,k,\lambda) \equiv -\frac{\ln(1-\epsilon)}{\lambda k}$.

$$\text{VaR}(1) = -\sigma\sqrt{k}\Phi^{-1}\left(1 - (1 - \epsilon)e^{\lambda k}\right) - \mu k \tag{5}$$

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. 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 $\sqrt{\eta}\text{VaR}(1)$ with the true VaR number $\text{VaR}(\eta)$. In order to do so, define the absolute and relative underestimation functions:

$$f_a(\eta) = \text{VaR}(\eta) - \sqrt{\eta}\text{VaR}(1) \tag{6}$$

$$f_r(\eta) = \frac{\text{VaR}(\eta)}{\sqrt{\eta}\text{VaR}(1)} \tag{7}$$

Now it is evident 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. Indeed, it suffices to set $\lambda = 0$ in (4) and (5) to see that $f_a(\eta) = (1 - \sqrt{\eta})k\mu\sqrt{\eta} < 0$ in the realistic case of positive expected returns $\mu > 0$ and extrapolation $\eta > 1$: the *approximative* critical loss $\sqrt{\eta}\text{VaR}(1)$ is *larger* than the *actual* critical loss $\text{VaR}(\eta)$. From here we can verify when the square–root–of–time rule holds, i.e. iff returns are iid normal with zero expected returns. Setting $\lambda = 0, \mu = 0$ we see that $f(\eta) = 0$ for all $\eta \geq 0$. Alternatively, if $\lambda = 0$, then

$$\text{VaR}(\eta) = -\sigma\sqrt{k\eta}\Phi^{-1}(\epsilon)$$

which is the standard VaR formula used in the financial industry.

3 A Zero Drift

In most applications of the square–root–of–time rule to risk management, it is assumed that the drift is zero. 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 the drift, (see e.g. Merton, 1981), and any estimation of the drift term is bound to be very inaccurate. In this section we assume that the drift is zero, and in the next section extended our results to the case where the drift is positive.

It is easy to see from (4) and (5) that, in the absence of strictly positive higher–order drift terms, the square–root–of–time rule underestimates the true VaR:

Proposition 1 *Assume $\lambda > 0$ and $\mu = 0$. Then $\text{VaR}(\eta) > \sqrt{\eta}\text{VaR}(1)$ iff $\eta > 1$.*

The extent of underestimation, whether measured in absolute (f_a) or relative (f_r) terms is worse the larger the extrapolation η.

In other words, the actual critical loss at the ϵ confidence level, $\text{VaR}(\eta)$, is *larger* than the extrapolated critical loss, $\sqrt{\eta}\text{VaR}(1)$, iff we extrapolate, i.e. $\eta > 1$. This is true for any level of volatility σ and any level of confidence. For instance consider daily risk, i.e., $k = 250^{-1}$. The square–root–of–time rule recommended by the Basel Committee suggests that $\eta = 10$, which means that the VaR on which capital requirements are based is given by $\sqrt{10}\text{VaR}(1 \text{ day})$, while the correct expression is $\text{VaR}(10 \text{ days}){=}\text{VaR}(10k)$, and we just showed that $\text{VaR}(10 \text{ days}) > \sqrt{10}\text{VaR}(1 \text{ day})$.

We can see the intuition from (3). A rise in η to $\eta' > \eta$ raises the probability of a crash to $1 - e^{-\lambda k \eta'}$. Since the probability of a fall in wealth below $\text{VaR}(\eta)$ is higher conditional on a crash than conditional on no crash, the unconditional probability of losses exceeding the previous $\text{VaR}(\eta)$ rises, and it rises above ϵ. Therefore VaR will have to be revised upwards, but a rise of VaR proportional to $\sqrt{\eta}$ is not sufficient. Indeed, if VaR and $\sqrt{\eta}$ rise proportionally, only the additional Brownian risk, which arises because of the longer horizon of evaluation $\eta' - \eta$, is considered, and $\Phi\left(\frac{-\text{VaR}(\eta)}{\sigma\sqrt{\eta k}}\right)$ is constant. Therefore the unconditional probability of a loss exceeding even the new VaR scaled up by the square-root is still larger than ϵ, and in fact $\text{VaR}(\eta)/\eta$ needs to rise.

But Proposition 1 not only shows that a longer extrapolation horizon ultimately leads to underestimation, the magnitude of the error committed rises as well. In some instances, the guidance papers to the Basel Accord suggest using the rule even for quarterly periods and longer. One should be careful about such extrapolations and use conservative scaling parameters.

3.1 Rule of Thumb

The bias introduced by the square–root–of–time rule can quickly be approximated by simple rules of thumb of the form:

$$\text{VaR}(\eta) = \eta^c \text{VaR}(1) \tag{8}$$

With a jump process the VaR is given by $\text{VaR}(\eta) = -\sigma\sqrt{k\eta}\Phi^{-1}\left(1 - (1-\epsilon)e^{\lambda k\eta}\right)$, implying that in (8), $c > 1/2$. From (4) and (5) we know that $\text{VaR}(\eta) = \text{VaR}(1)$ $\sqrt{\eta}\frac{\Phi^{-1}(1-(1-\epsilon)e^{-k})}{\Phi^{-1}(1-(1-\epsilon)e^{-k})} \equiv \text{VaR}(1)\eta^c$. We find that to first order around $\eta = 1$,

$$c = \frac{1}{2} + \frac{\ln\left(\frac{\Phi^{-1}(1-(1-\epsilon)e^{-k})}{\Phi^{-1}(1-(1-\epsilon)e^{-k})}\right)}{\ln\eta} \approx \frac{1}{2} + \frac{b}{2}(1+\eta)$$

10

where b is a rapidly increasing function of the probability of a systemic crash over a one day horizon, λk, and is given by

$$b = \frac{\lambda k(1 - \epsilon) \exp\left(\lambda k + 2^{-1/2}[\Phi^{-1}(1 - (1 - \epsilon)e^{\lambda k})]^2\right)}{2^{-1/2}(-\Phi^{-1}(1 - (1 - \epsilon)e^{\lambda k}))} > 0$$

Viewed as a function of the horizon η, the square–root–of–time rule fails to scale VaR(1) sufficiently up due to the risk of a collapse, and the proper scaling is faster than the square-root by an amount that depends strongly and positively on η and λk. For instance, we get the following table for different values of λ:

λ	1/2	1/5	1/10	1/25	1/50
b	.0211638	.00765089	.00371025	.00145808	.000724831

Figure 1 illustrates the improved VaR computations when correcting the square–root–of–time rule by the approximation method laid out above. We assumed $\lambda = 1/25$, $\mu = 0$, $k = 1/250$ and $\epsilon = .01$.

4 The Benchmark Model with a Positive Drift

We argued that the zero drift assumption in the common practice of risk management is partly due to the fact that estimating μ reliably is difficult. For instance, a hypothesis of $\mu = 0$ for the S&P 500 cannot be rejected at the 10% level. Suppose however that the drift μ is strictly positive and known. Since some practitioner 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. Therefore even in that case, jumps lessen, and in most practical situations reverse, this overestimation, leading to a situation where the concerned financial institutions are potentially ill-prepared and/or undercapitalized.

