Scale Models

The simple approximations used to explore a portfolio's behaviour over various time horizons hold traps for the unwary. Francis Diebold, Andrew Hickman, Atsushi Inoue and Til Schuermann explain

What is the relevant horizon for risk management? This obvious question has no obvious answer.[1] Horizons such as seven to 10 days for equity and foreign exchange and 30 days for interest rate instruments are routinely discussed. In fact, even horizons as long as a year are not uncommon.[2]

Operationally, risk is often assessed at a one-day horizon, and even shorter (intra-day) horizons have been discussed. Short-horizon risk measures are commonly converted to other horizons, such as 10-day or 30-day, by scaling.[3] For example, to obtain a 10-day volatility we would multiply the one-day volatility by $\sqrt{10}$. Moreover, the horizon conversion is often significantly longer than 10 days. Many banks, for example, link the measurement of trading volatility to internal capital allocation and risk-adjusted performance measurement schemes, which rely on annual volatility estimates. The temptation is to scale one-day volatility by $\sqrt{252}$.

The routine and uncritical use of scaling is also widely accepted by regulators. For example, it is a prominent feature of the Basle Committee on Banking Supervision's January 1996 *Amendment to the Capital Accord to Incorporate Market Risks*. Specifically, the amendment requires a 10-day holding period and advises conversion by scaling: "In calculating value-at-risk, an instantaneous price shock equivalent to a 10-day movement in prices is to be used, ie, the minimum "holding period" will be 10 trading days. Banks may use value-at-risk numbers calculated according to shorter holding periods scaled up to 10 days by the square root of time..." (page 44, section B4, paragraph c).

In this paper we sound an alarm: such scaling is inappropriate and misleading. We show below that converting volatilities by scaling is statistically appropriate only under strict conditions that are routinely violated by high-frequency (eg, one-day) asset returns. Even in the unlikely event that the conditions for its statistical legitimacy are met, scaling is nevertheless problematic for economic reasons associated with fluctuations in portfolio composition.

Short and long horizons

Our basic findings can be summarised as: scaling works in independent, identically distributed (iid) environments but fails otherwise. Here we describe the restrictive environment in which scaling is appropriate. Let v_t be a log price at time t and suppose that changes in the log price are independently and identically distributed:

$$v_t = v_{t-1} + \varepsilon_t; \quad \varepsilon_t \overset{iid}{\sim} (0, \sigma^2)$$

where the one-day return is ε_t with standard deviation σ. Similarly, the h-day return is:

$$v_t - v_{t-h} = \sum_{i=0}^{h-1} \varepsilon_{t-i}$$

with variance $h\sigma^2$ and standard deviation $\sqrt{h}\sigma$. Hence the "rule": to convert a one-day standard deviation to an h-day standard deviation, simply scale by $\sqrt{h}$. For some applications, a percentile of the distribution of h-day returns may be desired; percentiles also scale by $\sqrt{h}$ if log changes are not only iid but also normally distributed

The prescribed scaling rule relies on one-day returns being independent and identically distributed. The literature on mean reversion in stock returns appreciates this, and scaling is often used as a test for whether returns are iid, ranging from early work (eg, Cootner, 1964) to recent work (eg, Campbell, Lo and MacKinlay, 1997). But high-frequency financial asset returns are distinctly *not* independent and identically distributed. Even if high-frequency portfolio returns are conditional-mean independent (which has been the subject of intense debate in the efficient markets literature), they are certainly not conditional-variance independent, as shown in hundreds of recent papers documenting strong volatility persistence in financial asset returns.[4]

