Russell Investments

Viewpoint

Russell Research

By: Bob Collie, FIA, Chief Research Strategist, Americas Institutional
Charles Anselm, CFA, Senior Portfolio Manager

MARCH 2013

Measuring the success of a managed volatility investment strategy

Finding metrics that fit

A new breed of investment mandates is growing in popularity. Going by names such as multi-asset, outcome-oriented and solutions-based investing, these mandates build a closer tie between the investor's ultimate objectives and the instructions given to the investment manager. They replace the manager's traditional goal – achieving incremental returns over those of the broad market – with a goal that has a more direct connection to the investor's end objectives.

There are many implications of such a change. In this paper, we focus on one: how to report investment results. We approach the question by considering one example of a multi-asset strategy: a managed volatility portfolio. Although the reporting metrics we propose are specific to this mandate, the principles we describe apply more widely. We also broaden the discussion, therefore, to consider some of the implications of this approach for multi-asset portfolios in general. In particular, we will argue that multi-asset mandates need metrics that fit the specific objectives in mind, and that no single metric is necessarily sufficient to capture all of the goals of a multi-asset mandate.

We will argue that multi-asset mandates need metrics that fit the specific objectives in mind, and that no single metric is necessarily sufficient to capture all of the goals of a multi-asset mandate.

Managed volatility

Managed volatility strategies[1] use a combination of techniques to reduce risk in a portfolio. Specifically, managed volatility strategies aim to achieve lower volatility and reduced

[1] For more detail of managed volatility strategies, see Thomas and Collie (2012).

downside in falling markets, as compared to managers' typical investment mandates, while still delivering comparable levels of return.

The objective of the managed volatility approach is defined in terms of both risk and return. This adds a new dimension to a more traditional investment mandate's goal of simply delivering better returns than those of a market benchmark.[2] This leads to the question of how best to measure the results achieved.

The reporting of investment results is at present almost universally based on the comparison of portfolio returns to those of a broad market index (or combination of indices). This comparison gives the asset owner an easy measure of the investment manager's success against the goal of delivering higher returns than those of the broad market – i.e., did the manager beat the passive alternative?[3] But that approach is inadequate where there is a dual objective covering not only returns but also risk.

Measuring the things that you want to be managed

We must start by reminding ourselves that even in a more traditional investment mandate, the market-based benchmark is more a reflection of what is realistically achievable than it is of the investor's desired outcome *per se*. Hence, we should set aside for the moment the notion that success is when the portfolio return exceeds the performance benchmark, and failure is when it does not. This notion is the basis for just about all performance reporting at present. The simplicity of this approach is clearly appealing. And, since higher returns are better than lower, it might seem that this notion is unarguable. For a managed volatility portfolio, however, it captures only one aspect of the mandate. Yes, higher returns are better than lower, but less risk is better than more risk, and that, too, must be reflected in how the results of the portfolio are to be reported.

Yes, higher returns are better than lower, but less risk is better than more risk, and that, too, must be reflected in how the results of the portfolio are to be reported.

To many readers, it might seem that the natural extension of the traditional return-based objective is a risk-adjusted return metric that collapses the two dimensions into one. This would allow us to retain the simplicity and clarity of the existing approach: if the risk-adjusted return beats that of the policy benchmark, the investment manager has added value.

Alas, the relationship between risk and return is not so simple. While some versions of risk – such as volatility and drawdowns – can be measured, these are really only proxies for what the investor truly means by "risk"; likewise, the trade-off between risk and return is hard to capture definitively. What is more, valuable information is lost in the process of collapsing two dimensions into one: a managed volatility strategy might be regarded as having succeeded if it delivers returns equal to those of the market over a certain period, but with less risk, and it similarly might be regarded as having succeeded if it delivers returns above those of the market, without having taken on more risk. It may even be regarded as having succeeded if it delivers returns below those of the market, if they are achieved with less risk, or if it takes on risk greater than that of the market, if the returns are commensurately strong. But these situations are different from one another, and the outcomes represent different types of success. Good reporting should distinguish between them. So a single measure of risk-adjusted return is inevitably unsatisfactory.

The problem is easily solved if we are willing to abandon a second unhelpful notion that has become entrenched: that reporting must collapse down to a single metric. While that is certainly convenient, it is hardly essential. There are, indeed, very few areas of our lives where success can truly be measured on a single dimension. Are you good at your job?

[2] For reasons explained in the appendix, we have ignored here the tracking error targets that are usually appended to return objectives.

[3] In the case of a passive mandate, the intent is to match rather than beat the market return. The metric used to measure success – performance relative to the policy benchmark – is the same.

How was your last vacation? Is this *Viewpoint* helpful?[4] While any of these things can be judged to be a success or a failure, that judgment is based on more than one input. In most situations, we human beings really don't have that much of a problem with using more than one metric at a time; we need to start doing that with investment results.