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. We relegate the proof to the Appendix.

Proposition 2 *Assume that $\lambda > 0$.*

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

The degree of underestimation is worse the larger the extrapolation, η, as long as the drift term μ is not too large, $\mu \leq \underline{\mu} \equiv \underline{\mu}(\eta, k, \lambda, \sigma, \epsilon)$. We have $\bar{\mu} > \underline{\mu}$ if $\eta > 1$ and $\bar{\mu} = \underline{\mu}$ if $\eta = 1$.

11

Assume also that $\epsilon \leq 1 - \frac{1}{2}e^{-\lambda k}$. The degree of underestimation is worse the larger the likelihood of a crash, λ, no matter how large the drift μ or the horizon of extrapolation η.

The expressions for $\bar{\mu}$ and $\underline{\mu}$ can be found in the proof, and the functions are also plotted in Figures 2 and 3 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.

Proposition 2 says that as long as the drift term is small enough, the square–root–of–time rule underestimates the VaR, and this error is worse the larger the likelihood of a systemic event. Just as in Proposition 1, the larger the horizon of extrapolation, the greater will the underestimation be, at least as long as it is not counterbalanced by a strong drift term.

In order to assess the bias induced by the square–root–of–time, 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 3 that $\underline{\mu} \geq .0548$ 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.

To acquire a deeper 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, we provide a closer look at the function $\bar{\mu}(\eta, k, \lambda, \sigma, \epsilon)$ as depicted in Figure 2. The shape of the function is qualitatively unaffected by the choice of parameters k, λ, σ and ϵ (as can be verified from (10)). We see that $\bar{\mu}$ and $\underline{\mu}$ rise as the degree of extrapolation increases, and tend to infinity as $\eta \to \bar{\eta}$. The fact that the critical drifts are increasing functions of η reflects the fact that for small intervals of time, the probability of a systemic event is negligible, while the drift term, while being very small, is less so. As the horizon of extrapolation widens, the drift term scales up linearly with time while the probability of a systemic event as given by the exponential distribution inherent in Poisson processes, increases exponentially with time. Therefore, there will be, for each given drift, an η large enough for which the jump term dominates the drift term, and the square–root–of–time rule underestimates the true VaR. In particular for η close enough to $\bar{\eta}$, *no* drift can be large enough to induce overestimation, contrary to what was found above in the $\lambda = 0$ case of no jumps.

5 Near-Catastrophic Systemic Crashes

The models discussed above assume that in a crash all wealth is wiped out. That is appropriate assumption for a non diversified portfolio, where, say, the default of a single firm, country or world region, may render all equity or debt worthless. Alternatively, in a systemic crisis, even well diversified portfolios may become worthless. But more generally, we now analyze the scenario under which only part of the wealth is wiped out in a systemic event.

In the case of a partial wipe–out of wealth, the results remain qualitatively the same. We explicitly only consider the case where the drift is zero. In this case, regardless of how small the fraction of the wealth that is wiped out in a crash, an application of the square–root–of–time rule always leads to a downward bias in quantile forecasts. The reason is intutively because the algebra presented in Section 3 remains the same if the support of the continuously compounded return over the relevant horizon ηk, in the event of a systemic crash, is contained in $(-\infty, -VaR]$. This implies that if the systemic crash is severe enough, the portfolio return almost surely cannot recover losses greater than the VaR over the specified scaling horizon. However, since in actuality there is always some probability that Brownian motion moves out of the range $(-\infty, -VaR]$, we now extend our model by considering a crash occurring at a random time $\tau \in [0, \infty)$ that only wipes out a fraction $1 - \delta$ of total financial wealth. The recovery rate δ is constant and deterministic. $\delta = 0$ corresponds to full wealth wipe out in a crash. We assume that, to first order, only one systemic event occurs in the given interval of time. The possibility that more than one systemic events can happen in the time interval simply reinforces our results.

If a systemic event occurs at a time $\tau \in [t, T]$ when wealth is Y_τ, wealth drops by a fraction $1 - \delta$. Thereafter it follows the lognormal process given by the diffusion part of the stochastic differential equation (1). Thus at time $T > t$, conditional on a systemic event happening in $[t, T]$, $Y_T = Y_\tau \delta \exp(\mu(T - \tau) + \sigma(W_T - W_\tau)) = \delta Y_t \exp(\mu(T - t) + \sigma(W_T - W_t))$. Combining all possible cases, we get:

$$
\begin{aligned}
Y_T &= Y_t \exp\left(\mu(T - t) + \sigma(W_T - W_t)\right) \mathbf{1}_{\{q_T - q_t = 0\}} \\
&\quad + \delta Y_t \exp\left(\mu(T - t) + \sigma(W_T - W_t)\right) \mathbf{1}_{\{q_T - q_t > 0\}} \\
&= Y_t \exp\left(\mu(T - t) + \sigma(W_T - W_t)\right) \left[\mathbf{1}_{\{q_T - q_t = 0\}} + \delta \mathbf{1}_{\{q_T - q_t > 0\}}\right]
\end{aligned}
$$

with a returns process of

$$
X(t, t + \eta k) = \mu \eta k + \sigma(W_{t+\eta k} - W_t) + \ln[\delta + (1 - \delta)\mathbf{1}_{\{q_T - q_t = 0\}}]
$$

VaR is then the solution to

$$
\epsilon = \Phi\left(\frac{-\mathrm{VaR}(\eta) - \mu \eta k}{\sigma \sqrt{\eta k}}\right) e^{-\lambda k \eta} + \Phi\left(\frac{-\mathrm{VaR}(\eta) - \ln \delta - \mu \eta k}{\sigma \sqrt{\eta k}}\right)(1 - e^{-\lambda k \eta}) \quad (9)
$$

This solution always exists and is unique and continuously differentiable. As before, we need to show that $\mathrm{VaR}(\eta) > \sqrt{\eta}\mathrm{VaR}(1)$ and that $\mathrm{VaR}(\eta) > \mathrm{VaR}(\eta')$ when $\eta > \eta'$, i.e. that $\nu(\eta) \equiv \frac{\mathrm{VaR}(\eta)}{\sqrt{\eta}} > \nu(1) = \mathrm{VaR}(1)$ and that ν is increasing in η, for $\eta > 1$ in a relevant range. In differential form, this amounts to showing that $\frac{\partial \nu(\eta)}{\partial \eta} > 0$. The following Proposition is the extension of Proposition 1 for an economy in which systemic events can have various degrees of severity $\delta \geq 0$. However, since there is no explicit solution for VaR, and since the relation defined by (9) involves essentially non–algebraic implicit functions, definite results can only be shown under more restrictive, but plausible and realistic, assumptions:

Proposition 3 *Assume $\mu = 0$ and $\lambda > 0$. There is a $\underline{\eta} > 1$ such that $\mathrm{VaR}(\eta)/\sqrt{\eta}$ is increasing in η, and in particular $\mathrm{VaR}(\eta) > \sqrt{\eta}\mathrm{VaR}(1)$ for $\eta \in (1, \underline{\eta})$, if any one of the following assumptions applies:*

(i) $\delta < \underline{\delta}$, for some $\underline{\delta} \in (0, 1)$.

