The Impact of Volatility Targeting

Campbell R. Harvey, Edward Hoyle, Russell Korgaonkar, Sandy Rattray, Matthew Sargaison and Otto Van Hemert

*JPM* 2018, 45 (1) 14-33
doi: https://doi.org/10.3905/jpm.2018.45.1.014
http://jpm.iijournals.com/content/45/1/14

This information is current as of November 1, 2018.

**Supplementary Material**  http://jpm.iijournals.com/content/suppl/2018/11/01/45.1.14.DC1

**Email Alerts**  Receive free email-alerts when new articles cite this article. Sign up at:  http://jpm.iijournals.com/alerts
The Impact of Volatility Targeting

CAMPBELL R. HARVEY, EDWARD HOYLE, RUSSELL KORGAONKAR, SANDY RATTRAY, MATTHEW SARGAISON, AND OTTO VAN HEMERT

One of the key features of volatility is that it is persistent, or clusters. High volatility over the recent past tends to be followed by high volatility in the near future. This observation underpins Engle’s [1982] pioneering work on autoregressive conditional heteroskedasticity (ARCH) models. In this article, we study the risk and return characteristics of assets and portfolios that are designed to counter the fluctuations in volatility. We achieve this by leveraging the portfolio at times of low volatility and scaling down at times of high volatility. Effectively, the portfolio is targeting a constant level of volatility, rather than a constant level of notional exposure.

Conditioning portfolio choice on volatility has attracted considerable recent attention. The financial media have zoomed in on the increasing popularity of risk parity funds. In recent work, Moreira and Muir (2017) found that volatility-managed portfolios increase the Sharpe ratios in the case of the broad equity market and a number of dynamic, mostly long–short stock strategies.

Although most research has concentrated on equity markets, we investigate the impact of volatility targeting across more than 60 assets, with daily data beginning as early as 1926. We find that Sharpe ratios are higher with volatility scaling for risk assets (equities and credit), as well as for portfolios that have a substantial allocation to these risk assets, such as a balanced (60–40 equity–bond) portfolio and a risk parity (equity–bond–credit–commodity) portfolio.

Risk assets exhibit a so-called leverage effect (i.e., a negative relation between returns and volatility), and so volatility scaling effectively introduces some momentum into strategies. That is, volatility often increases in periods of negative returns, causing positions to be reduced, which is in the same direction as what one would expect from a time-series momentum strategy. Historically such a momentum strategy has performed well (see, e.g., Hamill, Rattray, and Van Hemert 2016). For other assets, such as bonds, currencies, and commodities, volatility scaling has a negligible effect on realized Sharpe ratios.

We show that volatility targeting consistently reduces the likelihood of extreme returns (and the volatility of volatility) across our 60+ assets. Under reasonable investor preferences, a thinner left tail is much
preferred (for a given Sharpe ratio).\(^3\) Volatility targeting also reduces the maximum drawdowns for both the balanced and risk parity portfolios.

**PRELIMINARIES**

**Data**

Our study relies on daily return data as a starting point. Often, monthly data are available for longer histories, but these data are less suitable for obtaining responsive volatility estimates. Exhibit 1 provides an overview.

In the next section, we will consider U.S. equity data. The earliest daily return dataset available to us is from July 1, 1926, and is obtained from Kenneth French’s website.\(^4\) It is the value-weighted returns of firms listed on the NYSE, AMEX, and NASDAQ, henceforth referred to as *Equities All U.S.* We will also use the returns of the 10 industry portfolios, available from the same source and start date. We additionally use S&P 500 futures data from 1988, which allows us to estimate volatility based on intraday data.

In the following section, we focus our attention on fixed income and other assets. For U.S. Treasury bonds, daily yields are available since 1962 from the Federal Reserve.\(^5\) We construct proxy daily returns by assuming that 10-year yields are par yields and computing the return of par coupon bonds.\(^6\) We also use 10-year on Saturday as well, and thus the data include the Saturday returns up to then.

\(^3\) Under the common assumption of concave utility, investors dislike the left tail more than they like the right tail. Hence, for a given Sharpe ratio, investors are willing to give up some of the right tail to reduce the left tail.

\(^4\) See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Until 1952, stocks on the NYSE traded

\(^5\) Federal Reserve Economic Data (see https://fred.stlouisfed.org). To illustrate the return dynamics over a longer period of time, we will use monthly data in Exhibit 10, obtained from Global Financial Data from July 1926 (to match the start date of the equity data).

\(^6\) Assuming that the 10-year yield is the par yield on a semi-annual coupon paying bond, we reprice the bond the following day using that day’s 10-year yield and assuming that all cash flows are now 1/261 years closer (with 261 the assumed number of weekdays per year). The return over the one-day period is the new price minus one (par).
Treasury futures data from 1988, which will again allow us to evaluate volatility estimates based on intraday data. We also explore credit returns, hedged with Treasuries, creating a long time series for an exposure that should resemble the synthetic CDX investment-grade index that is available today. To this end, we use the Bank of America Merrill Lynch U.S. Corporate Master Total Return index, and the hedging methodology follows Cook et al. (2017).