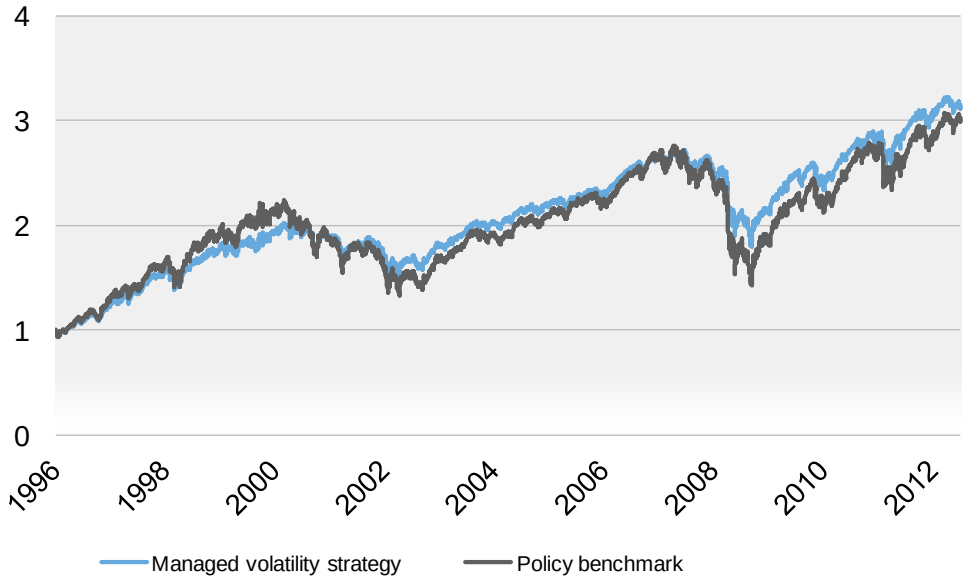
The measures of success of a managed volatility strategy

Let us apply this principle to the case of the managed volatility portfolio.[5] As in Thomas and Collie (2012) we will consider a portfolio that has been designed with the intention of delivering returns comparable to those of an 85/15 mix of U.S. large cap equities and bonds,[6] but doing so with less risk.

We show below the back-tested performance of the managed volatility strategy and of the 85/15 mix.

Chart 1: Managed volatility strategy: growth of a dollar, July 1996–December 2012

Growth of a dollar



In most situations, we human beings really don't have that much of a problem with using more than one metric at a time; we need to start doing that with investment results.

Examining the graph, we can see that over this period, the strategy has been largely successful in achieving what it set out to do. The portfolio value did not fluctuate as much as that of the policy benchmark, and during periods of falling markets (Sep-2000 to Mar-2003, and May-2008 to Mar-2009), the managed volatility strategy fell less. The cumulative return

[4] One arena where a single dimension does generally define success is the field of sports, and it seems quite likely that this simplicity is one reason why sports are so popular, offering as they do a welcome contrast to the messy and nuanced goals we must deal with in just about everything else we do. In addition to being multidimensional, most aspects of our lives demand some degree of qualitative, subjective assessment. However, investment returns, like sports scores, can be measured both quantitatively and objectively. In the case of investment returns, this objectivity is probably less important than it might at first seem. Just because something is mathematically correct does not mean that the numbers tell the full story. A "final score" can contribute to the illusion that no interpretation is required. In reality, both the soft measures and the objective measures are simply a foundation for judgments to be made.

[5] The managed volatility strategy return has been calculated as the composite of a one-third allocation (rebalanced monthly) to each of a defensive equity strategy (as represented by the Russell 1000 defensive equity index), a call overwriting strategy (as represented by the CBOE S&P 500 buy-write index, commonly known as the BXM index) and a volatility-responsive asset allocation strategy (which invests in the Russell 3000 index and the Barclays U.S. Treasury index in varying weights, with a higher allocation to the equity market following periods of low market volatility and a higher allocation to fixed income following periods of high market volatility, following the methodology of Collie, Sylvanus and Thomas (2011)).

[6] Specifically, the policy return is calculated from a mix of 85% Russell 1000® Index and 15% Barclays U.S. Treasury Index, Rebalanced monthly

over the entire period (Jun-1996 to Dec-2012) was, however, slightly above that of the policy benchmark.

How do we capture this visual assessment with appropriate performance metrics?

For this particular strategy, the first metric to use would be volatility, or the standard deviation of returns. This is in recognition of the central role of volatility management in the portfolio objectives. Volatility is an imperfect measure of risk and an incomplete statement of the portfolio's full list of objectives, but that is OK; it is not the only metric we will use. A more practical challenge this metric presents is that at least 20 data points are needed to permit a robust calculation of standard deviation. So for periods of less than two years, it is the daily (rather than monthly) return series that must be used as the basis for this calculation.