(ii) $\delta > \bar{\delta}$, for some $\bar{\delta} \in (0, 1)$, and $\lambda k \eta \leq 1$.

What Proposition 3 says is that for recovery rates δ close enough to either 0 or 1, the square–root–of–time rule underestimates the true VaR, and does so increasingly poorly for larger η. The result for the case where δ is close to 0 follows by continuity from the benchmark case $\delta = 0$. The case δ close to 1 is shown by a separate argument. Extensive numerical simulations performed by the authors suggest that these results are valid much more generally and seem to hold for all values of $\delta \in [0, 1)$, as can be seen in Figures 7 and 8. Looking at those two figures, further discussed below, it becomes clear that the relative risk errors $f_r(\eta) = \mathrm{VaR}(\eta)/(\sqrt{\eta}\mathrm{VaR}(1))$ arising from the square–root–of–time rule are roughly the same for any potential losses above 25%, and that for all levels of λ and all reasonable extrapolation horizons η.

6   Discussion

Our results indicate that applying the square–root–of–time rule to the time scaling of quantiles results in a downward bias of risk forecasts, provided the scaling horizon is sufficiently large. Intuitively, the explanation lies in how the three key factors, Brownian motion, drift, and Poisson process behave under time scaling. Roughly speaking, the Brownian part increases at the rate $\sqrt{t}$, with the drift growing at a linear rate t, and the potentiality of the jump component growing exponentially. This suggests that the jump component eventually dominates.

6.1 Numerical Simulations

The key question must be how relevant this is in practice. In order to assess the bias induced by the square–root–of–time, we perform numerical simulations where the drift and volatility are calibrated to historical annualized values of the S&P–500 index, i.e., 5.48% and 15.84%, respectively.

6.1.1 Zero Drift and Full Wealth Wipe Out

In the first set of numerical simulations we follow standard practice in the risk management industry and assume the drift is zero. Table 1 shows the VaR from using the square–root–of–time rule ($\sqrt{\eta}\text{VaR}(1)$) as well as the correct VaR ($\text{VaR}(\eta)$), while Figure 4 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. The reason is that as the probability of a crash in a given time period goes to zero, so does the importance of the jump component. 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. For λ and η both large, there is a region in which the probability of the systemic event approaches ϵ. In that region the true VaR tends to infinity, implying that the relative error tends to infinity as well.

6.1.2 Positive Drift and Full Wealth Wipe Out

If the drift is positive the bias is less severe. While for large enough η the square–root–of–time rule always leads to a downward bias in risk estimates, this may not be the case if η is small since in that case the Poisson term is dominated by the drift term. Consider Figure 2 where we plot the critical drift value, $\bar{\mu}$ as the holding period, η, increases. The region beneath the curve is the region in which the square–root–of–time rule underestimates the true VaR. At the 10–day level, the critical drift term equals 7%, which roughly translates into a number above the mean real rate of return of the index (5.5%). In other words, during the periods with average or lower than average expected returns, the square–root–of–time rule underestimates the true VaR in the particular case where $\eta = 10$, and conversely for the periods with expected returns on the higher end of the historical scale. Results are however more clear-cut for η larger than a month. Indeed, the Poisson term dominates the drift term for reasonable parameter values, even though we chose λ to be quite small in Figure 2.

This can also be seen in Figure 5, which is identical to Figure 4, 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 above 1 almost everywhere, but it is below 1 for holding periods of around 10 days and expected years to a systemic event of above 25 years. As the holding periods increase, the ratio exceeds 1 at an increasing rate. 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.

6.1.3 Partial Wealth Wipe Out

Finally, we consider the case where only part of the wealth is wiped out. Figures 7 and 8 show the relative error of the square–root–of–time rule as we change the fraction of wealth wiped out, the holding period, and the crash frequency. Interestingly, as the wiped out level $(1 - \delta)$ increases from zero, there is first a rapid increase in the square–root–of–time rule error, but the full effect is reached fairly soon, in all cases before the loss factor $1 - \delta$ exceeds 20%. This is why we can claim that the predictions of our model carry over to more realistic environments with partial crashes. For reasonable parameter values and levels of extrapolation, the *relative error* with a 25% wipeout is qualitatively the same as with a complete wipeout.

In Table 2 we consider 1% VaR where 25% of wealth is wiped out on average every 25 years, where the holding periods range from 10 to 60 days. The error introduced by the square–root–of–time rule increases along with the holding periods, with the ratio of the true VaR to the time scaled VaR increasing from 1.02 to 1.42, respectively.

As similar picture emerges in Table 3, where we consider a similar event, for a 25 day holding period, where the probability level ranges from 5% to 0.004%. The relative error introduced by the square–root–of–time rule increases from 1.02 to 1.62.

7   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. While this may be a relatively innocuous assumption for many applications, it can be a critically flawed assumption in other cases, such as in pricing and regulation in market environments prone to systemic events. The time scaling of risk has emerged as an important tool in the calculation of regulatory capital (it is all but required by the Basel Capital Accords (Basel Committee on Banking Supervision, 1996, 1998)), and sees an important role in the regulation of pension funds.

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. We feel that jump diffusions are an appropriate methodology for capturing the risk of systemic events, which while rare are large in magnitude. They are also instrumental in understanding the smile and smirk effects in implied option volatilities (refer for instance to Pan (2002) or Carr and Wu (2003)), which are oftentimes argued to reflect the market's anticipation of such crashes.

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, 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, i.e. the systemic part of the market risk. We therefore believe the square root of time to be fundamentally flawed, not because it is false per se, but because it is systematically biased and fails, depending on the application, to account for the exact events that it set out to model in the first place, namely the potential for catastrophic or near-catastrophic one-off events. There are indications that the Basel Committee is in fact beginning to realize such problems, for in a more recent technical guidance paper (Basel Committee on Banking Supervision, 2002) it is no longer suggested that the square–root–of–time rule be used, but that "in constructing VaR models estimating potential quarterly losses, institutions may use quarterly data or convert shorter period data to a quarterly equivalent using an *analytically appropriate method supported by empirical evidence*" (emphasis added).

In an exercise of revealed preferences, the anecdotal evidence that very few major financial institutions seem to have ever violated the square-root of time scaled Basel constraints may indicate, alongside alternative reputational considerations, that they too are aware of the fact that the mechanism underlying those regulations biases the VaR downwards by not properly accounting for the "big ones."

