Chapter 9

Risk Measures

Reference: Hull (2009) 20; McDonald (2014) 31; McNeil, Frey, and Embrechts (2005); Alexander (2008)

9.1 Symmetric Dispersion Measures

9.1.1 Mean Absolute Deviation

The variance (and standard deviation) is very sensitive to the tails of the distribution. For instance, even if the standard normal distribution and a student-t distribution with 4 degrees of freedom look fairly similar, the latter has a variance that is twice as large (recall: the variance of a t_n distribution is $n/(n-2)$ for $n > 2$). This may or may not be what the investor cares about. If not, the mean absolute deviation is an alternative. Let μ be the mean, then the definition is

$$\text{mean absolute deviation} = \mathrm{E}\,|R - \mu|. \tag{9.1}$$

This measure of dispersion is much less sensitive to the tails—essentially because it does not involve squaring the variable.

Notice, however, that for a normally distributed return the mean absolute deviation is proportional to the standard deviation—see Remark 9.1. Both measures will therefore lead to the same portfolio choice (for a given mean return). In other cases, the portfolio choice will be different (and perhaps complicated to perform since it is typically not easy to calculate the mean absolute deviation of a portfolio).

Remark 9.1 *(Mean absolute deviation of $N(\mu,\sigma^2)$ and t_n) If $R \sim N(\mu,\sigma^2)$, then*

$$\mathrm{E}\,|R-\mu| = \sqrt{2/\pi}\sigma \approx 0.8\sigma.$$

If $R \sim t_n$, then $\mathrm{E}\,|R| = 2\sqrt{n}/[(n-1)B(n/2, 0.5)]$*, where B is the beta function. For* $n = 4$, $\mathrm{E}\,|R| = 1$ *which is just 25% higher than for a $N(0,1)$ distribution. In contrast, the standard deviation is* $\sqrt{2}$*, which is 41% higher than for the* $N(0,1)$.

9.1.2 Index Tracking Errors

Suppose instead that our task, as fund managers, say, is to track a benchmark portfolio (returns R_b and portfolio weights w_b)—but we are allowed to make some deviations. For instance, we are perhaps asked to track a certain market index. The deviations, typically measured in terms of the variance of the tracking errors for the returns, can be motivated by practical considerations and by concerns about trading costs. If our portfolio has the weights w, then the portfolio return is $R_p = w'R$, where R are the original assets. Similarly, the benchmark portfolio (index) has the return $R_b = w_b'R$. If the variance of the tracking error should be less than U, then we have the restriction

$$\mathrm{Var}(R_p - R_b) = (w - w_b)'\Sigma(w - w_b) \leq U, \tag{9.2}$$

where Σ is the covariance matrix of the original assets. This type of restriction is fairly easy to implement numerically in the portfolio choice model.

9.2 Downside Risk

9.2.1 Value at Risk

The mean-variance framework is often criticized for failing to distinguish between downside of the return distribution (considered to be risk) and upside (considered to be potential). The Value at Risk is one way of focusing on the downside.

Remark 9.2 *(Quantile of a distribution) The 0.05 quantile is the value such that there is only a 5% probability of a lower number,* $\Pr(R \leq \text{quantile}_{0.05}) = 0.05$.

The 95% Value at Risk ($\mathrm{VaR}_{95\%}$) is a number such that there is only a 5% chance that the loss ($-R$) is larger that $\mathrm{VaR}_{95\%}$

$$\Pr(-R \geq \mathrm{VaR}_{95\%}) = 5\%. \tag{9.3}$$

Here, 95% is the confidence level of the VaR. For instance, if $\mathrm{VaR}_{95\%} = 18\%$, then we are 95% sure that we will not lose more than 18% of our investment. To convert the value

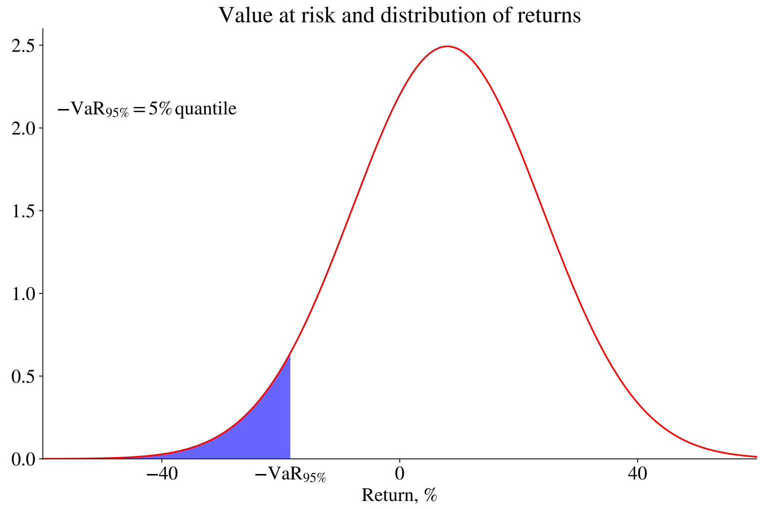

Figure 9.1: Value at risk

at risk into value terms (CHF, say), just multiply the VaR for returns with the value of the investment (portfolio).