A second measure of portfolio risk is drawdown, or the largest peak-to-trough drop in value that the portfolio experienced. Even though this ignores portfolio performance at all times except for the one peak-to-trough episode, it is closely linked to what investors often mean when they talk of "risk," since it focuses on downside and loss of value. For the same reason, it is a good complement to the volatility metric as a measure of risk.

Next, we turn to returns. The return earned on the portfolio relative to the policy benchmark may no longer be the sole gauge of success, but it remains an important component of it. So this is our third metric.

Finally, we show the Sharpe ratio, which is the ratio of the portfolio return to the portfolio volatility (both based on returns in excess of cash). We noted above that a risk-adjusted return measure would be unsatisfactory as the sole measure of success. As a complement to the other metrics listed here, however, a risk-adjusted return measure (such as the Sharpe ratio) can play a useful supplementary role as a gauge of whether the amount of return that has been earned is commensurate with the risk taken. As in the case of volatility, there is a technical challenge; in this case, the challenge is associated with periods where the return on the portfolio is negative. For such periods, Sharpe ratios can be confusing, as illustrated in the sidebar, Exhibit A. For that reason, we have replaced negative Sharpe ratios with "N/A" in our recommended reporting structure.

Even though [drawdown] ignores portfolio performance at all times except for the one peak-to-trough episode, it is closely linked to what investors often mean when they talk of "risk," since it focuses on downside and loss of value.

Exhibit A

A higher Sharpe ratio normally implies either that returns were higher or volatility was lower. Hence in the hypothetical example below, portfolios B and C both achieved a Sharpe ratio superior to that of Portfolio A, one doing so by generating a higher return and the other through lower volatility.

Period 1	Portfolio A	Portfolio B	Portfolio C
Return (above cash)	6%	8%	6%
Volatility	12%	12%	9%
Sharpe ratio	0.50	0.67	0.67

However, if returns are negative, this pattern no longer holds. In period 2, shown below, Portfolios B and C each achieve the same Sharpe ratio, yet Portfolio B was inferior to Portfolio C, with both a lower return and higher volatility.

Period 2	Portfolio A	Portfolio B	Portfolio C
Return (above cash)	-6%	-8%	-6%
Volatility	12%	12%	9%
Sharpe ratio	-0.50	-0.67	-0.67

So the interpretation of a negative Sharpe ratio is difficult.

A sample report

Based on these metrics, we would therefore report the performance shown in Chart 1 as follows:

	Jan'00 - Dec'00 (1-Yr)	Jan'10 - Dec'12 (3-Yr annualized)	Jan'08 - Dec'12 (5-Yr annualized)	Jul'96 - Dec'12 (Full Period annualized)
Volatility. Objective: 80% or less volatility than the benchmark.				
Managed Vol. Strategy	13.87%	11.43%	15.49%	12.63%
Benchmark	19.50%	15.54%	22.07%	17.43%
Ratio	71.12%	73.55%	70.17%	72.47%
Drawdown. Objective: 80% or less drawdown than the benchmark.				
Managed Vol. Strategy	-7.23%	-10.53%	-19.21%	-13.04%
Benchmark	-14.44%	-15.58%	-27.85%	-20.52%
Ratio	50.04%	67.60%	68.97%	63.55%
Return. Objective: Return in line with the benchmark over longer periods.				
Managed Vol. Strategy	4.31%	8.30%	3.26%	7.21%
Benchmark	-4.69%	10.54%	2.74%	6.94%
Sharpe Ratio. Objective: Sharpe Ratio above that of the benchmark.				
Managed Vol. Strategy	N/A	0.71	0.18	0.39
Benchmark	N/A	0.67	0.10	0.27

The calendar year 2000 is shown here as an example of a period where the benchmark return was negative. The other periods shown are the most recent 3- and 5-calendar year periods and the full period of backtested history.
Hypothetical example provided for illustrative purposes only and not meant to reflect any actual investment.

For the full 16.5-year period in question, this report confirms the conclusion reached earlier: the portfolio was successful against each of the relevant metrics. Any one of the above metrics (although individually accurate) may not fully convey the true outcome of the portfolio. When viewed together, the metrics provide a clearer sense of the outcomes delivered by this strategy, i.e., long-term returns that are similar to the benchmark, but with lower volatility and lower drawdowns relative to the benchmark.

Clearly, this will not always be the case. In particular, this strategy is not expected to participate fully in market rallies: its defensive nature means that it is expected to underperform relative to the policy benchmark at times when the market delivers strong performance. As with all investment performance measurements, results should be expected to vary considerably over time, and results over short time periods can be especially variable. So the task of interpreting the results requires judgment on the part of the asset owner – all the more so because there are now four metrics, not just one, to take into account.