A Appendix: Proofs

Proof of Proposition 2

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

$$\sigma \sqrt{\eta k} \left[\Phi^{-1}(1 - (1 - \epsilon)e^{\lambda k}) - \Phi^{-1}(1 - (1 - \epsilon)e^{\lambda k \eta}) \right] - \mu k \sqrt{\eta}(\sqrt{\eta} - 1) \geq 0$$

i.e. iff

$$\mu \leq \bar{\mu}(\eta, k, \lambda, \sigma, \epsilon) \equiv \frac{\sigma \left[\Phi^{-1}(1 - (1 - \epsilon)e^{\lambda k}) - \Phi^{-1}(1 - (1 - \epsilon)e^{\lambda k \eta}) \right]}{\sqrt{k}(\sqrt{\eta} - 1)} \tag{10}$$

The proof of monotonicity of f_r and f_a with respect to η, the second assertion, goes along similar lines, denoting the standard normal density by ϕ:

$$\frac{\partial f_r(\eta)}{\partial \eta} > 0$$

$$\Leftrightarrow \frac{\partial \text{VaR}(\eta)}{\partial \eta} > \frac{1}{2}\eta^{-1/2}\text{VaR}(1)$$

$$\Leftrightarrow \mu < \frac{2\sigma\sqrt{k\eta}(1-\epsilon)\lambda e^{\lambda k\eta}}{\phi(\Phi^{-1}(1-(1-\epsilon)\lambda e^{\lambda k\eta}))} \equiv \underline{\mu}$$

It is then easy to see that the monotonicity of f_r implies the monotonicity of f_a.

As to the last assertion,

$$\frac{\partial(\text{VaR}(\eta) - \sqrt{\eta}\text{VaR}(1))}{\partial \lambda} = \sigma\sqrt{k\eta}\frac{(\epsilon-1)ke^{\lambda k}}{\phi(\Phi^{-1}(1-(1-\epsilon)e^{\lambda k}))}$$
$$- \sigma\sqrt{k\eta}\frac{(\epsilon-1)k\eta e^{\lambda k\eta}}{\phi(\Phi^{-1}(1-(1-\epsilon)e^{\lambda k\eta}))}$$

The partial derivative is strictly positive iff

$$\frac{e^{\lambda k\eta}\eta}{\phi(\Phi^{-1}(1-(1-\epsilon)e^{\lambda k\eta}))} > \frac{e^{\lambda k}}{\phi(\Phi^{-1}(1-(1-\epsilon)e^{\lambda k}))}$$
$$\Leftrightarrow 2\ln\eta + 2\lambda k(\eta-1) - [\Phi^{-1}(1-(1-\epsilon)e^{\lambda k})]^2 + [\Phi^{-1}(1-(1-\epsilon)e^{\lambda k\eta})]^2 > 0$$

The first two terms are positive, while the last two terms taken together are positive if $\epsilon \leq 1 - \frac{1}{2}e^{-\lambda k}$, a sufficient condition of which is to require $\epsilon \leq \frac{1}{2}$. 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 $\frac{1}{2}$. ∎

Proof of Proposition 3

We want to show, at least for reasonably short horizons and for reasonable values of δ, that $\nu(\eta) \equiv \text{VaR}(\eta)/\sqrt{\eta}$ is larger than $\nu(0)$. Furthermore, in order to show that underestimation worsens with η, we'd like to show that $\nu(\eta)$ is increasing in η.

Now ν solves the following equation, modified from equation (9):

$$F(\nu, \eta, ; \cdot) \equiv \Phi\left(\frac{-\nu}{\sigma\sqrt{k}}\right)e^{-\lambda k\eta} + \Phi\left(\frac{-\nu}{\sigma\sqrt{k}} - \frac{\ln(\delta)}{\sigma\sqrt{\eta k}}\right)(1 - e^{-\lambda k\eta}) - \epsilon = 0 \qquad (11)$$

For simplicity, denote $\xi \equiv \frac{-\nu}{\sigma\sqrt{k}} - \frac{\ln(\delta)}{\sigma\sqrt{\eta k}}$. The strategy of the proof is as follows. Assume that $\eta = 1$ and show that there exist values of δ for which the statements hold, and then infer that the same must hold in neighbourhoods of those δ.

The first and obvious reference δ is to choose $\delta = 0$, and the following lemma confirms the importance of the benchmark economy for it shows that the VaR in the economy with partial losses converges to the economy with total loss as losses worsen:

Lemma 4 $\lim_{\delta \to 0} \nu(\eta; \delta) = -\sigma\sqrt{k}\Phi^{-1}(1 - (1-\epsilon)e^{\lambda k\eta})$, $\forall(\eta, \lambda, k, \epsilon)$ such that $\lambda k\eta \leq -\ln(1 - \epsilon)$.

Proof :
First of all, we show that $\lim_{\delta \to 0} \nu(\cdot; \delta) < \infty$. We argue by contradiction. Assume that $\lim_{\delta \to 0} \nu(\cdot; \delta) = \infty$. Then in the limit we get from (9) that $0 + \Phi(\cdot)(1 - e^{-\lambda k\eta}) = \epsilon$. Since we assume that $\lambda k\eta \leq -\ln(1 - \epsilon)$, $\epsilon > 1 - e^{-\lambda k\eta}$, and since $\Phi \in [0, 1]$, we get a contradiction. Therefore $\lim_{\delta \to 0} \xi = \lim_{\delta \to 0} \frac{-\nu}{\sigma\sqrt{k}} - \frac{\ln(\delta)}{\sigma\sqrt{\eta k}} = +\infty$, and the statement follows from (11). ∎

We can conclude, by the continuity of ν, that the square–root–of–time rule underestimates the true VaR for all δ sufficiently close to 0, and for all η satisfying the condition in the statement of the previous lemma.

From the implicit function theorem applied to (9) we find that $\frac{\partial\nu(\eta)}{\partial\eta} > 0$ iff

$$
\begin{aligned}
&G(\eta; \delta, \cdot) \equiv \\
&\Phi\left(\frac{-\nu(\cdot)}{\sigma\sqrt{k}} - \frac{\ln(\delta)}{\sigma\sqrt{\eta k}}\right) - \epsilon + \frac{1}{2}\frac{\ln(\delta)}{\lambda\sigma(k\eta)^{3/2}}(1 - e^{-\lambda k\eta})\phi\left(\frac{-\nu(\cdot)}{\sigma\sqrt{k}} - \frac{\ln(\delta)}{\sigma\sqrt{\eta k}}\right) > 0
\end{aligned}
$$

It can easily be seen that $G > 0$ for all δ sufficiently close to 0, since $\lim_{\delta \to 0} G = 1 - \epsilon$. Indeed, from the previous lemma we know that $\lim_{\delta \to 0} \nu < \infty$, and therefore that $\lim_{\delta \to 0} \xi = +\infty$, and so the first two terms in G converge to $1 - \epsilon$. The last term in G can be written as $-\alpha \ln(\delta)\frac{1}{\sqrt{2\pi}}\exp(-\frac{\nu}{\sigma\sqrt{k}} - \frac{\ln(\delta)}{\sigma\sqrt{\eta k}})^2$, and we now show that it converges to zero. Dropping constants on the way, $\lim_{\delta \to 0} -\ln(\delta)\exp(-[\nu^2 + 2\ln(\delta)\nu + (\ln(\delta))^2])$ $\leq \lim -\ln(\delta)\exp(-2\ln(\delta)\nu - (\ln(\delta))^2) \leq \lim -\ln(\delta)\exp(-2\ln(\delta)\bar{\nu} - (\ln(\delta))^2$, since ν is bounded in δ by some $\bar{\nu}$. Therefore, by Hotelling's rule, $\lim \frac{-\ln(\delta)}{\exp((\ln(\delta))^2 + 2\ln(\delta)\bar{\nu})}$ $=\lim \frac{1}{2[-\ln(\delta) - \bar{v}]\exp((\ln(\delta))^2 + 2\ln(\delta)\bar{\nu})} = 0$. This tells us, for δ sufficiently close to 0, that an increase in η leads to an increase in $\nu(\eta)$, i.e. to a worsening of the underestimation.