We also use daily futures and forwards data for 50 liquid assets from Cook et al. (2017). This dataset covers commodities (six agricultural, six energy, and seven metal contracts), nine currencies (all against the U.S. dollar), 10 equity indexes, nine bonds, and three interest rate contracts.

**Volatility Scaling**

We focus on excess returns because they capture the compensation for bearing risk, not the time value of money. Excess returns are a type of unfunded returns (e.g., a long equities position financed by borrowing at the risk-free T-bill rate). The unfunded nature of excess returns makes evaluating scaled position returns particularly straightforward: Volatility-scaled returns are simply inversely proportional to a conditional volatility estimate that is known a full 24 hours ahead of time, using returns up to \( t - 2 \).

That is,

\[
r_{t}^{\text{scaled}} = r_{t} \times \frac{\sigma_{\text{target}}}{\sigma_{t-2}} \times k^{\text{scaled}}
\]

where we added a constant \( k \) (approximately 1), chosen such that ex post, over the full sample period, the target volatility is realized. We do this to facilitate comparison across different securities, methods, and sample periods. We will set the volatility target to 10% annualized throughout.

Unscaled returns involve no conditional volatility estimate, just a constant to achieve the same ex post 10% realized volatility as scaled returns:

\[
r_{t}^{\text{unscaled}} = r_{t} \times k^{\text{unscaled}}
\]

Notice that futures and forwards trade on margin, so their returns are already essentially unfunded; thus, the risk-free rate is not deducted. In addition, the Treasury-hedged credit returns are unfunded by construction.

To estimate volatility, we use the standard deviation of daily returns, with exponentially decaying weights to returns at different lags. We find similar results when using equal weights to returns over a rolling window of fixed length or using estimates based on three- or five-day overlapping returns (not reported).

Because volatility may be more precisely estimated with higher-frequency data, we also examine the effects of scaling by intraday volatility for the S&P 500 and 10-year Treasury futures since 1988. We obtain a volatility estimate from five-minute returns over the liquid 9:15 a.m. to 2:00 p.m. (Chicago time) time window.

We aggregate squared returns to a daily realized variance value, average these daily values with exponentially decaying weights, and then take the square root.

For the Equities All U.S. data, we work with calendar-day data, to account for Saturday returns before 1952. For other assets, we work with weekday data. In all cases, we annualize the volatility estimate, adjusting for the number of data points, to ensure comparability.

**Performance Statistics**

In Exhibit 2, we list the performance statistics on which we focus. In most cases, we evaluate these statistics at the monthly frequency, which we believe is

\[8\] We compute the standard deviation with a stated zero mean (i.e., based on squared returns) to prevent relying on mean returns estimated with large error over short time windows. In all cases, we require 270 trading days of data before we form volatility-scaled returns. This ensures that the slowest volatility estimate (using exponential-decaying weights with a 90-day half-life) has at least three half-lives’ worth of data to use as a warm-up.

\[9\] Andersen et al. (2003) showed that realized intraday volatility predicts daily return volatility well for a number of currencies.

\[10\] This is the period with consistently liquid trading conditions over the full sample period and across both securities. Adding up the squared overnight return leads to slightly less persistence (not reported), which is consistent with the findings of Bollerslev et al. (2018), who reported greater persistence of intraday volatility.

\[11\] Saturdays were partial trading days before 1952. We doubled the squared returns on Saturdays before calculating the exponentially weighted moving average for estimating volatilities.
more relevant to investors than, for example, the daily frequency. More precisely, we use 30-calendar-day (in the case of Equities All U.S.) or 21-weekday overlapping returns. Only the mean and turnover of the notional exposure are evaluated using daily data.

In the exhibits with tables, we report the Sharpe ratio both gross and net of transaction costs. We use the following transaction cost estimates, expressed as a fraction of the notional value traded: 1.0 bp (or 0.01%) for equities, 0.5 bp for bonds, 0.5 bp for credit, 1.0 bp for gold, 2.0 bps for oil, and 3.5 bps for copper. In exhibits with figures, we will just show returns gross of transaction costs, but results are very similar on a net basis.

The amount of trading needed to implement the volatility scaling can be inferred from the mean notional exposure times the turnover of the notional exposure. The latter is obtained as the mean absolute daily exposure change, annualized and divided by twice the mean exposure. That is, turnover is expressed as the annual number of roundtrips of the mean exposure. We do not consider turnover incurred from rolling futures or forwards contracts. Notice that an unscaled position in a particular asset will have zero turnover.

The volatility of the rolling one-year realized volatility (i.e., vol of vol) is the statistic that most directly measures the extent to which the volatility scaling results in more-constant risk exposure.

The mean shortfall is the realized counterpart of expected shortfall, also known as conditional value at risk. In contrast, the usual value-at-risk metric simply measures how bad the $p$-th percentile of the returns distribution is; that is, it ignores returns below the $p$-th percentile. The mean shortfall is preferred because it uses all the returns below the $p$-th percentile. The mean shortfall measures left-tail behavior, which is most relevant for investors. However, we will also show the mean exceedance, the equivalent metric for the right tail, to illustrate how volatility scaling cuts both the left and the right tails.

