Diebold, F.X. Hickman, A., Inoue, A. and Schuermann, T. (1998),
"Converting 1-Day Volatility to h-Day Volatility: Scaling by Root-h is Worse than You Think,"
Wharton Financial Institutions Center, Working Paper 97-34.
Published in condensed form as "Scale Models," Risk, 11, 104-107, 1998.

Converting 1-Day Volatility to h-Day Volatility: Scaling by $\sqrt{h}$ is Worse than You Think

Francis X. Diebold
University of Pennsylvania
fdiebold@mail.sas.upenn.edu

Andrew Hickman
Oliver, Wyman and Company
ahickman@owc.com

Atsushi Inoue
University of Pennsylvania
inoue@econ.sas.upenn.edu

Til Schuermann
Oliver, Wyman and Company
tschuermann@owc.com

December 1996
This Print: July 3, 1997

Correspondence to:
F.X. Diebold
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297

Abstract: We show that the common practice of converting 1-day volatility estimates to h-day estimates by scaling by $\sqrt{h}$ is inappropriate and produces overestimates of the variability of long-horizon volatility. We conclude that volatility models are best tailored to tasks: if interest centers on long-horizon volatility, then a long-horizon volatility model should be used. Economic considerations, however, confound even that prescription and point to important directions for future research.

Acknowledgments: 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 System Conference on Financial Market Structure in a Global Environment, but all errors are ours alone. We gratefully acknowledge support from the National Science Foundation, the Sloan Foundation, and the University of Pennsylvania Research Foundation.

1. Motivation and Background

What is the relevant horizon for risk management? This obvious question has no obvious answer.[1] Horizons such as 7 to 10 days for equity and foreign exchange, and longer horizons such as 30 days for interest rate instruments, are routinely discussed. In fact, horizons as long as a year are not uncommon.[2] Operationally, risk is often assessed at a 1-day horizon, and shorter (intra day) horizons have even been discussed. Short-horizon risk measures are converted to other horizons, such as 10-day or 30-day, by scaling.[3] For example, to obtain a 10-day volatility we multiply the 1-day volatility by $\sqrt{10}$. Moreover, the horizon conversion is often significantly longer than 10 days. Many banks, for example, link trading volatility measurement to internal capital allocation and risk-adjusted performance measurement schemes, which rely on annual volatility estimates. The temptation is to scale 1-day volatility by $\sqrt{252}$.

The routine and uncritical use of scaling is also widely accepted by regulators. For example, the Basle Committee's January 1996 "Amendment to the Capital Accord to Incorporate Market Risks" features it prominently. 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, i.e. the minimum "holding period" will be ten trading days. Banks may use value-at-risk numbers calculated according to shorter holding periods scaled up to ten days by the square root of time ...
> (p. 44, section B.4, paragraph c)

---

[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 and Minton (1996a, b) and J.P. Morgan (1996).

In this paper we sound an alarm: such scaling is inappropriate and misleading. We show in section 2 that converting volatilities by scaling is statistically appropriate only under strict conditions that are routinely violated by high-frequency (e.g., 1-day) asset returns. In section 3, we provide a detailed illustrative example of the failure of scaling. We conclude in section 4, in which we note that, 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 .

2. The Links Between Short-Horizon and Long-Horizon Risk: Statistical Considerations

Scaling Works in 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,

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

Then the 1-day return is

$$v_t - v_{t-1} = \varepsilon_t,$$

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 "$\sqrt{h}$ rule": to convert a 1-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 iid, and in addition, normally distributed.

The scaling rule relies on 1-day returns being iid. 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 (e.g., Cootner, 1964) to recent work (e.g., Campbell, Lo and MacKinlay, 1996). But high-frequency financial asset returns are distinctly *not* iid. 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 evidenced by hundreds of recent papers documenting strong volatility persistence in financial asset returns.[4]

The Failure of Scaling in non-iid Environments

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 1-day returns,

$$y_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \omega + \alpha y_t^2 + \beta \sigma_{t-1}^2$$

$$\varepsilon_t \sim NID(0,1),$$

t = 1, ..., 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 nonzero. The model

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

has been fit 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 the most intensely, yielding a wealth of results that enable us to illustrate the failure of scaling both analytically and by simulation.

Drost and Nijman (1993) study the temporal aggregation of GARCH processes.[6] Suppose we begin with a sample path of a 1-day return series, $\{y_{(1)t}\}_{t=1}^{T}$, which follows the GARCH(1,1) process above.[7] Then Drost and 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

$$\sigma^2_{(h)t} = \omega_{(h)} + \beta_{(h)}\sigma^2_{(h)t-1} + \alpha_{(h)}y^2_{(h)t-1}$$

where

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

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

[5] Again, see the surveys of GARCH models in finance by Bollerslev, Chou and Kroner (1992) and Diebold and 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.

