TESTING HEDGE EFFECTIVENESS FOR FAS 133: THE VOLATILITY REDUCTION MEASURE

by Andrew Kalotay and Leslie Abreo, Andrew Kalotay Associates, Inc.

Corporations routinely use derivative products such as interest rate swaps to hedge business risk. For example, a financial institution with floating-rate assets might synthetically convert its fixed-coupon bond obligation into a floating-rate obligation by entering into an appropriate interest rate swap. But companies can also use derivative products for speculation, with unpredictable and possibly disastrous consequences, as in the case of the Procter & Gamble fiasco. The problem for shareholders is that this kind of exposure typically gets buried in footnotes to the corporation's financial statements, even if it has the potential—realized in the case of P&G—to have a material effect on reported earnings.

In an attempt to make a company's exposure to its derivative positions more transparent, the Financial Accounting Standards Board (FASB) has issued Statement 133 (and subsequent amendments). This statement requires full disclosure of the derivative transactions entered into by the corporation. In addition, all derivatives must be marked to market and the changes in their market value reported in the income statement. FAS 133 has been in effect since June 16, 2000.

Recognizing that marking derivatives to market introduces volatility to earnings (anathema to Wall Street's equity analysts), FAS 133 allows the corporation to mark the hedged liabilities or assets to market as well. But, in order to qualify for this favorable treatment, the hedge must first be shown to be highly effective. This article focuses on testing for hedge effectiveness. While FAS 133 distinguishes between "fair value" and "cash flow" hedges, we confine our discussion to fair value hedges for brevity.

After providing an overview of hedge effectiveness testing, we review and critique the FASB guidelines from both a conceptual and a technical perspective. The spirit of the guidelines is commendable, but we find the recommended tests to be flawed. In some cases they pass hedges that should fail, and in others they fail hedges even though such a conclusion is unwarranted.

The main part of the article presents and illustrates the application of an alternative test, one that is based on a volatility reduction measure that we call "VRM." We show this VRM-based test to be fully compliant with the spirit of the FASB's recommendations, while correcting their major shortfalls. We demonstrate the use of VRM for both prospective and retrospective testing, and then discuss how these measures should be combined into a single hedge effectiveness score.

While the VRM approach is applicable to hedges of most kinds of financial instruments and exposures, the illustrations in this article are taken from the realm of fixed income, such as the hedge of a fixed-coupon bond with a fixed-to-floating interest rate swap. Because of this focus, we also include an appendix that discusses the complexities of implementing hedge effectiveness testing for fixed income instruments. These complexities include the proper treatment of accrued interest and accounting for aging. With regard to prospective testing, we suggest a method for simulating future yield curves and discuss the number of simulations required for statistically significant results.

Finally, we explore the relationship of VRM to Value at Risk, or VaR, the risk measure that is used by financial institutions, as well as a growing number of industrial companies, in assessing and managing their exposures. We show that VRM provides a common analytical framework for accounting and risk management, and that it has direct applicability to optimizing hedges.

HEDGE EFFECTIVENESS TESTING

Having laid down the pre-requisite that a hedge must be shown to be highly effective, FASB has provided only broad guidelines as to how to actually test for effectiveness. Before reviewing these guidelines, it might be helpful to discuss the rationale for hedging from an economic, as opposed to an accounting, perspective.

The basic motivation for hedging is to reduce or eliminate the variability of profits and firm value that arises from market changes. In the real world of imperfect hedges, practitioners tend to be concerned only with deviations that are significant from a business perspective. What constitutes a "significant" change is typically specified by internal corporate guidelines.

The effectiveness of a hedge becomes relevant only in the event there is a significant change in the value of the hedged item, whether it be a specific collection of assets or an expected cash flow stream. In principle, a hedge is effective if the price movements of the hedged item and the hedging derivative roughly offset each other, so the net change of the *package* (hedged item plus derivative) is negligible relative to the change solely of the hedged item. If the changes are insignificant, the effectiveness of the hedge is irrelevant. We shall return to this observation below when we discuss the "80/125 rule" suggested by the FASB.