Ideally, we'd like to show monotonicity for all δ, but the essentially non-algebraic nature of the problem does not allow one to do so. But we can single out the other end of the spectrum as well. Now assume that $\delta = 1$. Then we know that $\nu(\eta) = \nu(1)$, all $\eta \geq 1$. Denoting $K \equiv \delta^{-1}$, we now show that $\partial G/\partial K > 0$ evaluated at $K = 1$ iff $e^{\lambda k\eta} < 2\eta k\lambda + 1$.

Denote $\alpha \equiv \frac{1 - e^{-k}}{2\lambda\sigma(k\eta)^{3/2}}$, in which case $G(K) = \Phi(\xi(K)) - \epsilon - \alpha\ln(K)\phi(\xi(K))$. From the

implicit function theorem,

$$\frac{d\nu}{dK} = \frac{(1 - e^{-\lambda k \eta})\phi(\xi)}{K\sqrt{\eta}\left[e^{-\lambda k \eta}\phi(-\frac{\nu}{\sigma\sqrt{k}}) + (1 - e^{-\lambda k \eta})\phi(\xi)\right]}$$

so that

$$\left.\frac{d\xi}{dK}\right|_{K=1} = \frac{e^{-\lambda k \eta}}{\sigma\sqrt{\eta k}}$$

Inserting these expressions into

$$\frac{dG}{dK} = \phi(\xi)\left[\frac{d\xi}{dK}[1 + \alpha\ln(K)\xi] - \frac{\alpha}{K}\right]$$

we get

$$\left.\frac{dG}{dK}\right|_{K=1} = \phi\left(-\frac{\nu}{\sigma\sqrt{k}}\right)\left[\frac{e^{-\lambda k \eta}}{\sigma\sqrt{k\eta}} - \alpha\right] > 0$$

iff $e^{\lambda k \eta} < 2\eta k \lambda + 1$.

Since $G(\eta; 1) = 0$, $\forall \eta \geq 1$, and G is decreasing in δ, $G(1; \delta) > 0$ for some $\delta < 1$. Now as η increases from 1, $\frac{\partial \nu(\eta)}{\partial \eta} > 0$ and therefore $\nu(\eta) > \nu(1)$. In addition, slightly raising η further will in turn increase ν further since G is still strictly positive, and so forth. This shows that for δ in a neighbourhood of 1, the square–root–of–time rule is underestimating the true VaR, at least for reasonable extrapolation horizons, $e^{\lambda k \eta} < 2\eta k \lambda + 1$, a sufficient condition being $\lambda k \eta \leq 1$.

B   Appendix: Tables and Figures

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

Table 3: 1% VaR on a $1000 portfolio with S&P–500 volatility, varying ϵ

VaR on a $1000 portfolio of S&P–500 Volatility. $\eta = 25$, $\mu = 0$, $\sigma = 0.1584$, $k = 0.004$. $1/\lambda = 25$ indicates the expected time of crash in years, and $1 - \delta = 25\%$ indicates the fraction of wealth wiped out in crash. Ratio is $\text{VaR}(\eta)/(\sqrt{\eta}\text{VaR}(1))$.

Measure	Quantile probability ϵ		
	0.05	0.01	0.004
VaR(25)	84.3	125.7	216.1
$\sqrt{25}$VaR(1)	82.5	116.8	133.5
Ratio	1.02	1.08	1.62

Figure 1: Absolute error $\mathrm{VaR}(\eta) - \eta^c \mathrm{VaR}(1)$

Notes: This graph depicts the absolute errors arising from the use of scaling approximations to $\mathrm{VaR}(\eta)$. The top curve represents the errors using the square–root–of–time rule ($c = 1/2$), the bottom curve represents the errors arising with the rule-of-thumb correction ($c = \frac{1}{2} + \frac{b}{2}(1+\eta)$) mentioned in the text.
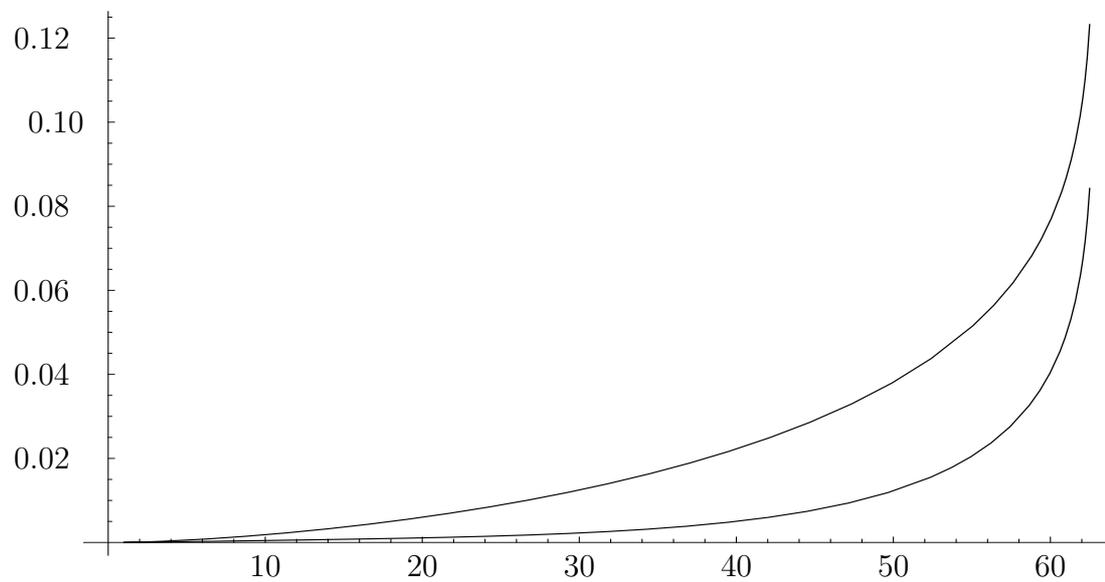
Figure 2: The critical drift function $\bar{\mu}$ as a function of η