To highlight the failure of scaling in non-iid environments and the nature of the associated erroneous long-horizon volatility estimates, we will use a simple Garch(1,1) process for one-day returns:

$$y_t = \sigma_t \varepsilon_t; \quad \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2; \quad \varepsilon_t \sim \text{NID}(0,1)$$

where $t = 1, \ldots, T$. We impose the usual regularity and covariance stationarity conditions, $0 < \omega < \infty$, $\alpha \geq 0$, $\beta \geq 0$ and $\alpha + \beta < 1$. The key feature of the Garch(1,1) process is that it allows for time-varying conditional volatility, which occurs when α and/or β is non-zero. The model has been fitted to hundreds of financial series and has been tremendously successful empirically; hence its popularity.[5] We hasten to add, however, that our general thesis – that scaling fails in the non-iid environments associated with high-frequency asset returns – does not depend on any way on a Garch(1,1) structure. Rather, we focus on the Garch(1,1) case because it has been studied most intensively, yielding a wealth of results that enable us to illustrate the failure of scaling both analytically and by simulation.

Drost & Nijman (1993) study the temporal aggregation of Garch processes.[6] Suppose we begin with a sample path of a one-day return series, $\{y_{(1)t}\}_{t=1}^{T}$, which follows the Garch(1,1) process above.[7] Then Drost & Nijman show that, under regularity conditions, the corresponding sample path of h-day returns, $\{y_{(h)t}\}_{t=1}^{T/h}$, similarly follows a Garch(1,1) process with:

$$\sigma_{(h)t}^2 = \omega_{(h)} + \beta_{(h)} \sigma_{(h)t-1}^2 + \alpha_{(h)} y_{(h)t-1}^2$$

where:

[1] Chew (1994) provides insightful early discussion

[2] A leading example is Bankers Trust's Raroc system; see Falloon (1995)

[3] See, for example, Smithson & Minton (1996a, b) and JP Morgan (1996)

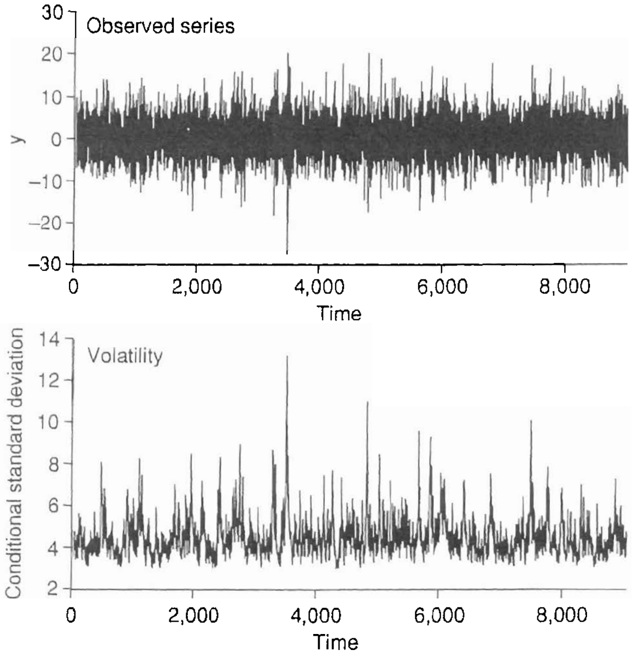
[4] See, for example, the surveys by Bollerslev, Chou & Kroner (1992) and Diebold & Lopez (1995)

[5] Again, see the surveys of Garch models in finance by Bollerslev, Chou & Kroner (1992) and Diebold & Lopez (1995)