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

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

where

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

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

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 1-day volatility to h-day volatility. It is painfully obvious, moreover, that the 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, such 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.

[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,

3. A Detailed Example

Let us examine the failure of scaling in a specific example. We parameterize the GARCH(1,1) process to be realistic for daily returns by setting α=0.10 and β=0.85, which are typical of the parameter values obtained for estimated GARCH(1,1) processes. Choice of ω is arbitrary and amounts to a normalization, or choice of scale. We set ω=1, which implies that the unconditional variance of the process equals 20. We set $\sigma_0^2 = \frac{\omega}{1-\alpha-\beta}$, discard the first 1,000 realizations to allow the effects of the initial condition to dissipate, and keep the following T=9,000 realizations. In Figure 1 we show the series of daily returns and the corresponding series of 1-day conditional standard deviations, σ_t. The daily volatility fluctuations are evident.

Now we examine 10-day and 90-volatilities, corresponding to h=10 and h=90. In Figure 2 we show 10-day volatilities computed in two different ways. We obtain the first (incorrect) 10-day volatility by scaling the 1-day volatility, σ_t, by $\sqrt{10}$. We obtain the second (correct) 10-day volatility by applying the Drost-Nijman formula.[10] In Figure 3, we repeat the comparison of Figure 2, except we display 90-day rather than 10-day volatilities.

It is clear that although scaling produces volatilities that are correct on average, it magnifies the volatility fluctuations, whereas they should in fact be damped. 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

---

which implies that volatility fluctuations must vanish.

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

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? First, 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 computed 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 1-day volatilities. If we had analytic formulae, we could apply them, but we don't. Hence if h-day volatilities are of interest, it makes sense to use an h-day model.

Third, when the 1-day return model is *not* correctly specified, things 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 1-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.

4. Concluding Remarks

The relevant horizon may vary by asset class (e.g., equity vs. fixed income), industry (e.g., banking vs. insurance), position in the firm (e.g., trading desk vs. CFO), and motivation

[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 and Foster (1994).

(e.g., private vs. regulatory), and thought must be given to the relevant horizon on an application-by-application basis. Modeling 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 directions 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 and 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

---

[12] Moreover, Christoffersen and 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.

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 in order 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 behavioral 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 can't 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 behavioral models for changes in traders' positions. That, in our view, is a pressing direction for future research.

References

Bollerslev, T. (1986), “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics*, 31, 307-327.

Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992), "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5-59.

Campbell, J.Y., Lo., A.W. and MacKinlay, A.C. (1997), *The Economics of Financial Markets*. Princeton: Princeton University Press.

Chew, L. (1994), “Shock Treatment,” *Risk*, 7, September.

Christoffersen, P.F. and Diebold, F.X. (1997), “Are Volatility Dynamics Relevant for Risk Management?,” Manuscript, Research Department, International Monetary Fund, and Department of Economics, University of Pennsylvania.

Cootner, P. (1964), *The Random Character of Stock Market Prices.* Cambridge: M.I.T. Press.

Diebold, F.X. (1988), *Empirical Modeling of Exchange Rate Dynamics*. New York: Springer-Verlag.

Diebold, F.X. (1998), *Elements of Forecasting in Business, Economics, Government and Finance*. Cincinnati, Ohio: South-Western College Publishing.

Diebold, F.X. and Lopez, J. (1995), "Modeling Volatility Dynamics," in Kevin Hoover (ed.), *Macroeconometrics: Developments, Tensions and Prospects*. Boston: Kluwer Academic Press, 427-472.

Drost, F.C. and Nijman, T.E. (1993), “Temporal Aggregation of GARCH Processes,” *Econometrica*, 61, 909-927.

Falloon, W. (1995), “2020 Visions,” *Risk*, 8, October.

Findley, D.F. (1983), “On the Use of Multiple Models for Multi-Period Forecasting,” *American Statistical Association, Proceedings of the Business and Economic Statistic Section,* 528-531.

J.P. Morgan (1996) “RiskMetrics --Technical Document,” Fourth Edition, New York.

Kupiec, P. and O’Brien, J. (1995), “Internal Affairs,” *Risk*, 8, May.

Nelson, D.B. and Foster, D.P. (1994), "Asymptotic Filtering Theory for Univariate ARCH Models," *Econometrica*, 62, 1-41.