$\sigma = .1584, \delta = 0, k = 1/250, \epsilon = 1/250, \lambda = 1/55.$
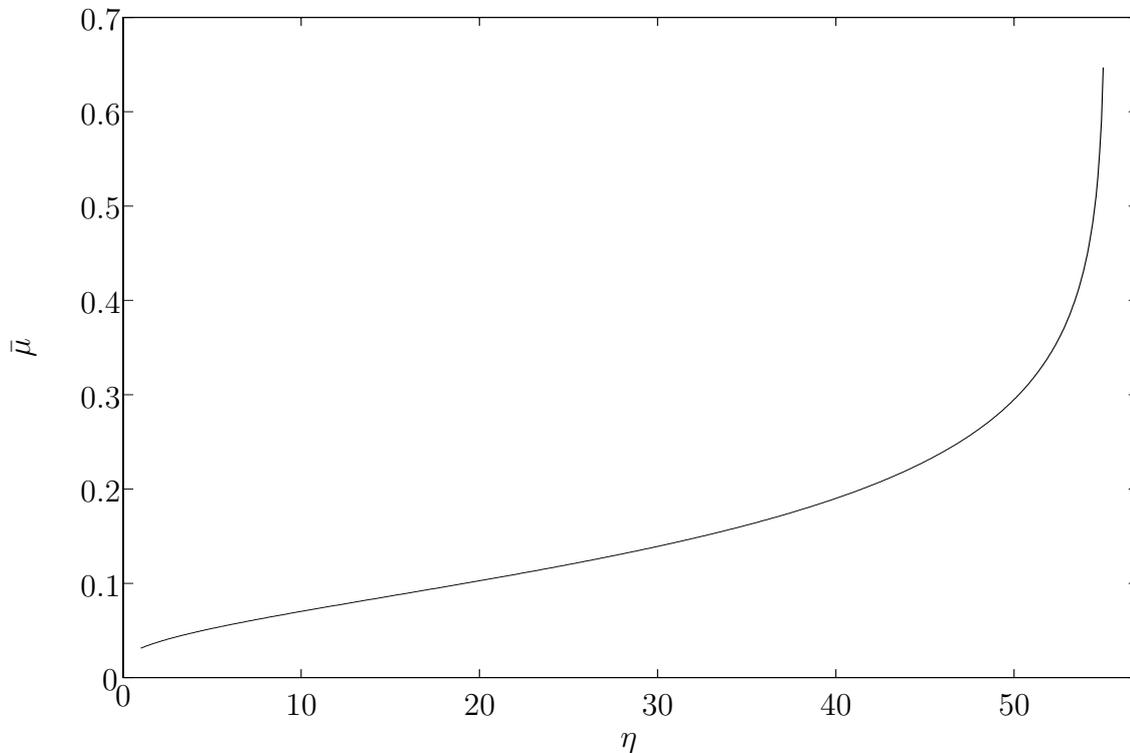
Figure 3: The critical drift functions $\underline{\mu}$ (bottom) and $\bar{\mu}$ (top) as functions of η.
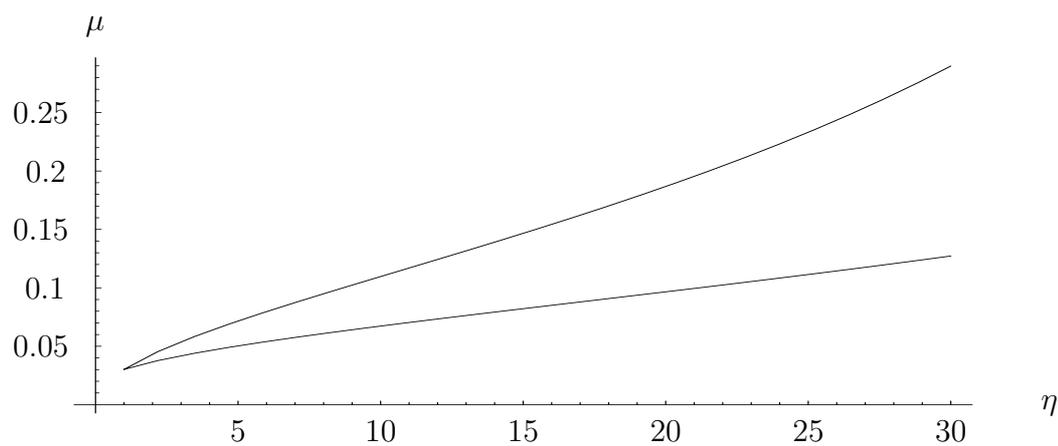
Figure 4: Relative error in the square–root–of–time rule, S&P-500 annual volatility, zero mean, full wealth wipeout, and 1% VaR

$\mu = 0, \sigma = .1584, \delta = 0, k = 1/250, \epsilon = 0.01$. Ratio is $\mathrm{VaR}(\eta)/(\sqrt{\eta}\mathrm{VaR}(1))$. $1/\lambda$ indicates the expected time of crash in years.
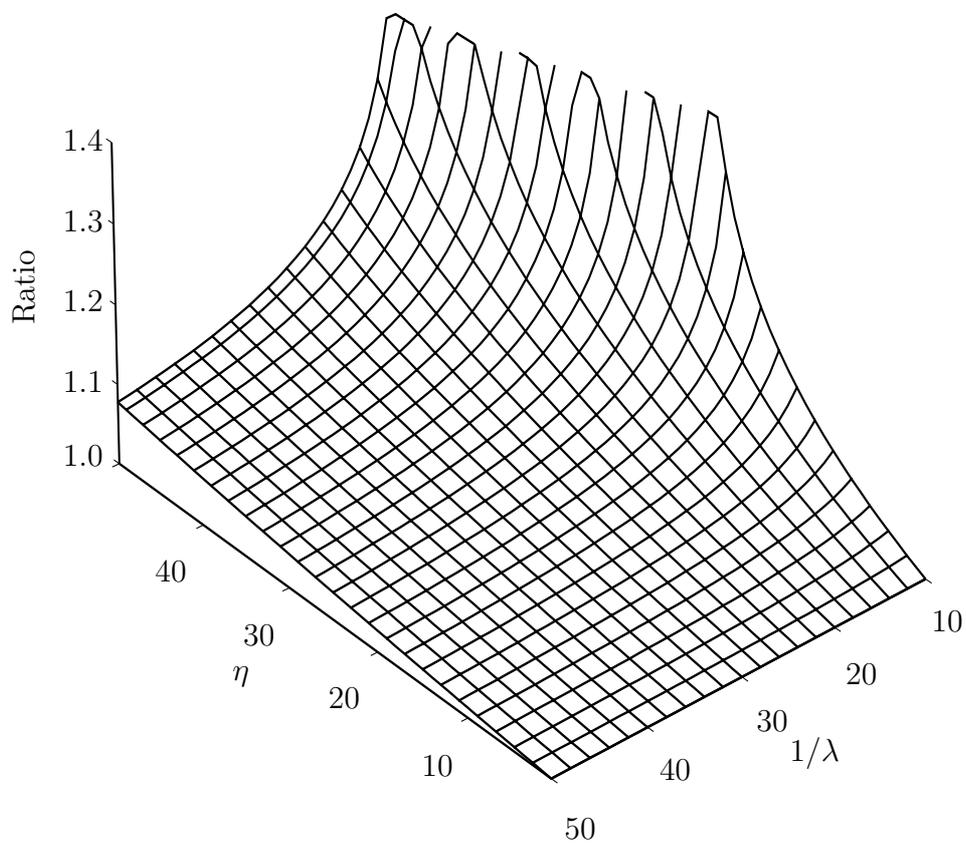
Figure 5: Relative error in the square–root–of–time rule, S&P-500 annual volatility and mean, full wealth wipeout, and 1% VaR