Clearly, $-R \geq \mathrm{VaR}_{95\%}$ is true when (and only when) $R \leq -\mathrm{VaR}_{95\%}$, so (9.3) can be written

$$\Pr(R \leq -\mathrm{VaR}_{95\%}) = 5\%. \tag{9.4}$$

This says that $-\mathrm{VaR}_{95\%}$ is a number such that there is only a 5% chance that the return is below it. That is, $-\mathrm{VaR}_{95\%}$ is the 0.05 quantile (5th percentile) of the return distribution. Using (9.4) allows us to work directly with the return distribution (not the loss distribution), which is often convenient. See Figure 9.1 for an illustration. If the return is normally distributed, $R \sim N(\mu, \sigma^2)$ then

$$\mathrm{VaR}_{95\%} = -(\mu - 1.64\sigma). \tag{9.5}$$

Example 9.3 *(VaR with $R \sim N(\mu, \sigma^2)$) If daily returns have $\mu = 8\%$ and $\sigma = 16\%$, then the 1-day $VaR_{95\%} = -(0.08 - 1.64 \times 0.16) \approx 0.18$; we are 95% sure that we will not lose more than 18% of the investment over one day, that is, $VaR_{95\%} = 0.18$.*

More generally, we can consider the confidence level α instead of just 0.95, so

$$\Pr(R \leq -\mathrm{VaR}_\alpha) = 1 - \alpha, \text{ so} \tag{9.6}$$

$$\mathrm{VaR}_\alpha = -[(1-\alpha)^{th} \text{ quantile of } R]. \tag{9.7}$$

In particular, if the return is normally distributed, $R \sim N(\mu, \sigma^2)$, then

$$\text{VaR}_\alpha = -(\mu + c\sigma), \tag{9.8}$$

where c is the $(1-\alpha)^{th}$ quantile for a $N(0,1)$ distribution (for instance, $c = -1.64$ for $1-\alpha = 0.05$). See Figure 9.2 for an illustration.

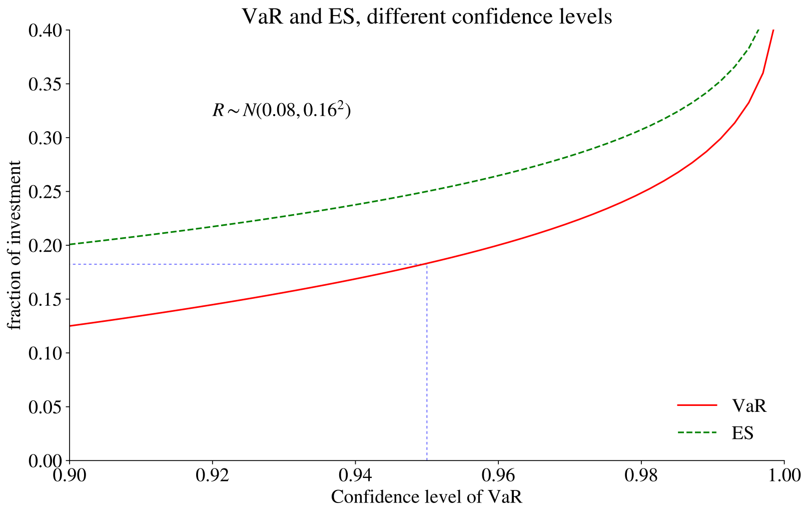

Figure 9.2: Value at risk, different probability levels

Remark 9.4 *(Critical values of $N(\mu, \sigma^2)$) If $R \sim N(\mu, \sigma^2)$, then the probability that $R \leq \mu + c\sigma$ is 5% for $c = -1.64$, 2.5% for $c = -1.96$, and 1% for $c = -2.33$.*

Example 9.5 *(VaR with $R \sim N(\mu, \sigma^2)$) If $R \sim N(\mu, \sigma^2)$ with $\mu = 8\%$ and $\sigma = 16\%$, then $VaR_{97.5\%} = -(0.08 - 1.96 \times 0.16) \approx 0.24$.*

Figure 9.3 shows the distribution and VaRs (for different probability levels) for the daily S&P 500 returns. Two different types of VaRs are shown: *(i)* based on a normal distribution and *(ii)* as the empirical VaR (from the empirical quantiles of the distribution).