Smithson, C. and Minton, L. (1996a), “Value at Risk,” *Risk*, 9, January.

Smithson, C. and Minton, L. (1996b), “Value at Risk (2),” *Risk*, 9, February.

Figure 1
GARCH(1,1) Realization and Conditional Standard Deviation

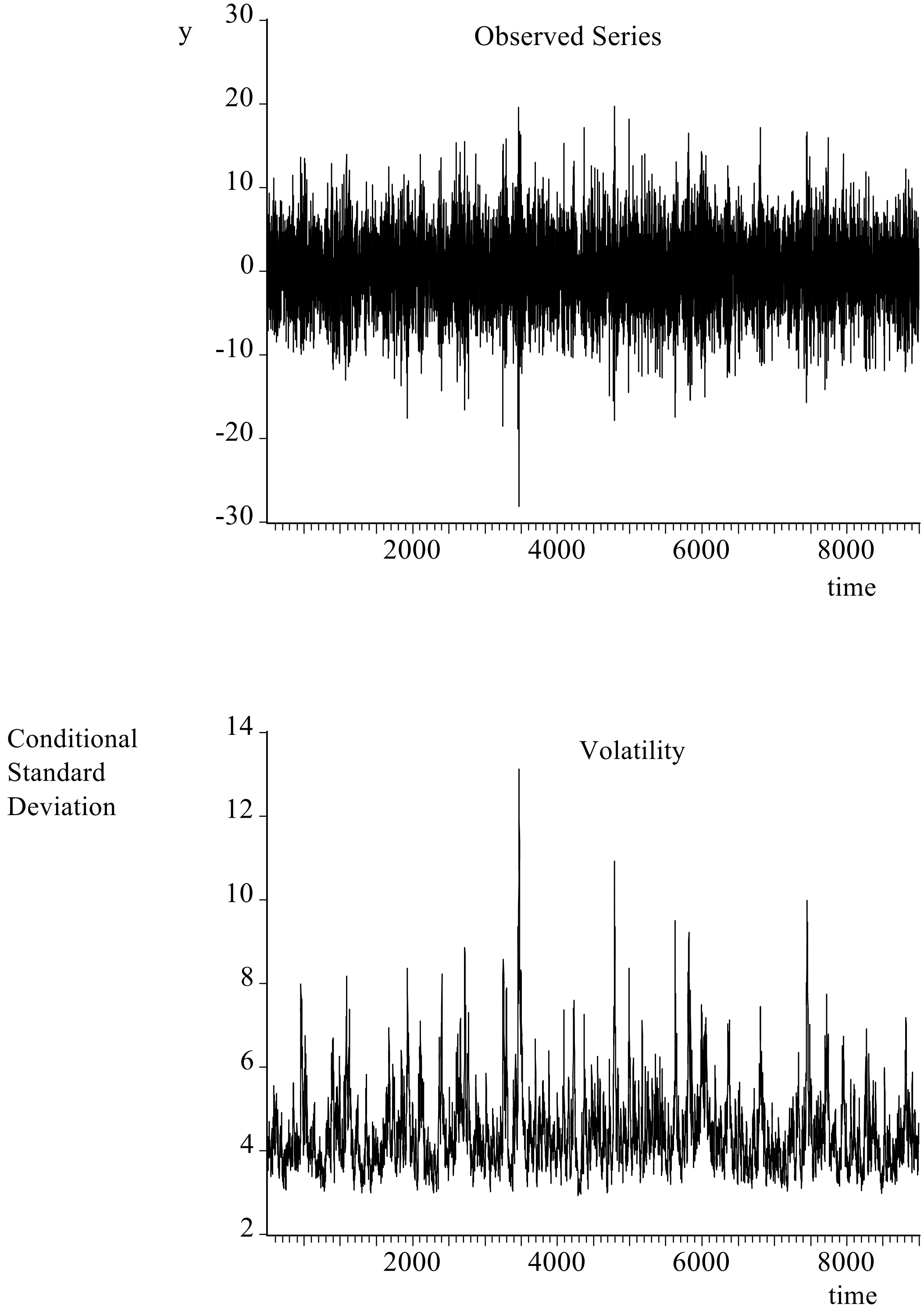
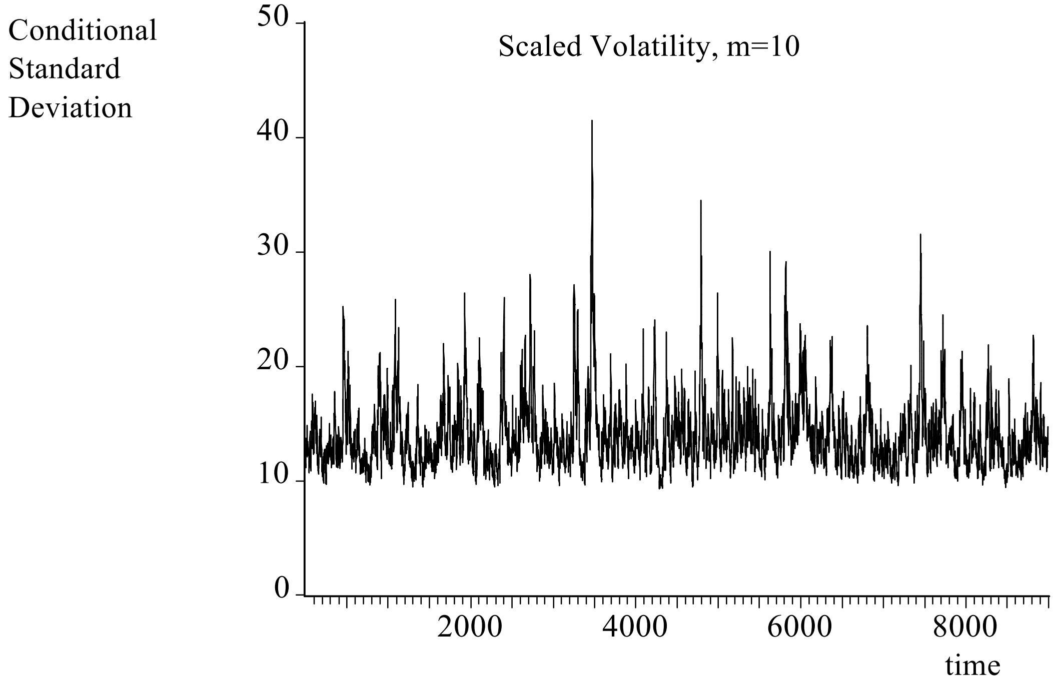