$\mu = 0.0548$, $\sigma = .1584$, $\delta = 0$, $k = 1/250$, $\epsilon = 0.01$. Ratio is $\text{VaR}(\eta)/(\sqrt{\eta}\text{VaR}(1))$. $1/\lambda$ indicates the expected time of crash in years.
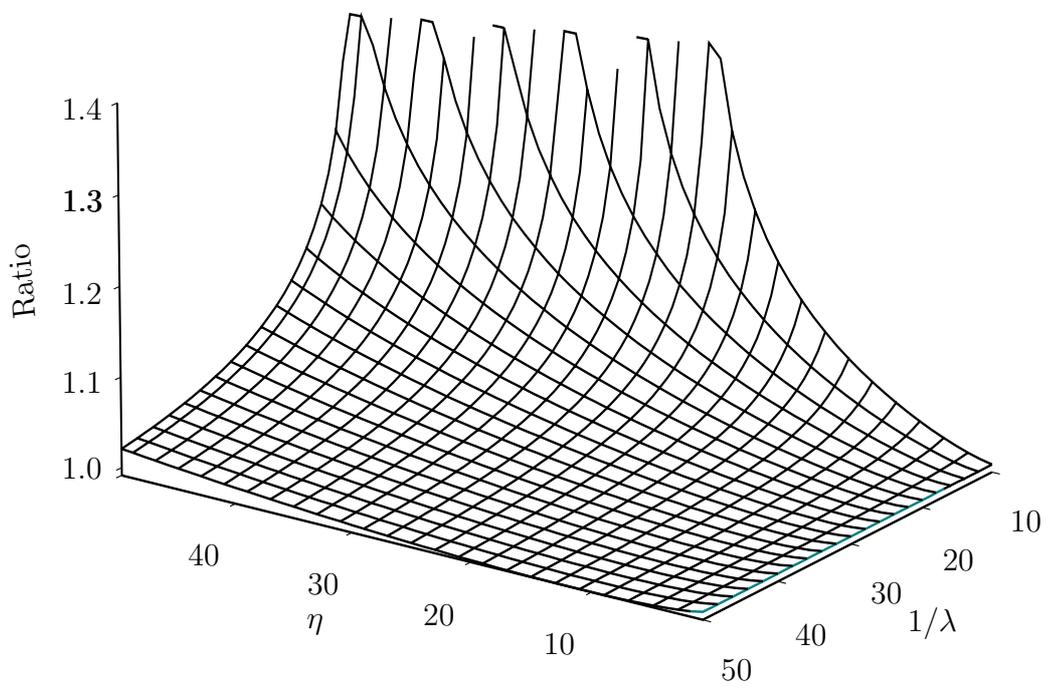
Figure 6: Relative error in the square–root–of–time rule, S&P-500 annual volatility and mean, full wealth wipeout, and 1% 10 day VaR

$\mu = 0.0548, \sigma = .1584, \delta = 0, k = 1/250, \epsilon = 0.01, \eta = 10.$ $1/\lambda$ indicates the expected time of crash in years.
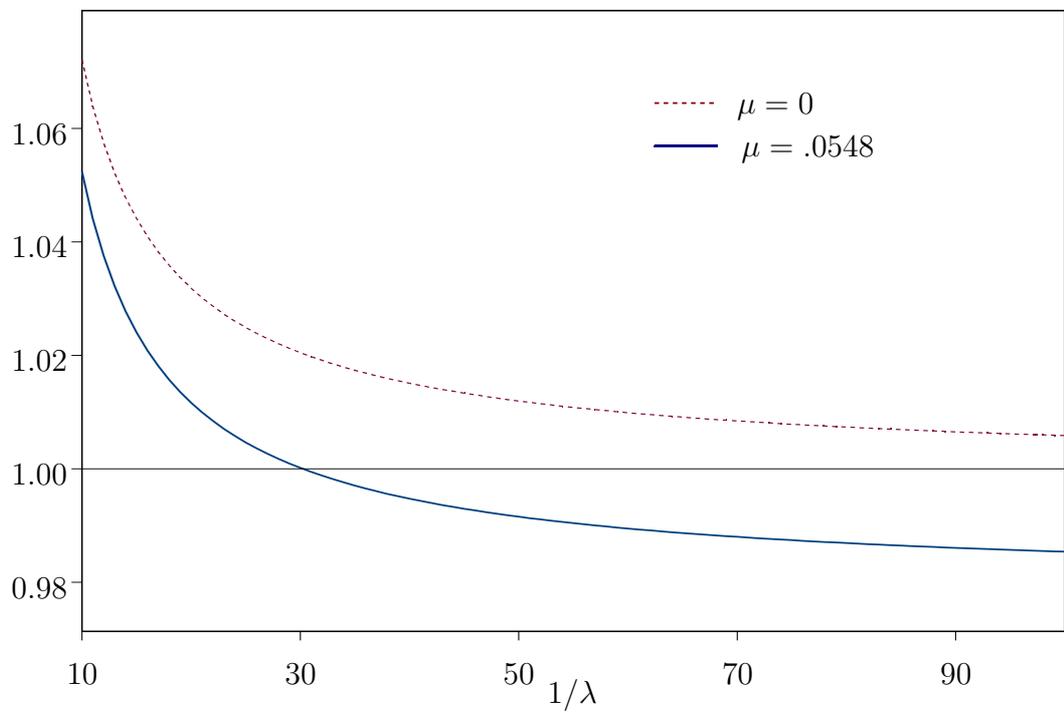
Figure 7: Relative error in the square–root–of–time rule, S&P-500 annual volatility, zero mean, partial wealth wipeout, and 1% 10 day VaR