Example 9.6 *(VaR and regulation of bank capital) Bank regulations have used 3 times the 99% VaR for 10-day returns as the required bank capital.*

Notice that the return distribution depends on the investment horizon, so a VaR is typically calculated for a stated investment period (for instance, one day). Multi-period VaRs are calculated by either explicitly constructing the distribution of multi-period returns, or

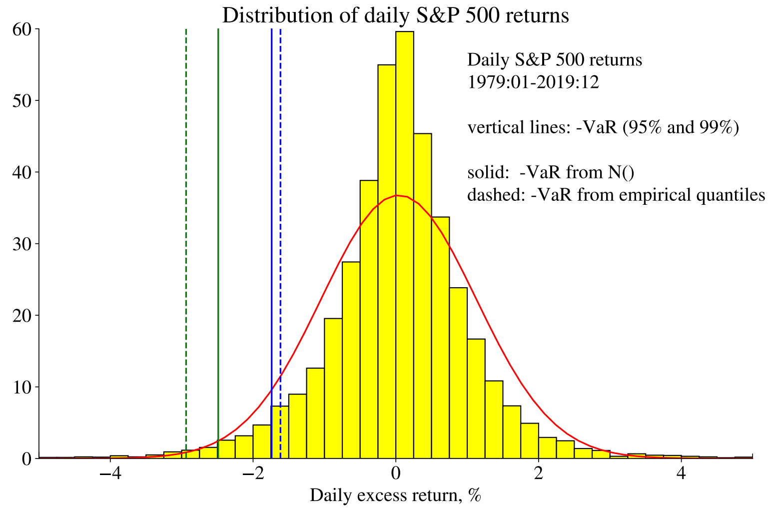

Figure 9.3: Return distribution and VaR for S&P 500

by making simplifying assumptions about the relation between returns in different periods (for instance, that they are iid).

Remark 9.7 *(Multi-period VaR) If the returns are iid, then a q-period return has the mean $q\mu$ and variance $q\sigma^2$, where μ and σ^2 are the mean and variance of the one-period returns respectively. If the mean is zero, then the q-day VaR is $\sqrt{q}$ times the one-day VaR.*

Example 9.8 *(The London whale) The broad outline of the "London whale" (JPM) story is as follows: at the end of 2011, top management instructed the division to bring down the RWA (risk weighted asset) exposure to (various) credit derivatives. However, that would (a) have caused high execution costs and (b) the portfolio had recently performed well, so the division invented a new VaR method and pushed it through the Risk Office without the usual parallel testing. They went on to triple the positions (and lose \$719 million in 2012Q1). Interestingly, the two VaR models show divergent paths for the value at risk.*

Example 9.9 *(Recommendations about risk budgeting). Here are some general recommendations about risk budgeting: (a) define risk limits and stick to them; (b) do not rely on a single risk measure; (c) be extra careful when changing risk measure; (d) do stress tests, eg. using the most extreme events during the last 15 years and/or identify which scenarios would hurt the most; (e) do not forget that risk changes (high volatility of individual assets, higher correlation among them).*

9.2.2 Backtesting a VaR model

While the results in Figure 9.3 are interesting, they are just time-averages in the sense of being calculated from the unconditional distribution: time-variation in the distribution is not accounted for. The problem with that is illustrated in Figure 9.4.

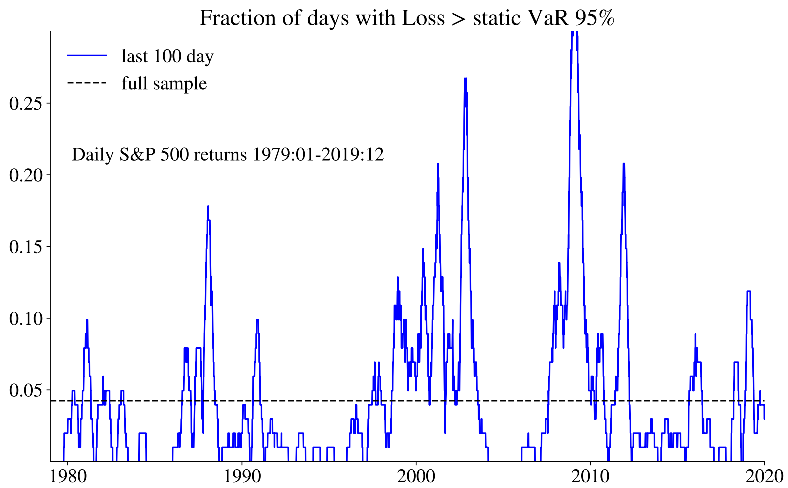

Figure 9.4: Backtesting a static VaR model on a moving data window

Figure 9.5 illustrates the VaR calculated from a time series model for daily S&P returns. In this case, the VaR changes from day to day as both the mean return (the forecast) as well as the standard error (of the forecast error) do. Since *volatility clearly changes over time*, this is crucial for a reliable VaR model. In short, the model is

$$\mu_t = \lambda \mu_{t-1} + (1-\lambda) R_{t-1} \tag{9.9}$$

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda)(R_{t-1} - \mu_{t-1})^2. \tag{9.10}$$

The mean uses a "RiskMetrics" approach of updating yesterday's mean with yesterday's return. This is the same as a weighted average of past returns, but where recent data have higher weights that old data. The variance is a similar updating of yesterday's variance with the square of yesterday's surprise.