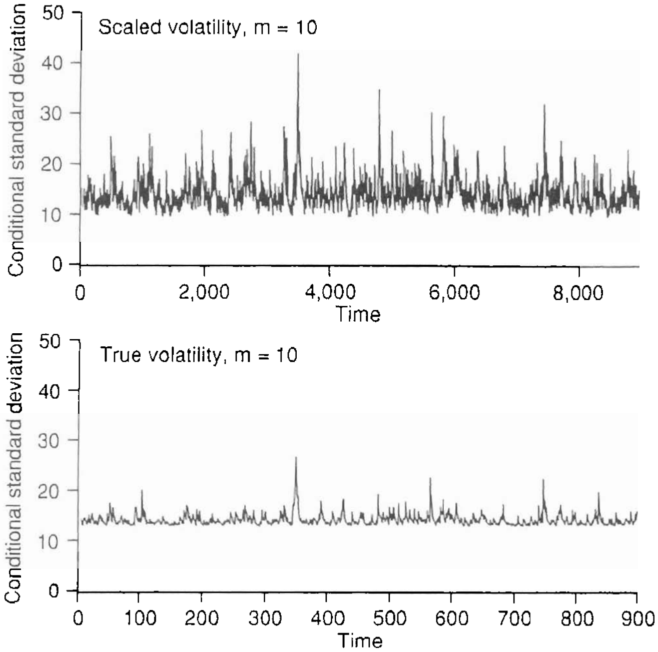
[6] More precisely, they define and study the temporal aggregation of weak Garch processes, a formal definition of which is beyond the scope of this paper. Although the distinction is not crucial for our purposes, technically inclined readers should read "weak Garch" whenever they encounter the word "Garch"

[7] Note the new and more cumbersome, but necessary, notation, the subscript in which keeps track of the aggregation level

1. One-day returns and volatilities



2. True and scaled 10-day volatilities



$$\omega_{(h)} = h\omega \frac{1-(\beta+\alpha)^h}{1-(\beta+\alpha)}$$

$$\alpha_{(h)} = (\beta + \alpha)^h - \beta_{(h)}$$

and $|\beta_{(h)}| < 1$ is the solution of the quadratic equation:

$$\frac{\beta_{(h)}}{1+\beta_{(h)}^2} = \frac{a(\beta+\alpha)^h - b}{a\left(1+(\beta+\alpha)^{2h}\right) - 2b}$$

where:

$$a = h(1-\beta)^2 + 2h(h-1)\frac{(1-\beta-\alpha)^2(1-\beta^2-2\beta\alpha)}{(\kappa-1)(1-(\beta+\alpha)^2)} + 4\frac{\left(h-1-h(\beta+\alpha)+(\beta+\alpha)^h\right)(\alpha-\beta\alpha(\beta+\alpha))}{1-(\beta+\alpha)^2}$$

$$b = (\alpha - \beta\alpha(\beta+\alpha))\frac{1-(\beta+\alpha)^{2h}}{1-(\beta+\alpha)^2}$$

and κ is the kurtosis of y_t.[8] The Drost-Nijman formula is neither pretty nor intuitive, but it is important, because it is the key to correct conversion of one-day volatility to h-day volatility. It is painfully obvious, moreover, that the simple scaling formula does not look at all like the Drost-Nijman formula.

If, however, the scaling formula were an accurate approximation to the Drost-Nijman formula, it would still be very useful because of its simplicity and intuitive appeal. Unfortunately, this is not the case. As $h \to \infty$, analysis of the Drost-Nijman formula reveals that $\alpha_{(h)} \to 0$ and $\beta_{(h)} \to 0$, which is to say that temporal aggregation produces gradual disappearance of volatility fluctuations.[9] Scaling, in contrast, magnifies volatility fluctuations.

A detailed example

Let us examine the failure of scaling in a specific example. We parameterise the Garch(1,1) process to be realistic for daily returns by setting a = 0.10 and b = 0.85, which are typical of the parameter values obtained for estimated Garch(1,1) processes. The choice of w is arbitrary and amounts to a normalisation, or choice of scale. We set w = 1, which implies that the unconditional variance of the process equals 20. We also set:

$$\sigma_0^2 = \frac{\omega}{1-\alpha-\beta}$$

discard the first 1,000 realisations to allow the effects of the initial condition to dissipate, and keep the following T = 9,000 realisations. In figure 1, we show the series of daily returns and the corresponding series of one-day conditional standard deviations, σ_t. The daily volatility fluctuations are evident.