A well-designed test should pass a hedge that is truly effective, and fail one that is not. At first glance, a relatively lenient test—one that allows most of its derivative positions to qualify for hedge accounting treatment—would seem desirable to a corporation. But passing the test is only the initial step: the proof of the pudding is in the effect of the hedge on *reported earnings*. A test with a low power of discrimination—one that is easy to pass—can result in significant volatility in earnings.

A related problem may arise even for hedges that automatically qualify for the short-cut method described below, and therefore are not subject to testing. We will discuss why some of these qualifying hedges would actually fail a rigorously designed testing procedure. While the short-cut method eliminates quarter-to-quarter earnings volatility, these hedges may not eliminate business risk.

FASB Guidelines

According to the FASB, a properly designed effectiveness test should take into account both the historical performance (retrospective test) and the anticipated future performance (prospective test) of the hedge. With respect to historical performance, the FASB has suggested two approaches, which are discussed below. The FASB has not explicitly recommended methods for prospective testing, nor has it indicated how the results of retrospective and prospective tests should be combined. We will address each of these issues below.

The idea of combining historical performance with expected future performance has important practical ramifications. Consider a corporation that enters into a long-term, say 10-year, hedge. It is conceivable that six months into the hedge, a retrospective test will fail, due to some idiosyncrasy of the market. At the same time, prospective testing for the remaining 9 1/2 years may establish that the hedge is effective. By properly combining the results, one is likely to conclude that the hedge is effective, in spite of its poor historical performance. Continuing with the example, the same hedge may fail a prospective test with two years to go, even though it has performed well over the previous eight years. By including credit for previous good behavior, the effectiveness of the hedge may be justified.

According to our experience, most corporations perform only a retrospective test. However, as discussed above, this could severely understate the actual effectiveness of a hedge. In our opinion, for both theoretical and practical reasons, a well-designed test should be based on the *combination* of prospective and retrospective test results.

Let us examine the FASB's suggested approaches to retrospective testing. One is the so-called "80/125 rule." A hedge is deemed effective if the ratio of the change in value of the derivative to that of the hedged item falls between 80% and 125%.

An unintended and unfortunate consequence of the 80/125 rule is that during periods of market stability virtually any hedge is likely to fail, even though the resulting price movements are insignificant from a business perspective. Consider, for example, a $100 million bond hedged with an interest rate swap. A $10,000 change in the value of the bond and a $4,000 opposite change in the value of the swap results in a ratio 0.4. Hence, under the 80/125 rule, the hedge will be deemed *ineffective*

The spirit of the FAS guidelines is commendable, but we find the recommended tests to be flawed. In some cases they pass hedges that should fail, and in others they fail hedges even though such a conclusion is unwarranted. The VRM-based test is fully compliant with the spirit of the FASB's recommendations, while correcting their major shortfalls.

TABLE 1
VOLATILITY REDUCTION MEASURE (RETROSPECTIVE TEST)

	Change in Value ($ Millions)		
Quarter	**Swap**	**Bond**	**Package (Bond + Swap)**
1	(1.0)	1.1	0.1
2	(0.8)	1.0	0.2
3	(1.6)	2.0	0.4
4	2.5	(2.8)	(0.3)
5	2.6	(2.1)	0.5
Standard Deviation		1.921	0.332
Ratio of Standard Deviations 17.26%			
Volatility Reduction Measure 82.74%			

even though the net change of $6,000 is a miniscule 0.006% of the face amount.

We will remedy this weakness of the 80/125 rule by two means. First, we make it more robust by extending it backwards to incorporate *all* historical data, and not only the latest observations. In addition, for the reasons given above, we combine the result with that of the prospective test.[1]

The second approach suggested by the FASB is based on the correlation of the changes in value of the hedged item and that of the derivative. Roughly speaking, a hedge is deemed effective if the R-squared of the regression line explaining the data is sufficiently high, say 80%. But a high R-squared by itself is not a reliable indicator of effectiveness. In addition, the changes in value should be roughly offsetting; that is, the *slope* of the regression line should be close to 1.0, a consideration not explicitly referred to in FAS 133. (A related issue that is not addressed by the FASB is the intercept of the regression line. Should the intercept be constrained to 0 or should the "best fit" regression determine it? Constraints on the regression process lower the R-squared, increasing the likelihood that the hedge will fail the effectiveness test.)