Although skewness and kurtosis are commonly reported, we have omitted them from our main analysis for two reasons. First, skewness and kurtosis are very sensitive to outliers because their computation involves taking the third and fourth power of returns, respectively. Second, skewness is affected by both left- and right-tail behavior, but investors are likely much more concerned with the left tail of the return distribution. Readers interested in more detail can refer to the online supplement, where we also discuss tail skewness, tail kurtosis, and (maximum) drawdown.

**U.S. EQUITIES**

**Unscaled Equity Returns since 1926**

The top three panels of Exhibit 3 present daily, monthly, and annual excess equity returns. It is evident

---

**EXHIBIT 2**

**Performance Statistics**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharpe ratio</td>
<td>Ratio of the mean and standard deviation of the excess return (annualized)</td>
</tr>
<tr>
<td>Mean notional exposure</td>
<td>Mean daily exposure</td>
</tr>
<tr>
<td>Turnover notional exposure</td>
<td>Mean absolute daily exposure change, annualized and divided by twice the mean exposure</td>
</tr>
<tr>
<td>Vol of vol</td>
<td>Standard deviation of the rolling one-year standard deviation of 21-weekday or 30-calendar-day overlapping returns</td>
</tr>
<tr>
<td>Mean shortfall (left tail)</td>
<td>Mean of returns below the $p$-th percentile ($p = 1$ and $5$ will be considered)</td>
</tr>
<tr>
<td>Mean exceedance (right tail)</td>
<td>Mean of returns above the $p$-th percentile ($p = 95$ and $99$ will be considered)</td>
</tr>
</tbody>
</table>

Notes: As a default, we compute the Sharpe ratio, volatility of volatility (vol of vol), mean shortfall, and mean exceedance using a one-month (21 weekdays or 30 calendar days) evaluation frequency. The mean and turnover of the notional exposure are evaluated using daily data.
EXHIBIT 3
Equities All U.S. Returns (1926–2017)

Notes: The first three panels of the exhibit are daily, monthly, and annual U.S. equity returns in excess of the T-bill rate for the 1926–2017 period. No volatility scaling has been applied. The bottom panel shows cumulative (nominal, excess, and real) returns on a log-scale.
that volatility tends to cluster, being persistently high during the 1930s (Great Depression), the early 2000s (following the bursting of the tech bubble), and 2007–2009 (global financial crisis). The most negative day, October 19, 1987 (Black Monday), is also clearly visible.

The bottom panel of Exhibit 3 contrasts nominal, excess, and real cumulative returns, with the nominal return markedly higher during the high-inflation 1970s and 1980s. Notice that the excess returns (the focus in this article) are slightly below real returns. This is intuitive because the short-rate deducted to arrive at excess returns captures both an inflation component (the larger effect empirically) and a real rate component.13

Persistence of Equity Volatility

In Exhibit 4, we sort returns into quintiles based on the previous month’s volatility. The left panel shows the mean excess return and the right panel the volatility (both annualized) for the subsequent month. The persistence of volatility is evident in the right panel. However, the mean return shows no clear pattern across different quintiles (left panel). Thus, expected returns do not seem to reflect the persistence in volatility; that is, they do not provide a substantially higher reward in the case of predictably higher volatility. This is the first indication that volatility scaling may improve the Sharpe ratio of a long equities investment, as we will establish in the next subsection.

To further illustrate that equity volatility clusters, we show in the second online supplement the autocorrelation of the monthly squared volatility (i.e., variance) of daily returns.

Performance of Volatility-Scaled Equity Returns

In Exhibit 5, we show performance statistics for unscaled (top row) and volatility-scaled (other rows) Equities All U.S. investments (1927–2017). We use exponentially decaying weights for the volatility estimate with a half-life indicated in parentheses in the first column of the table.

The Sharpe ratio improves from 0.40 (unscaled) to between 0.48 and 0.51 (volatility scaled) and is not very sensitive to the choice of volatility estimate.14,15 The gross and net Sharpe ratio are the same for the reported precision, with the caveat that we use transaction cost estimates.

13 We use Consumer Price Index (all urban consumers) data from the U.S. Department of Labor, Bureau of Labor Statistics. See https://fred.stlouisfed.org/graph/?id=CPIAUCSL,CPIAUCNS.

14 To test for the statistical significance of this improvement, we run a regression of volatility-scaled daily returns (20-day half-life) on unscaled returns. We find an intercept of 0.64 bp with a t-stat of 3.05 (Newey–West corrected with 30 lags). The R² of the regression is 0.73.

15 Dopel and Ramkumar (2013) and Moreira and Muir (2017) also found that volatility targeting improves the Sharpe ratio for equities since 1927.
The Impact of Volatility Targeting

reflective of the current environment and apply this to the full history. As described in the first section, we use a rolling one-month (30 calendar days) evaluation frequency. We find very similar results for a rolling three-month (90 calendar days) evaluation frequency (not reported).

The mean exposure is higher with volatility scaling to achieve the same 10% full sample realized volatility because larger exposures are taken during low-volatility episodes. The turnover is zero for the unscaled investment and ranges from about five times a year for the most responsive and reactive volatility estimates to less than once a year for the least responsive volatility estimates.