Now we examine the 10-day volatility, corresponding to h = 10. In figure 2, we show 10-day volatilities calculated in two different ways. We obtain the first (incorrect) 10-day volatility by scaling the one-day volatility, σ_t, by $\sqrt{10}$. We obtain the second (correct) 10-day volatility by applying the Drost-Nijman formula.[10]

It is clear that although scaling produces volatilities that are correct on average, it magnifies the volatility fluctuations, whereas they should in fact be lessened. That is, scaling produces erroneous conclusions of large fluctuations in the conditional variance of long-horizon returns, when in fact the opposite is true. Moreover, we cannot claim that the scaled volatility estimates are "conservative" in any sense; rather, they are sometimes too high and sometimes too low.

If scaling is inappropriate, then what is appropriate? Firstly, as we have shown, if the short-horizon return model is correctly specified as a Garch(1,1) process, then long-horizon volatilities can be calculated using the Drost-Nijman formula. Second, if the short-horizon return model is correctly specified but does not fall into the family of models covered by Drost and Nijman, then the Drost-Nijman results do not apply, and there are no known analytic methods for computing h-day volatilities from one-day volatilities. If we had analytic formulas, we could apply them, but we don't. So, if h-day volatilities are of interest, it makes sense to use an h-day model.

Third, when the one-day return model is not correctly specified, things

[8] Bollerslev (1986) shows that a necessary and sufficient condition for the existence of a finite fourth moment, and hence a finite kurtosis, is $3\alpha^2+2\alpha\beta+\beta^2 < 1$

[9] The Drost-Nijman result coheres with the result of Diebold (1988), who shows that temporal aggregation produces returns that approach an unconditional normal distribution, which implies that volatility fluctuations must vanish

[10] We set $\sigma^2_{(10)1}$ at its unconditional mean

are even trickier.[11] For example, the best approximation to 10-day return volatility dynamics may be very different from what one gets by applying the Drost-Nijman formula to an (incorrect) estimated Garch(1,1) model for one-day return volatility dynamics (and of course very different as well from what one gets by scaling estimates of daily return volatilities by $\sqrt{10}$). This again suggests that if h-day volatilities are of interest, it makes sense to use an h-day model.

Conclusion

[redacted]he relevant horizon may vary by asset class (eg, equity versus fixed income), industry (eg, banking versus insurance), position in the firm (eg, trading desk versus chief financial officer), and motivation (eg, private versus regulatory), and thought must be given to the relevant horizon on an application-by-application basis. Modelling volatility only at one short horizon, followed by scaling to convert to longer horizons, is likely to be inappropriate and misleading, because temporal aggregation should reduce volatility fluctuations, whereas scaling amplifies them.[12] Instead, a strong case can be made for using different models at different horizons.[13]

We hasten to add that it is not our intent to condemn scaling always and everywhere. Scaling is charmingly simple, and it is appropriate under certain conditions. Moreover, even when those conditions are violated, scaling produces results that are correct on average, as we showed. Hence scaling has its place, and its widespread use as a tool for approximate horizon conversion is understandable. But as our sophistication increases, the flaws with such "first-generation" rules of thumb become more pronounced, and the possibilities for improvement become apparent. Our intent is to stimulate such improvement.

We believe that the use of different models for different horizons is an important step in the right direction. But even with that sophisticated strategy, the nagging and routinely neglected problem of portfolio fluctuations, pinpointed in a prescient article by Kupiec & O'Brien (1995), remains. Measuring the volatility of trading results depends not only on the volatility of the relevant market prices but also on the position vector that describes the portfolio. Estimates of h-day volatility are predicated on the assumption of a fixed position vector throughout the h-day horizon, which is unlikely.

Positions tend to change frequently in the course of normal trading, both within and across days, for a number of reasons. First, positions may be taken to facilitate a customer transaction, and then decline to normal "inventory" levels when offsetting customer orders come in, or when the positions are laid off in the market or hedged. Second, traders may put on (or take off) short-term speculative positions, or adjust long-term proprietary trading strategies. Finally, trading management may intervene to reduce positions in response to adverse market movements.