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.
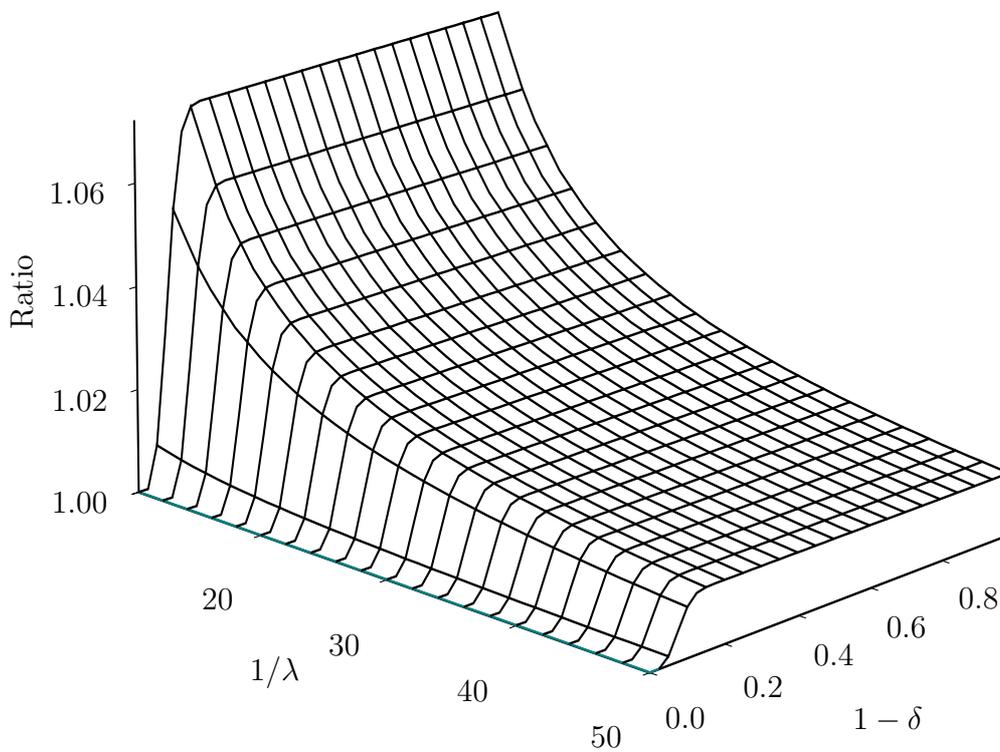
Figure 8: Relative error in the square–root–of–time rule, S&P-500 annual volatility, zero mean, partial wealth wipeout, and 1% VaR

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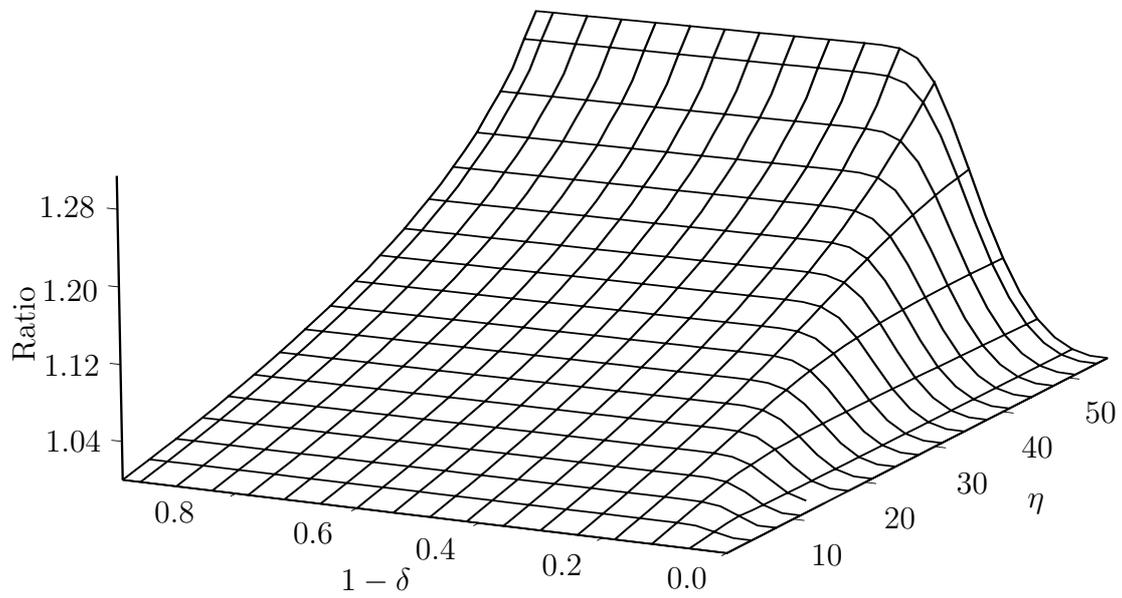
References

Ahn, D., Boudukh, J., Richardson, M., and Whitelaw, R. (1999). Optimal risk management using options. *Journal of Finance*, 54(1):359–375.

Akgiray, V. and Booth, G. (1988). Mixed diffusion-jump process modeling of exchange rate movements. *Review of Economics and Statistics*, 70(4):631–637.

Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.

Artzner, P., Delbaen, F., Eber, J., and Whitelaw, R. (1999). Coherent measure of risk. *Mathematical Finance*, 9(3):203–228.

Attari, M. (1999). Discontinuous interest rate processes: An equilibrium model for bond option prices. *Journal of Financial and Quantitative Analysis*, 34(3):293–322.

Bakshi, G., Cao, C., and Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, 52:2003–2049.

Ball, C. and Torous, W. (1983). A simplified jump process for common stock returns. *Journal of Financial and Quantitative Analysis*, 18:53–65.

Ball, C. and Torous, W. (1985). On jumps in common stock prices and their impact on call option pricing. *Journal of Finance*, 40:155–173.

Basak, S. and Shapiro, A. (2001). Value-at-risk based risk management: Optimal policies and asset prices. *Review of Financial Studies*, 8:371–405.

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

Basel Committee on Banking Supervision (1998). Framework for supervisory information about derivatives and trading activities. Technical report.

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*.

Crockett, A. (2000). Marrying the micro- and macro prudential dimensions of financial stability. Mimeo, Bank of International Settlements, http://www.bis.org/review/rr000921b.pdf.

Dacorogna, M., Müller, U., Pictet, O., and de Vries, C. (1999). Extremal forex returns in extremely large datasets. Mimeo.

Danielsson, J., Shin, H. S., and Zigrand, J.-P. (2003). The impact of risk regulation on price dynamics. *Forthcoming, Journal of Banking and Finance.*

Danielsson, J. and Zigrand, J.-P. (2001). What happens when you regulate risk? Evidence from a simple equilibrium model. FMG Discussion Paper 393.

Diamond, D. and Rajan, R. (2001). Liquidity shortages and banking crises. Mimeo, University of Chicago.

Drost, F. C. and Nijman, T. E. (1993). Temporal aggregation of GARCH processes. *Econometrica*, 61:909–927.

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

Duffie, D. and Singleton, K. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720.

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

Fama, E. (1965). The behaviour of stock-market prices. *Journal of Business*, 38(1):34–105.

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

Jarrow, R. and Rosenfeld, E. (1984). Jump risks and the intertemporal capital asset pricing model. *Journal of Business*, 57:337–351.

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

Johnson, G. and Schneeweis, T. (1994). Jump-diffusion processes in the foreign exchange markets and the release of macroeconomic news. *Computational Economics*, 7(4):427–445.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. *Review of Financial Studies*, 1:427–445.

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.

Kim, M., Oh, Y., and Brooks, R. (1994). Are jumps in stock returns diversifiable? Evidence and implications for option pricing. *Journal of Financial and Quantitative Analysis*, 29(4):609–631.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36:394–419.

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

Merton, R. (1976). Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

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

Naik, V. and Lee, M. (1990). General equilibrium pricing of options on the market portfolio with discontinuous returns. *Review of Financial Studies*, 3(4):493–521.

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.

Piazzesi, M. (2001). An econometric model of the yield curve with macroeconomic jump effects. Technical report, NBER working paper no. 8246.

Schwert, W. (1990). Stock volatility and the crash of 1987. *Review of Financial Studies*, 3:77–102.