Both the vol of vol and left tail (mean shortfall) materially improve with volatility scaling, and the improvement is greatest for the most responsive volatility estimates. The right tail (mean exceedance) is also, not surprisingly, reduced with volatility scaling. Hence, volatility scaling cuts both the left and right tails. Consistent with this, in the online supplement, we show

---

16For example, for the case of a 10-day half-life, the costs are about 2 × 1 bp × 71% × 4.66 = 0.066% per year for a 10% volatility strategy. The unrounded gross and net Sharpe ratios are then 0.4831 and 0.4766, but both are 0.48 after rounding.

17See also Hocquard, Ng, and Papageorgiou (2013), who studied how volatility targeting changes the tail-risk properties of an equity portfolio since 1990.
that kurtosis is much reduced when volatility scaling. The effect on skewness is more ambiguous because both the left and right tails are cut. Also the maximum drawdown is lower with volatility scaling.

In Exhibit 6, we further compare unscaled and volatility-scaled returns, the latter of which uses a volatility estimate based on a half-life of 20 days. In the left panel, we plot the cumulative return, which shows that the volatility-scaled investment generally outperformed, except during the middle part of the sample period. The impact of volatility scaling is illustrated in the right panel, where we depict the rolling one-year realized volatility for both unscaled and volatility-scaled 30-day overlapping returns. The realized volatility of volatility-scaled returns is much more stable over time. This is also evident from the vol of vol metric (i.e., the standard deviation of the rolling one-year realized volatility) reported in the legend: 4.6% for unscaled returns versus 1.8% for volatility-scaled returns.

Performance of Volatility-Scaled Equity Returns, Robustness across Subsamples and Industries

In Exhibit 7, we show the key statistics visually. We include equities broadly over the full sample period (1927–2017), equities broadly over three 30-year subsample periods, and 10 industry portfolios over the full sample period.
period. Subplots are such that in all cases observations above the dashed diagonal line correspond to situations in which volatility scaling improves the statistic.

The Sharpe ratio improves in all cases, except during the 1957–1987 subsample period. The vol of vol and mean shortfall consistently and materially improve with volatility scaling.

**Performance of Volatility-Scaled S&P 500 Futures Returns and the Use of Intraday Data**

Exhibit 8 explores the benefit of using higher-frequency S&P 500 futures data (five-minute intervals from 1988) to estimate volatility. The top panel uses daily data, and the bottom panel uses five-minute bars. The Sharpe ratio, vol of vol, and mean shortfall (left tail), and mean exceedance (right tail) use a rolling one-month (21 weekdays) evaluation frequency. To facilitate comparison, both the unscaled and volatility-scaled returns are shown at 10% full-sample volatility.

**OTHER ASSETS**

**Unscaled Bond Returns since 1926**

As we did for equities, we begin by examining bond returns since 1926, with (proxy) daily returns starting in 1962. In the top three panels of Exhibit 9, we plot the daily, monthly, and annual excess returns. Returns were less volatile pre-1980 and much less so pre-1967. Therefore, it seems that bond markets have gone through different volatility regimes, lasting multiple decades. This is important to note because structural breaks may render the evaluation of a bond volatility targeting strategy that includes data from before the mid-1980s less appropriate. In contrast, equity markets experience clusters of volatility, but without clear structural breaks. From the bottom panel of Exhibit 9, we can see that excess returns were flat for the first 55 years of our sample period and experienced a 40-year drawdown, ending in the 1980s.

**Persistence of Bond Volatility**

As we did for Exhibit 4, in Exhibit 10 we sort returns into quintiles based on the previous month’s volatility. The left panel shows the mean excess return and the right panel the volatility (both annualized) for the subsequent month. Volatility is persistent (right panel). However, in contrast to equities in Exhibit 4, the mean bond returns are not similar across different quintiles (left panel); rather, the returns are much higher in the high-volatility quintile.
**EXHIBIT 9**

**U.S. Bond Returns (1926–2017)**

10-Year Bond Daily Excess Returns

10-Year Bond Monthly Excess Returns

10-Year Bond Annual Excess Returns

10-Year Bond Cumulative Monthly Returns with Log-Scale

Notes: The top three panels show (proxy) daily, monthly, and annual U.S. bonds returns in excess of the T-bill rate for the 1926–2017 period. No volatility scaling has been applied. The bottom panel shows cumulative (nominal, excess, and real) returns against a log-scale.
Thus, it is not obvious that volatility scaling will affect the Sharpe ratio of a long bond investment.

To further illustrate that bond volatility clusters, we show in the second online supplement the autocorrelation of the monthly squared volatility (i.e., variance) of daily returns.

Performance of Volatility-Scaled Bond Returns (since 1963)

In Exhibit 11, we report the performance statistics for U.S. bonds over the 1963–2017 period. Consistent with the quintile analysis displayed in Exhibit 10, volatility scaling decreases the Sharpe ratio over this period.

The reason is straightforward: The 1960–1980 period was characterized by both negative returns and low volatility, so a volatility targeting approach would lead to relatively large exposures during this extended bond bear market. Volatility targeting does lead to a lower vol of vol during this period, as it did for equities.