Backtesting a VaR model amounts to checking if (historical) data fits with the VaR numbers. For instance, we first find the $\text{VaR}_{95\%}$ and then calculate what fraction of returns that is actually below (the negative of) this number. If the model is correct it should be 5%. We then repeat this for $\text{VaR}_{96\%}$: only 4% of the returns should be below (the negative of) this number.

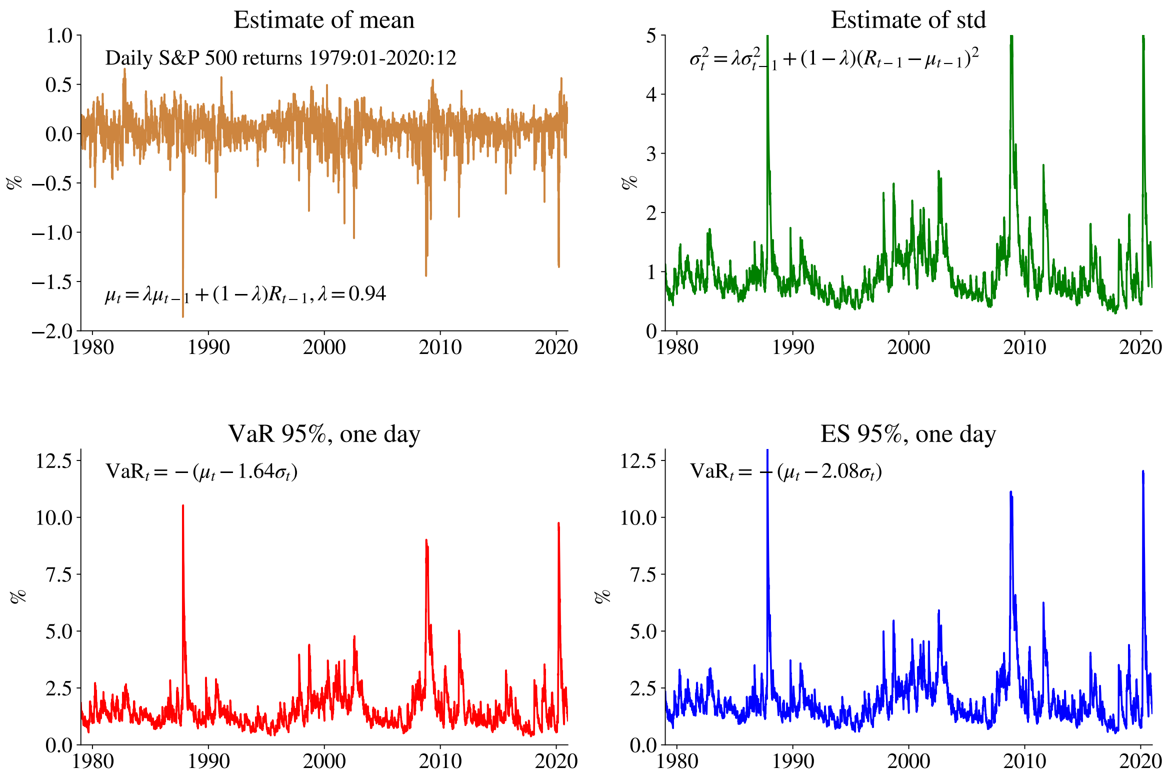

Figure 9.5: A dynamic VaR model

Figures 9.5–9.6 show results from backtesting a VaR model which assumes that one-day returns are normally distributed, but where the volatility is time varying. Clearly, this means that the VaR is also time varying: use (9.5) but allow σ (and less importantly, μ) to change from day to day. The evidence suggests that this model works relatively well at the 95% confidence level and that it is important to account for the time-varying volatility (or else there will be prolonged periods when the VaR performs poorly).

9.2.3 Value at Risk of a Portfolio

The general way of calculating the VaR of a portfolio is the same as for an individual asset (see above): first calculate (or estimate) the parameters of the distribution, then find the quantile.

However, in some special cases, there are ways to directly translate the VaR values of the individual assets to a portfolio VaR.

Remark 9.10 *Suppose the assets in the portfolio are jointly normally distributed with*