Any one of the above metrics (although individually accurate) may not fully convey the true outcome of the portfolio. When viewed together, the metrics provide a clearer sense of the outcomes delivered by this strategy.

Skill? Or luck?

So the numbers do not tell the full story, and evaluation requires the application of judgment on the part of the asset owner. Here we must pause and draw a clear distinction between two separate questions that investment reporting should help to answer. The first question is: *Were the investor's objectives met?* The second: *Did the investment manager do a good job?*

These are clearly related, but they are not the same. Achievement of the investor's objectives depends not only on how well the investment manager does, but also on factors over which a manager has no control. Most importantly, when investment markets in general deliver high returns, it is relatively easy to meet most investors' objectives. But in a year such as 2008, the best that could realistically be asked for – in hindsight – was

limitation of losses. Indeed, it is largely investors' desire to focus on the things under investment managers' control, and hence to achieve a fair measure of their value-add, which led to the prevalence of the benchmark-relative approach in the first place.

As noted at the start of this paper, the trend today is toward a greater focus on the investor's outcome, and asset owners want investment managers to think more broadly. Reporting, which has come to concentrate almost exclusively on the question of how good a job a manager did, needs to reflect this trend. As this happens, the need for application of judgment will increase. In our opinion, that's not a bad thing. The focus on benchmark-relative returns had become so great that it was arguably having a negative impact on investors' ability to achieve their objectives. The first – and ultimately more important – of the questions above (*Were the investor's objectives met?*) was overlooked. And the traditional benchmark-relative return number has never been as definitive an answer to our second question (*Did the investment manager do a good job?*) as it appeared to be: look at even the best long-term performance records, and you will find several three-year periods of returns trailing the benchmark.

The use of multiple metrics makes the distinction between our two questions more clear, makes the task of separating the effects of manager skill and luck more manageable, and hence provides a better foundation for the evaluation of both investment manager performance and progress toward the investor's goals.

Look at even the best long-term performance records, and you will find several three-year periods of returns trailing the benchmark.

On a related note, we would observe that even though the particular example we have used in this *Viewpoint* adds a risk dimension to reflect the investor's end objective, the managed volatility strategy remains essentially a market-relative mandate. This will not necessarily be the case for other multi-asset strategies. For example, a multi-strategy absolute-return mandate might target a particular level of return – say, 6% or 8% – and give a manager the freedom to invest in a wide range of strategies to achieve it. In that example, the importance of the policy benchmark would change with the time horizon: over the long term, the investor's objective (the absolute-return target) is most important, but over the short term, market experience will be the dominant element in performance. Thus, the market-relative benchmark is given greater weight over shorter time periods.

Conclusion: The right metrics matter

The trend toward multi-asset investing is an important development that creates a closer link between the end goals investors are working toward and the ways in which their portfolios are managed. Our analysis in this paper is based on a managed volatility strategy; other examples of multi-asset strategies include absolute-return mandates, liability-driven investing and retirement income products.

This trend has implications for the reporting and evaluation of investment performance. Each multi-asset portfolio is intended to achieve certain objectives, and hence requires reporting that is based on those specific objectives. In many cases, this will involve moving from reliance on a single measure of success – performance relative to a policy benchmark return – to the use of multiple metrics.

REFERENCES:

M. Thomas and R. Collie (2012, August). "Volatility management". *Russell Viewpoint*.

Appendix: A note on tracking error

Some readers may question our assertion that most reporting is currently based on a single measure of success and lacks any dimension of risk assessment. “What,” they might ask, “about tracking error?” It is true that most investment mandates specify a targeted level of tracking error, i.e., a limit on how great the standard deviation of the difference between the portfolio returns and the policy benchmark returns should be. We have ignored this in our main text, because we regard it as a measure that compounds – rather than balances – the investment manager’s incentive to place an inappropriate emphasis on the policy benchmark, at the expense of the investor’s objectives. Tracking error does not necessarily make it less likely that investors will meet their objectives. In any case, in practice, tracking error as a measure of investment manager success is paid so little attention that to call it secondary would be too generous.

While we do not consider tracking error to be among the objectives of a managed volatility strategy, there are other types of mandates in which it remains of interest. For example, those cases where the investment manager is intended to operate largely within a specific set of investments. While multi-asset mandates begin with the investor’s goals, they may still include some direction about how a particular portfolio is expected to contribute to meeting those goals – i.e., a broad definition of what the portfolio is expected to consist of. Tracking error is one way to measure the extent to which the investment manager is deviating from this core expected set of investments. For us, however, even in those cases, tracking error belongs well down on the list of the investor’s concerns.

For more information:

Call Russell at **800-426-8506** or
visit **www.russell.com/institutional**

First used: March 2013

USI-16182-03-16