In Exhibit 12, we contrast the cumulative return and realized volatility for an unscaled and volatility-scaled bond investment (for a comparison of the 1% and 5% left-tail return distributions, we refer to the online appendix). In all cases, the volatility scaling is done using exponentially decaying weights with a half-life of 20 days. Visible from the top–right panel is that the unscaled bond investment indeed has a low realized

---

**Exhibit 10**


<table>
<thead>
<tr>
<th>Volatility Quintile</th>
<th>Next Month Ann. Return</th>
<th>Next Month Ann. Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>1.23</td>
</tr>
<tr>
<td>2</td>
<td>0.00</td>
<td>1.56</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>1.89</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>2.22</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>2.55</td>
</tr>
</tbody>
</table>

Note: The left panel shows the mean excess return and the right panel the volatility (both annualized) when sorting on the previous month’s volatility for U.S. bonds over the 1962–2017 period.

**Exhibit 11**


<table>
<thead>
<tr>
<th>Scaling</th>
<th>Sharpe Ratio</th>
<th>Notional Exposure</th>
<th>Vol of Vol</th>
<th>Left Tail</th>
<th>Right Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross</td>
<td>Net</td>
<td>Mean</td>
<td>Turnover</td>
<td>1%</td>
</tr>
<tr>
<td>Unscaled</td>
<td>0.25</td>
<td>0.25</td>
<td>127%</td>
<td>0.00</td>
<td>3.9%</td>
</tr>
<tr>
<td>Scaled (10-day half-life)</td>
<td>0.05</td>
<td>0.04</td>
<td>180%</td>
<td>4.96</td>
<td>2.1%</td>
</tr>
<tr>
<td>Scaled (20-day half-life)</td>
<td>0.06</td>
<td>0.06</td>
<td>179%</td>
<td>2.51</td>
<td>2.1%</td>
</tr>
<tr>
<td>Scaled (40-day half-life)</td>
<td>0.08</td>
<td>0.08</td>
<td>177%</td>
<td>1.27</td>
<td>2.2%</td>
</tr>
<tr>
<td>Scaled (60-day half-life)</td>
<td>0.08</td>
<td>0.08</td>
<td>174%</td>
<td>0.86</td>
<td>2.4%</td>
</tr>
<tr>
<td>Scaled (90-day half-life)</td>
<td>0.09</td>
<td>0.09</td>
<td>170%</td>
<td>0.58</td>
<td>2.6%</td>
</tr>
</tbody>
</table>

Notes: Performance statistics detailed in Exhibit 2 for U.S. bonds, proxied from daily yield data (1963–2017). The gross and net (of estimated costs) Sharpe ratio, vol of vol, mean shortfall (left tail), and mean exceedance (right tail) use a rolling one-month (21 weekdays) evaluation frequency. To facilitate comparison, both the unscaled and volatility-scaled returns are shown at 10% full-sample volatility.
volatility during the 1964–1980 period. The underperformance of the volatility-scaled investment is solely due to the pre-1980 period.

One could argue that bond markets underwent a structural change in the mid-1980s, with monetary policy more geared toward inflation targeting. Hence, the post-1988 sample period considered in the next subsection may be more representative for today’s bond markets—although this period appears to be a single regime with declining rates.

**Performance of Volatility-Scaled Bond Returns (since 1988)**

In Exhibit 13, we report the performance statistics over 1988–2017, for which we have U.S. 10-year Treasury futures data, both daily and intraday. We see that in general the vol of vol is much lower with volatility scaling, but the Sharpe ratio and mean shortfall (left tail) are similar. Using intraday data for the volatility estimate produces a slight improvement.

In Exhibit 14, we contrast the cumulative return and realized volatility for an unscaled and volatility-scaled bond investment. In all cases, the volatility scaling is done using exponentially decaying weights with a half-life of 20 days. Consistent with Exhibit 13, the main difference between unscaled and scaled returns for Treasuries over the 1988–2017 period is the lower vol of vol when using volatility scaling (top-right panel).

**Performance of Volatility-Scaled Futures, Forwards, and Credit Returns**

We now turn our attention to 50 futures and forwards across global equities, fixed income, currencies (all against the U.S. dollar), and commodities, as well as credit. Performance statistics with and without volatility scaling are depicted in Exhibit 15. The Sharpe ratio improves slightly for equity indexes and credit when using volatility scaling, but it is similar for other assets. Credit is related to equities in the sense that both are exposed to firms’ cash flow risk (i.e., both are risk assets), and so it is intuitive to see credit exhibit similarity to equities in this respect. The vol of vol and mean shortfall improve materially for almost all assets with volatility scaling.