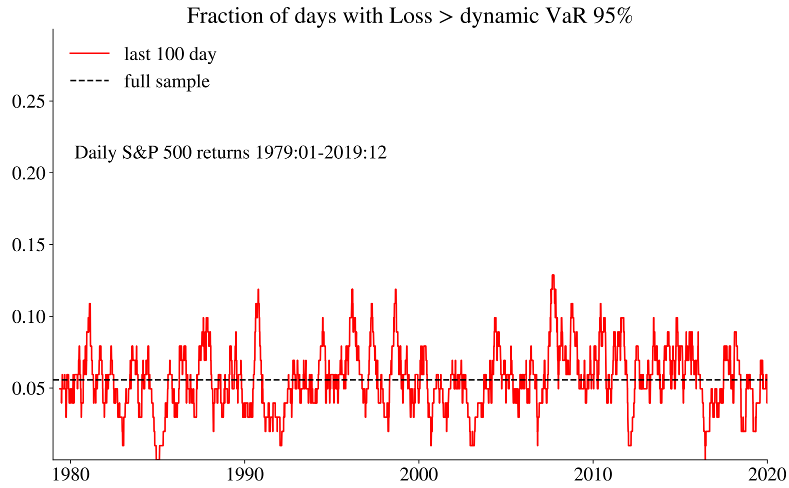

Figure 9.6: Backtesting a dynamic VaR model on a moving data window

zero means, so the VaR of asset i is $VaR_i = 1.64\sigma_i$. *Let* υ *be a vector where* $\upsilon_i = w_i VaR_i$, *where* w_i *is the portfolio weight. Then,* $VaR_p = [\upsilon' \operatorname{Corr}(R)\upsilon]^{1/2}$, *where* $\operatorname{Corr}(R)$ *is the correlation matrix of the assets. (To prove this, recall that* $VaR_p = 1.64\sigma_p$ *and that we can calculate* σ_p *from the* σ_i *values and correlations.)*

9.2.4 Index Models for Calculating the Value at Risk

Consider a multi-index model

$$\begin{aligned} R &= a + b_1 I_1 + b_2 I_2 + \ldots + b_k I_k + e, \text{ or} \\ &= a + b' I + e, \end{aligned} \tag{9.11}$$

where b is a $k \times 1$ vector of the b_i coefficients and I is a $k \times 1$ vector of the I_i indices. As usual, we assume $\mathrm{E}\, e = 0$ and $\operatorname{Cov}(e, I_i) = 0$. This model can be used to generate the inputs to a VaR model. For instance, the mean and standard deviation of the return are

$$\begin{aligned} \mu &= a + b' \mathrm{E}\, I \\ \sigma &= \sqrt{b' \operatorname{Cov}(I) b + \operatorname{Var}(e)}, \end{aligned} \tag{9.12}$$

which can be used in (9.8), that is, an assumption of a normal return distribution. If the return is of a well diversified portfolio and the indices include the key market indices, then the idiosyncratic risk $\operatorname{Var}(e)$ is close to zero. The RiskMetrics approach is to make this assumption.

Stand-alone VaR is a way to assess the contribution of different factors (indices). For instance, the indices in (9.11) could include: an equity indices, interest rates, exchange rates and perhaps also a few commodity indices. Then, an *equity VaR* is calculated by setting all elements in b, except those for the equity indices, to zero. Often, the intercept, a, is also set to zero. Similarly, an *interest rate VaR* is calculated by setting all elements in b, except referring to the interest rates, to zero. And so forth for an *FX VaR* and a *commodity VaR*. Clearly, these different VaRs do not add up to the total VaR, but they still give an indication of where the main risk comes from.

If an asset or a portfolio is a non-linear function of the indices, then (9.11) can be thought of as a first-order Taylor approximation where b_i represents the partial derivative of the asset return with respect to index i. For instance, an option is a non-linear function of the underlying asset value and its volatility (as well as the time to expiration and the interest rate). This approach, when combined with the normal assumption in (9.8), is called the *delta-normal method*.

9.2.5 Expected Shortfall

While the value at risk is a useful risk measure, it has the strange property that it does not make a distinction between a loss that is just below the VaR level and a loss that is a lot below it. The VaR only cares about whether the outcome is in the tail of the return distribution, not how far out.

In addition, the VaR concept has been criticized for having poor aggregation properties. In particular, the VaR for a portfolio is not necessarily (weakly) lower than the portfolio of the VaRs, which contradicts the notion of diversification benefits. (To get this unfortunate property, the return distributions must be heavily skewed.) The *expected shortfall* has better aggregation properties.

The expected shortfall (also called conditional VaR, average value at risk and expected tail loss) has better properties. It is the expected loss when the return actually is below the VaR_α, that is,

$$\text{ES}_\alpha = -\text{E}(R|R \leq -\text{VaR}_\alpha). \tag{9.13}$$

See Figures 9.7 and 9.5 for illustrations.

See Table 9.1 for an empirical comparison of the VaR, ES and some alternative downside risk measures (discussed below).