Whatever the cause of fluctuations in the position vector, it conflicts with the h-day buy-and-hold assumption. The degree to which this assumption is violated will depend on the trading desk's business strategy, the instruments it trades and the liquidity of the markets in which it trades. For example, even one day may be too long a horizon over which to assume a constant portfolio for a market maker in a major European currency – the end-of-day portfolio will bear little relation to the variety of positions that could be taken over the course of the next day, much less the next 10 days. To understand the risk over a longer horizon, we need not only robust statistical models for the underlying market price volatility but also robust behavioural models for changes in trading positions.

Finally, we stress the challenges associated with aggregating risks across positions and trading desks when the risks are assessed at different horizons. Obviously, one cannot simply add together risk measures at different horizons. Instead, conversion to a common horizon must be done through a combination of statistically appropriate h-day models of price volatility and behavioural models for changes in traders' positions. That, in our view, is a pressing direction for future research. ■

Francis Diebold is professor of economics and statistics at the University of Pennsylvania and a member of the Oliver, Wyman Institute. Andrew Hickman, formerly of Oliver, Wyman & Company, is an expert in quantitative risk management. Atsushi Inoue is a PhD candidate in economics at the University of Pennsylvania. Til Schuermann is head of research at Oliver, Wyman & Company.
The first author would like to thank the Oliver, Wyman Institute for its hospitality and inspiration. Peter Christoffersen, Paul Kupiec, Jose Lopez and Tony Santomero provided helpful discussion, as did participants at the Federal Reserve Bank of Atlanta Conference on Financial Market Structure in a Global Environment, but all errors are ours alone

[11] *A moment's reflection reveals misspecification to be the compelling case. The modern approach is to acknowledge misspecification from the outset, as for example in the influential paper of Nelson & Foster (1994)*

[12] *Moreover, Christoffersen & Diebold (1997) show that the predictable volatility dynamics in many asset returns vanish quickly with horizon, indicating that scaling can quickly lead one astray*

[13] *See Findley (1983) and Diebold (1998) for discussion of this same point in the context of more traditional forecasting problems*

REFERENCES

Bollerslev T, 1986
Generalized autoregressive conditional heteroskedasticity
Journal of Econometrics 31, pages 307–327

Bollerslev T, R Chou and K Kroner, 1992
Arch modeling in finance: a selective review of the theory and empirical evidence
Journal of Econometrics 52, pages 5–59

Campbell J, A Lo and A MacKinlay, 1997
The economics of financial markets
Princeton: Princeton University Press

Chew L, 1994
Shock treatment
Risk September, pages 63–70

Christoffersen P and F Diebold, 1997
How relevant is volatility forecasting for financial risk management?
Wharton Financial Institutions Center working paper 97-45, University of Pennsylvania

Cootner P, 1964
The random character of stock market prices
MIT Press, Cambridge

Diebold F, 1988
Empirical modeling of exchange rate dynamics
Springer-Verlag, New York

Diebold F, 1998
Elements of forecasting in business, economics, government and finance
South-Western College Publishing, Cincinnati, Ohio

Diebold F and J Lopez, 1995
Modeling volatility dynamics
in Macroeconometrics: developments, tensions and prospects, Kevin Hoover, editor, Kluwer Academic Press, Boston, pages 427–472

Drost F and T Nijman, 1993
Temporal aggregation of Garch processes
Econometrica 61, pages 909–927

Falloon W, 1995
2020 Visions
Risk October, pages 43–45

Findley D, 1983
On the use of multiple models for multi-period forecasting
American Statistical Association, Proceedings of the Business and Economic Statistic Section, pages 528–531

JP Morgan, 1996
RiskMetrics – technical document
Fourth edition, New York

Kupiec P and J O'Brien, 1995
Internal affairs
Risk May, pages 43–47

Nelson D and D Foster, 1994
Asymptotic filtering theory for univariate Arch models
Econometrica 62, pages 1–41

Smithson C and L Minton, 1996a
Value-at-risk
Risk January, pages 25–27

Smithson C and L Minton, 1996b
Value-at-risk (2)
Risk February, pages 38–39