**Portfolios**

So far, we have considered single-asset investments. In this section, we turn our attention to the 60–40 equity–bond balanced portfolio, and in the online...
The Impact of Volatility Targeting

EXHIBIT 13

<table>
<thead>
<tr>
<th>Scaling</th>
<th>Sharpe Ratio</th>
<th>Notional Exposure</th>
<th>Vol of Vol</th>
<th>1%</th>
<th>5%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unscaled</td>
<td>0.64</td>
<td>0.64</td>
<td>171%</td>
<td>0.00</td>
<td>2.5%</td>
<td>-7.3%</td>
<td>-5.5%</td>
</tr>
<tr>
<td>Volatility Used for Scaling Based on Daily Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaled (10-day half-life)</td>
<td>0.63</td>
<td>0.62</td>
<td>182%</td>
<td>4.33</td>
<td>1.2%</td>
<td>-7.5%</td>
<td>-5.6%</td>
</tr>
<tr>
<td>Scaled (20-day half-life)</td>
<td>0.63</td>
<td>0.63</td>
<td>182%</td>
<td>2.19</td>
<td>1.2%</td>
<td>-7.5%</td>
<td>-5.6%</td>
</tr>
<tr>
<td>Scaled (40-day half-life)</td>
<td>0.63</td>
<td>0.63</td>
<td>182%</td>
<td>1.11</td>
<td>1.4%</td>
<td>-7.5%</td>
<td>-5.6%</td>
</tr>
<tr>
<td>Scaled (60-day half-life)</td>
<td>0.63</td>
<td>0.63</td>
<td>181%</td>
<td>0.74</td>
<td>1.5%</td>
<td>-7.5%</td>
<td>-5.6%</td>
</tr>
<tr>
<td>Scaled (90-day half-life)</td>
<td>0.63</td>
<td>0.63</td>
<td>180%</td>
<td>0.49</td>
<td>1.6%</td>
<td>-7.5%</td>
<td>-5.6%</td>
</tr>
</tbody>
</table>

Volatility Used for Scaling Based on 5-Minute Data

<table>
<thead>
<tr>
<th>Scaling</th>
<th>Sharpe Ratio</th>
<th>Notional Exposure</th>
<th>Vol of Vol</th>
<th>1%</th>
<th>5%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaled (intraday, 5-day)</td>
<td>0.66</td>
<td>0.65</td>
<td>186%</td>
<td>3.80</td>
<td>1.1%</td>
<td>-7.4%</td>
<td>-5.4%</td>
</tr>
<tr>
<td>Scaled (intraday, 15-day)</td>
<td>0.64</td>
<td>0.64</td>
<td>186%</td>
<td>1.39</td>
<td>1.1%</td>
<td>-7.4%</td>
<td>-5.5%</td>
</tr>
<tr>
<td>Scaled (intraday, 25-day)</td>
<td>0.64</td>
<td>0.64</td>
<td>186%</td>
<td>0.87</td>
<td>1.2%</td>
<td>-7.5%</td>
<td>-5.5%</td>
</tr>
<tr>
<td>Scaled (intraday, 40-day)</td>
<td>0.63</td>
<td>0.63</td>
<td>185%</td>
<td>0.56</td>
<td>1.3%</td>
<td>-7.5%</td>
<td>-5.5%</td>
</tr>
<tr>
<td>Scaled (intraday, 60-day)</td>
<td>0.63</td>
<td>0.63</td>
<td>184%</td>
<td>0.38</td>
<td>1.4%</td>
<td>-7.6%</td>
<td>-5.5%</td>
</tr>
</tbody>
</table>

Notes: Performance statistics detailed in Exhibit 2 for U.S. 10-year Treasury futures (1988–2017). We consider volatility estimates based on daily data (top panel) and five-minute intraday data (bottom panel). The gross and net (of estimated costs) Sharpe ratio, vol of vol, mean shortfall (left tail), and mean exceedance (right tail) use a rolling one-month (21 weekdays) evaluation frequency. To facilitate comparison, both the unscaled and volatility-scaled returns are shown at 10% full-sample volatility.

EXHIBIT 14

Notes: Unscaled and volatility-scaled (exponential-weighted, 20-day half-life) 10-year Treasury futures returns for the 1988–2017 period. The left panel shows the cumulative return. The right panel shows the rolling one-year standard deviation of one-month (21 weekdays) overlapping returns. The standard deviation of the rolling one-year standard deviation is reported in parentheses in the legend. To facilitate comparison, both the unscaled and volatility-scaled returns are shown at 10% full-sample volatility.
supplement we show similar results for a 25–25–25–25 equity–bond–credit–commodity risk parity portfolio.\textsuperscript{18}

We will contrast three ways to implement such a portfolio:

1. Unscaled at both the asset and portfolio level
2. Volatility scaling at the asset level only
3. Volatility scaling at both the asset and portfolio level

In all cases the asset-level returns are subject to the full-sample scaling to 10% volatility, as discussed before, which means that as a starting point the allocation to the different asset classes is in proportion to full-sample volatility, and thus we can sensibly compare the different cases. For simplicity we assume portfolios are rebalanced to the target asset allocation mix each day.\textsuperscript{19}


\textsuperscript{19}For a discussion on the effect of the rebalancing frequency for a 60–40 balanced portfolio, see Granger et al. (2014).
In Exhibit 16, we report the performance statistics for the balanced 60–40 equity–bond portfolio, based on S&P 500 and 10-year Treasury futures return data. Because of the aforementioned asset-level scaling to 10% volatility in all cases, the 60–40 split here is in risk terms. The Sharpe ratio, vol of vol, and expected shortfall (left tail) all improve from asset-level volatility scaling and further improve from a second volatility scaling step at the portfolio level, which essentially adjusts for time variation in the correlation between different assets. Furthermore, the improvement in left-tail returns is greater than the reduction in right-tail returns.