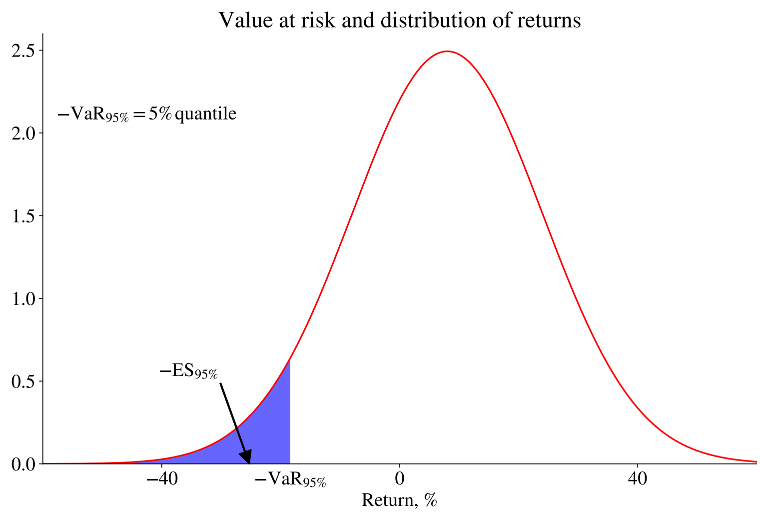

Figure 9.7: Value at risk and expected shortfall

	Small growth	Large value
Std	8.1	5.8
VaR (95%)	12.2	9.2
ES (95%)	18.0	13.2
SemiStd	5.6	3.9
Drawdown	78.6	63.2

Table 9.1: Risk measures of monthly returns of two stock indices (%), US data 1970:01-2020:12.

For a normally distributed return $R \sim N(\mu, \sigma^2)$ we have

$$\mathrm{ES}_\alpha = -\left[\mu - \frac{\phi(c)}{1-\alpha}\sigma\right], \tag{9.14}$$

where $\phi()$ is the pdf of a $N(0,1)$ variable and c is the $1-\alpha$ quantile of a $N(0,1)$ distribution (for instance, -1.64 for $1-\alpha = 0.05$). See Figure 9.2.

Proof. (of (9.14)) If $x \sim N(\mu, \sigma^2)$, then it is well known that $\mathrm{E}(x|x \leq b) = \mu - \sigma\phi(b_0)/\Phi(b_0)$ where $b_0 = (b-\mu)/\sigma$ and where $\phi()$ and $\Phi()$ are the pdf and cdf of a $N(0,1)$ variable respectively. To apply this, use $b = -\mathrm{VaR}_\alpha = \mu + c\sigma$ so $b_0 = c$. Clearly, $\Phi(c) = 1-\alpha$, so $\mathrm{E}(R|R \leq -\mathrm{VaR}_\alpha) = \mu - \sigma\phi(c)/(1-\alpha)$. Multiply by -1. ■

Example 9.11 *(ES) If $\mu = 8\%$ and $\sigma = 16\%$, the 95% expected shortfall is $ES_{95\%} = -(0.08 - 2.08 \times 0.16) \approx 0.25$ (since $\phi(-1.64)/0.05 \approx 2.08$) and the 97.5% expected* *shortfall is* $ES_{97.5\%} = -(0.08 - 2.34 \times 0.16) \approx 0.29$ *(since* $\phi(-1.96)/0.025 \approx 2.34$*)*

Instead, to estimate the expected shortfall from the empirical return distribution, use

$$\mathrm{ES}_{\alpha} = \frac{-1}{\sum_{t=1}^{T} \delta_t} \sum\nolimits_{t=1}^{T} \delta_t R_t, \text{ where } \delta_t = \begin{cases} 1 \text{ if } R_t \leq -\mathrm{VaR}_{\alpha} \\ 0 \text{ otherwise.} \end{cases} \tag{9.15}$$

This expression simply calculates the average $-R_t$ among those observations where $R_t \leq -\mathrm{VaR}_{\alpha}$.

See Figure 9.8 for an illustration.

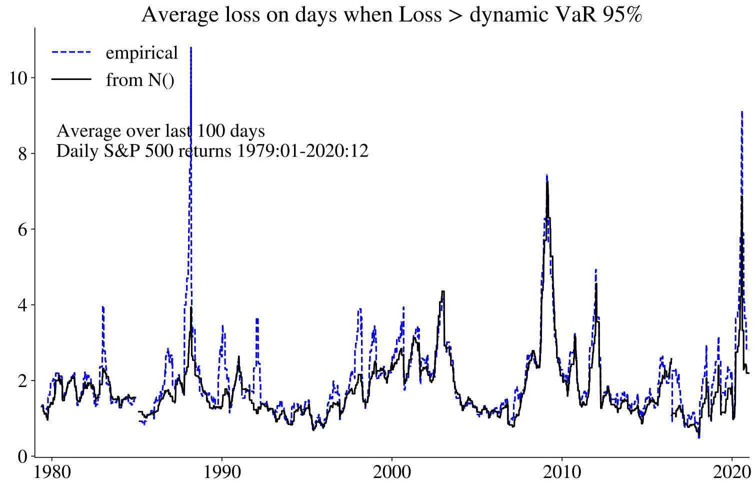

Figure 9.8: Backtesting a dynamic ES model on a moving data window

9.2.6 Target Semivariance (Lower Partial 2nd Moment) and Max Drawdown

Reference: Bawa and Lindenberg (1977) and Nantell and Price (1979)

The target semivariance (also called the lower partial 2nd moment) is defined as

$$\lambda(h) = \mathrm{E}[\min(R - h, 0)^2], \tag{9.16}$$

where h is a "target level" chosen by the investor. Also, $\sqrt{\lambda(h)}$ with $h = \mu$ is called the semi-standard deviation.