Figure 2
10-Day Volatility, Scaled and Actual



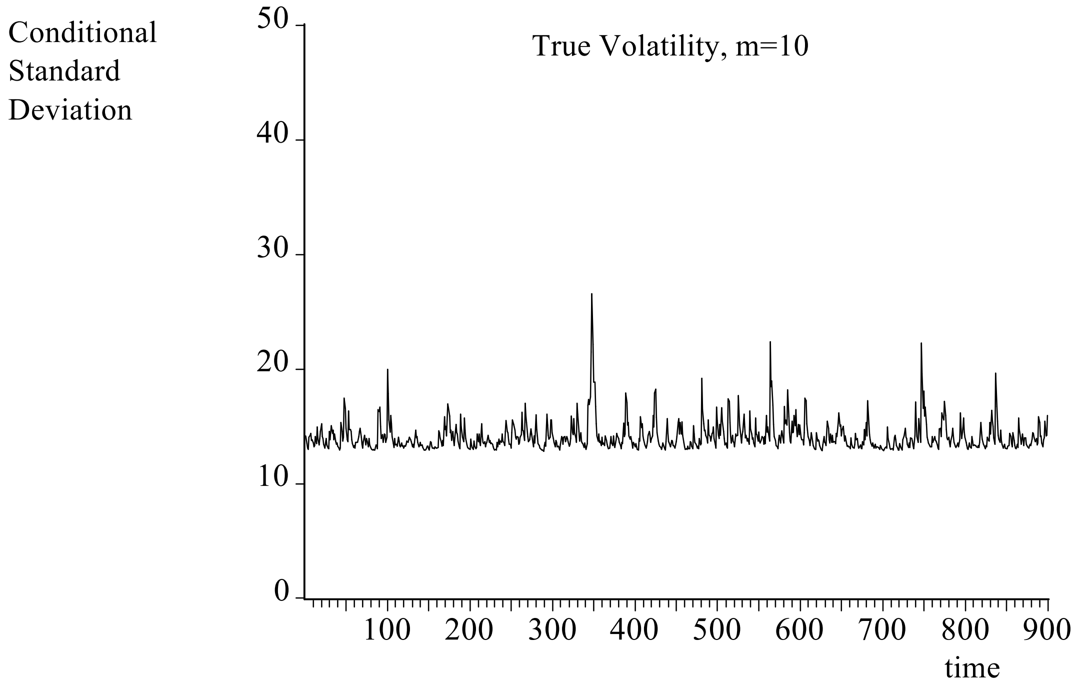


Figure 3
90-Day Volatility, Scaled and Actual