In Exhibit 17, we contrast the cumulative return and realized volatility for an unscaled and volatility-scaled (at both the asset and portfolio level) balanced portfolio. Volatility scaling is done using exponentially decaying weights with a half-life of 20 days. Consistent with the results of Exhibit 16, both the cumulative return and vol of vol improve with volatility scaling.

---

**Exhibit 16**


<table>
<thead>
<tr>
<th>Asset Scaling</th>
<th>Portfolio Scaling</th>
<th>Sharpe Ratio</th>
<th>Notional Exposure</th>
<th>Left Tail</th>
<th>Right Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Gross Net</td>
<td>Mean Turnover</td>
<td>Vol of Vol</td>
<td>1%</td>
</tr>
<tr>
<td>Unscaled</td>
<td>Unscaled</td>
<td>0.80 0.80</td>
<td>153% 0.00</td>
<td>3.4%</td>
<td>−9.7%</td>
</tr>
<tr>
<td>Scaled (20-day half-life)</td>
<td>Unscaled</td>
<td>0.87 0.87</td>
<td>179% 2.24</td>
<td>2.2%</td>
<td>−8.0%</td>
</tr>
<tr>
<td>Scaled (20-day half-life)</td>
<td>Scaled (20-day half-life)</td>
<td>0.91 0.90</td>
<td>183% 3.95</td>
<td>1.3%</td>
<td>−7.3%</td>
</tr>
</tbody>
</table>

Notes: Gross and net (of estimated costs) Sharpe ratio, vol of vol, mean shortfall (left tail), and mean exceedance (right tail) statistics described in Exhibit 2 for the 60–40 equity–bond balanced portfolio. We contrast an unscaled portfolio with a portfolio with volatility scaling at the asset level only and a portfolio with volatility scaling at both the asset and portfolio level. The volatility scaling is done using exponentially decaying weights with a half-life of 20 days. We use a rolling one-month (21 weekdays) evaluation frequency. To facilitate comparison, both the unscaled and volatility-scaled returns are shown at 10% full-sample volatility.

**Exhibit 17**


Notes: Unscaled and volatility-scaled (exponential-weighted, 20-day half-life) 60–40 equity–bond balanced portfolio returns for the 1988–2017 period. The left panel shows the cumulative return. The right panel shows the rolling one-year standard deviation of one-month (21 weekdays) overlapping returns. The standard deviation of the rolling one-year standard deviation is reported in parentheses in the legend. To facilitate comparison, both the unscaled and volatility-scaled returns are shown at 10% full-sample volatility.

---

Asvanunt, Nielsen, and Villalon (2015) considered various strategies to reduce the size of tail events for 60–40 equity–bond portfolios. These include options-based approaches and shifting to a risk parity asset class allocation based on risk exposures.
WHY DOES VOLATILITY SCALING PARTICULARLY IMPROVE THE SHARPE RATIO OF RISK ASSETS?

In this section, we examine possible explanations for why volatility scaling improves the Sharpe ratio for risk assets, such as equities and credit, but has no effect on the Sharpe ratio of other assets. Our analysis suggests an answer that can be split into three parts: (1) only risk assets empirically display a so-called leverage effect, (2) the leverage effect effectively introduces some momentum, and (3) such a momentum overlay is beneficial for the Sharpe ratio. Indeed, we will show that the tendency of volatility scaling to introduce some momentum empirically explains much of the cross-sectional variation in the Sharpe ratio improvement when using volatility scaling.

Leverage Effect Is Confined to Risk Assets

Equities and credit display a leverage effect, which is the tendency of returns to have a negative contemporaneous correlation to changes in volatility. The classic explanation by Black (1976) is that a negative equity return leads to a higher firm debt-to-equity ratio (more leverage in the capital structure of the firm), which in turn means equity volatility should increase (holding constant the firm’s cash flow volatility).21

---

21See also Christie (1982) for a discussion of the leverage effect. Bekaert and Wu (2000) argued that there is also a volatility feedback effect, in which the causality is reversed compared to the leverage effect; volatility increases give rise to higher risk premiums and thus negative returns.
In Exhibit 18, we indeed observe this leverage effect empirically for equities and credit, but not for other assets. The top panels show results for Equities All U.S. (1926–2017), the three subsamples, and the 10 industry portfolios considered earlier in the section on U.S. equities. The bottom panels show results for credit and the 50 futures and forwards for the 1988–2017 sample period. The right panels show the leverage effect: a negative correlation between monthly observations of the return and the change in variance. The left panels show a very similar picture for the correlation between monthly observations of return and the level of the variance.

**Leverage Effect Introduces Some Momentum**

When applied to assets exhibiting the leverage effect, volatility scaling effectively introduces some time-series momentum into strategies. That is, negative returns tend to be followed by a reduction in the position size (because volatility is higher in that case), and positive returns tend to be followed by an increase in the position size (because volatility is lower in that case).