In comparison with the variance

$$\sigma^2 = \mathrm{E}(R - \mathrm{E}\, R)^2, \tag{9.17}$$

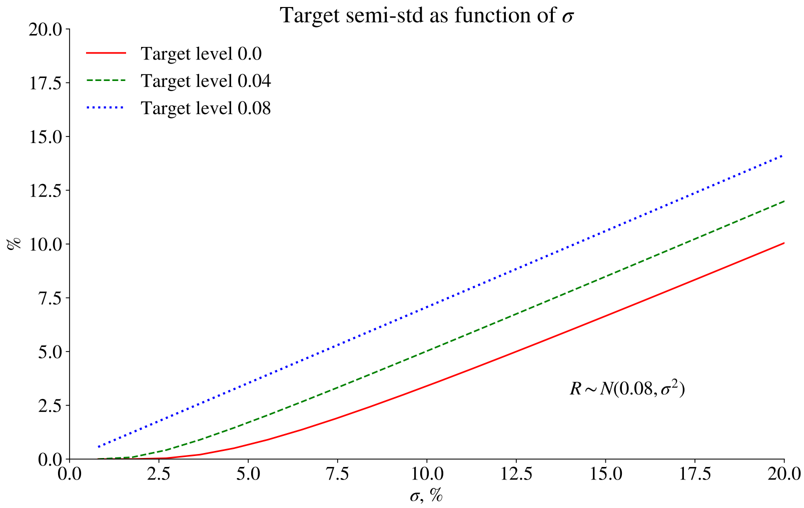

Figure 9.9: Target semivariance as a function of mean and standard deviation for a N(μ,σ^2) variable

the target semivariance differs in two aspects: *(i)* it uses the target level h as a reference point instead of the mean μ: and *(ii)* only negative deviations from the reference point are given any weight.

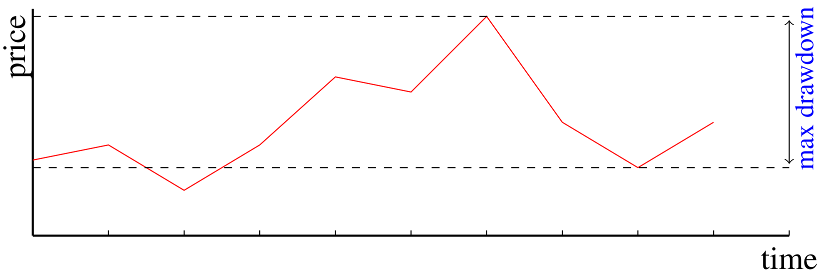

Figure 9.10: Max drawdown

For a normally distributed variable, the target semivariance $\lambda_p(h)$ is increasing in the standard deviation (for a given mean)—see Remark 9.12. See also Figure 9.9 for an illustration.

Instead, to estimate the target semivariance from the empirical return distribution, use

$$\lambda(h) = \frac{1}{T}\sum\nolimits_{t=1}^{T} \delta_t (R_t - h)^2, \text{ where } \delta_t = \begin{cases} 1 \text{ if } R_t \leq h \\ 0 \text{ otherwise.} \end{cases} \tag{9.18}$$

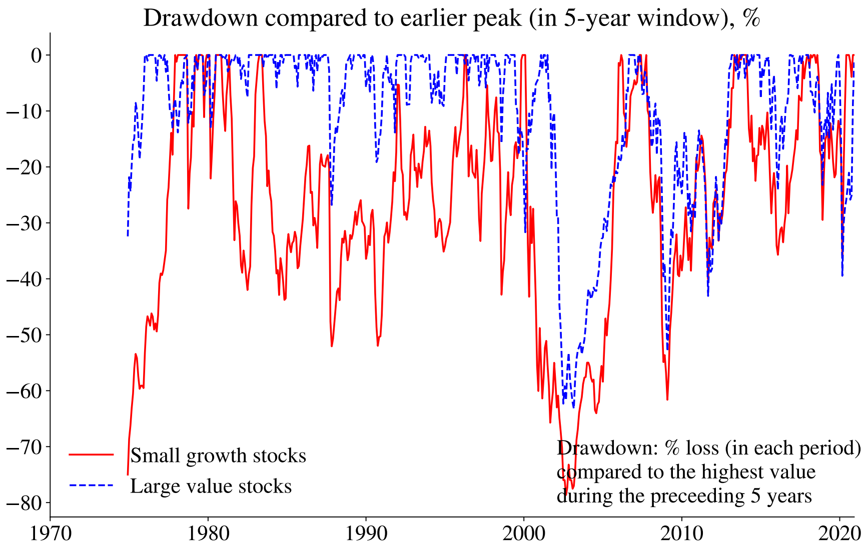

Figure 9.11: Drawdown

This expression simply calculates the average of $\min(R_t - h, 0)^2$. (Warning: some analysts define $\lambda(h)$ by just including those observations when $R_t \leq h$. This means multiplying $\lambda(h)$ in (9.18) by $T/\Sigma_{t=1}^{T}\delta_t$. Conceptually, this is estimating $\mathrm{E}[(R-h)^2|R_t \leq h]$.)