In Exhibit 19, we show more explicitly the assets for which volatility scaling leads to changes in position sizes that are in the momentum direction (i.e., smaller after negative returns, bigger after positive returns). Specifically, we show the correlation between the reciprocal of the volatility estimate (which is proportional to position sizing when volatility scaling is applied) and the past 21-, 65-, 130-, and 260-day returns (1-, 3-, 6-, 12-month momentum) in the four panels, respectively. The volatility estimate is based on exponential weighted returns with a 20-day half-life. We consider credit and the 50 futures and forwards for the 1988–2017 sample period.

**Notes:** Correlation between the reciprocal of the volatility estimate (which is proportional to position sizing when volatility scaling is applied) and the past 21-, 65-, 130-, and 260-day returns (1-, 3-, 6-, 12-month momentum) in the four panels, respectively. The volatility estimate is based on exponential weighted returns with a 20-day half-life. We consider credit and the 50 futures and forwards for the 1988–2017 sample period.
In fact, for other assets the correlation is predominantly negative, introducing a bet on mean reversion.

**Linking Momentumness of Returns and the Impact of Volatility Scaling on the Sharpe Ratio**

The final part of our investigation is to link directly the cross-sectional differences in the impact of volatility scaling on the Sharpe ratio and asset return properties.

The evidence suggests that time-series momentum strategies have historically performed well; see, for example, Hamill, Rattray, and Van Hemert (2016). In Exhibit 20, we show that it is indeed the *momentumness* of volatility scaling that explains a large part of the cross-sectional variation in the Sharpe ratio improvement when using volatility scaling for the various assets considered. We find the shorter-term, one-month momentum of returns to be most relevant here. Using a 20-day half-life for volatility scaling, the $R^2$ is 45%. For a slower volatility estimate using a 90-day half-life, the $R^2$ is even higher at 60%.

Understanding the precise relation between returns, volatility, and the Sharpe ratio of volatility-scaled returns is a topic of ongoing research.

**CONCLUDING REMARKS**

Recent research has demonstrated that volatility scaling improves the Sharpe ratios of equity portfolios. Our research shows it is a mistake to extrapolate this effect to other assets and that this boost is specific to so-called risk assets (e.g., equity and credit) or portfolios that have a sizable allocation to these risk assets. That is, for other assets, such as fixed income, currencies, and commodities, the effect of a simple volatility scaling on the Sharpe ratio is negligible.

Although the Sharpe ratio is important, most investors have broader investment objectives. We show that volatility scaling has one unambiguous effect across assets and asset classes: It reduces the likelihood of extreme returns (and the volatility of volatility). In particular, the lower probability of very negative returns (left-tail events) is valuable for investors.

Although we provided a detailed historical account of the impact of volatility targeting across 60+ assets and

---

**Exhibit 20**

Improvement in Sharpe Ratio When Volatility Scaling vs. Correlation (past returns, $1/vol$) for Various Assets

![Graph showing improvement in Sharpe ratio](image)

Notes: Improvement from volatility scaling (vertical axis) versus the momentumness of volatility scaling, determined as the correlation between the past 21-day returns and the reciprocal of the volatility estimate, on the horizontal axis. The volatility estimate is based on exponentially weighted squared returns with a 20-day half-life (left panel) and 90-day half-life (right panel). We consider credit and the 50 futures and forwards for the 1988–2017 sample period.

---

22Dachraoui (2018) also argues there is a link between the presence of a leverage effect and the extent to which volatility targeting improves the performance for an asset.
two multiasset portfolios, some topics are beyond the scope of this article. We will comment on three.

First, the detailed analysis for equity and bonds was done for U.S. assets, for which we have the longest daily return history. A caveat of this approach is that the United States is an ex post winner in the sense that over the past century it had robust economic growth and no major war on its mainland. This may particularly matter for bonds, which can start to resemble a credit investment when the creditworthiness of a government is questioned by investors. As such, our finding that volatility scaling does not meaningfully improve the Sharpe ratio of a bond investment should also be caveated, and going forward volatility scaling may improve the Sharpe ratio of bonds that unexpectedly start to behave in a more credit-like manner.

Second, although the focus in this article was on volatility scaling, there are other methods with the potential to improve the risk management of a long portfolio. Hamill, Rattray, and Van Hemert (2016) showed that trend-following strategies tend to work particularly well at times of equity and bond market sell-offs. Hence, a trend-following overlay may further improve the risk and return of a long portfolio.

Finally, although we explored intraday data for S&P 500 and Treasury futures and found some benefits vis-à-vis daily data, we believe we only scratched the surface of this topic in this article. More assets now have good-quality intraday data, and for more hours of the day. In addition, advances in statistical modeling may help us to use the intraday data to get more timely estimates; see, for example, Noureldin, Shephard, and Sheppard (2012) for a discussion on multivariate high-frequency-based volatility models.

ACKNOWLEDGMENTS

The authors would like to thank Xinuo Chen, Antoine Forterre, Nick Granger, Keith Haydon, Tim Meakins, Graham Robertson, Darrel Yawitch, and participants at the 2018 J.P. Morgan Macro Quantitative and Derivatives Conference for valuable comments.

REFERENCES


*To order reprints of this article, please contact David Rowe at d.rowe@pageantmedia.com or 646-891-2157.*

---

The Journal of Portfolio Management  
Fall 2018