An alternative measure is the (percentage) *maximum drawdown* over a given horizon, for instance, 5 years, say. This is the largest loss from peak to bottom within the given horizon–see Figure 9.10. This is a useful measure when the investor do not know exactly when he/she has to exit the investment—since it indicates the worst (peak to bottom) outcome over the sample. See Figure 9.11 for an illustration.

Remark 9.12 *(Target semivariance calculation for normally distributed variable*) For an $N(\mu, \sigma^2)$ variable, target semivariance around the target level h is*

$$\lambda_p(h) = \sigma^2 a\phi(a) + \sigma^2(a^2+1)\Phi(a), \text{ where } a = (h-\mu)/\sigma,$$

where $\phi()$ and $\Phi()$ are the pdf and cdf of a $N(0,1)$ variable respectively. Notice that $\lambda_p(h) = \sigma^2/2$ for $h = \mu$. See Figure 9.9 for a numerical illustration. It is straightforward (but a bit tedious) to show that

$$\frac{\partial \lambda_p(h)}{\partial \sigma} = 2\sigma\Phi(a),$$

so the target semivariance is a strictly increasing function of the standard deviation.

Remark 9.13 *(Sortino ratio) The Sortino ratio is an alternative to the Sharpe ratio (as a measure of performance). It is* $(\mathrm{E}\, R - h)/\sqrt{\lambda(h)}$.

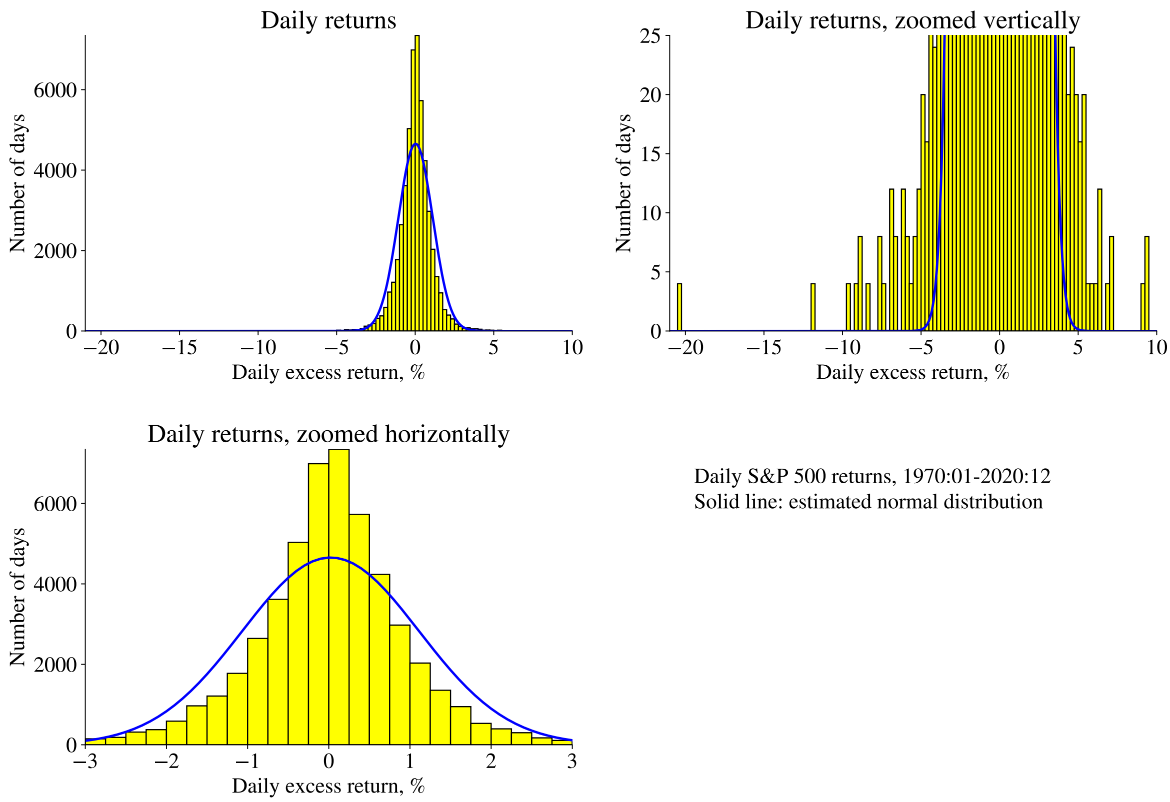

Figure 9.12: Distribution of daily S&P returns

See Table 9.2 for an empirical comparison of the different risk measures.