As indicated, the FASB has provided only broad guidelines to effectiveness testing. Accountants, as examiners of the results, not implementers of tests, are understandably reluctant to pronounce *a priori* what would be an acceptable test. Instead, clients are often directed to management consultants. Meanwhile, a proliferation of software packages claiming "FAS 133 compliance" fail to address hedge effectiveness testing in any satisfactory manner. In the absence of definitive guidance, corporations are expected to devise, apply, and defend their own tests. Given this uncertain state of affairs, we propose the following approach.

The Volatility Reduction Measure (VRM)[2]

Traders and portfolio managers, whose compensation is affected by the actual performance of hedges, judge the effectiveness of a hedge in terms of *volatility reduction*. The volatility of the item being hedged in the *absence* of a hedge is the obvious point of reference against which this reduction should be measured. Recall that the FASB guidelines focus on *pair-wise* (date-by-date) comparison of changes in value, without explicit reference to overall volatility with and without the hedge. The VRM method described below captures the significance of hedging to practitioners while retaining the basic intent of the FASB.

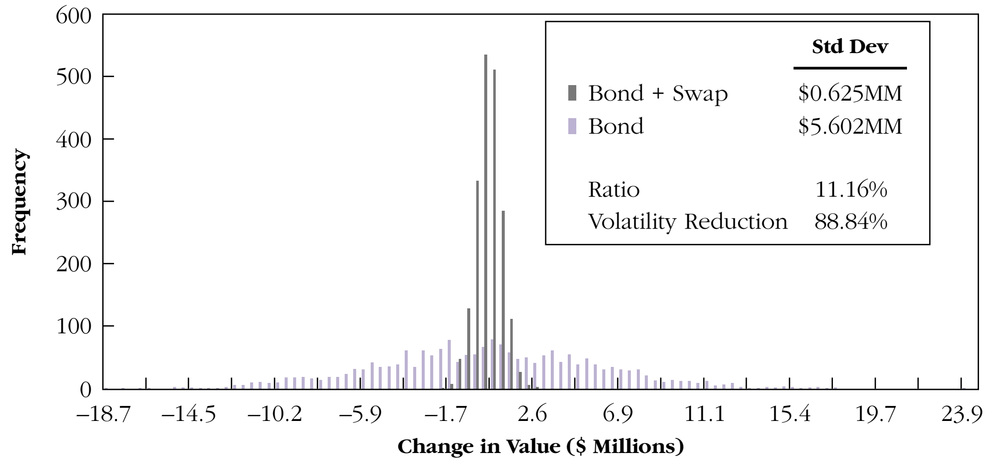
The following example of a retrospective test conveys the essence of the VRM approach. A fixed coupon bond[3] is being hedged with a LIBOR swap. The changes in value of each have been recorded for five quarters. Additionally, in each quarter we add the changes in the bond and the swap to arrive at a change in the value of the package (bond + swap). As shown in Table 1, the standard deviation (volatility) of the changes in value of the bond alone is $1.921 million, that of the package is $332 thousand.

1. The authors acknowledge helpful comments from Lisa Filomia and Jay Glacey of Ernst & Young with regard to combining retrospective and prospective tests.

2. The Volatility Reduction Measure (patent pending) for hedge effectiveness testing was invented by Andrew Kalotay Associates, Inc. It has been implemented at a leading telecommunications company and has been accepted by a Big-5 accounting firm. The authors gratefully acknowledge the significant contribution of Deane Yang to the development of VRM.

3. Under amended FAS 133 rules, the hedger is allowed to claim a hedge against changes in value of the bond attributable to changes in benchmark rates (in this case, LIBOR).

FIGURE 1
VOLATILITY REDUCTION MEASURE (PROSPECTIVE TEST): DISTRIBUTION OF 2000 SIMULATED QUARTERLY CHANGES IN VALUE ($100MM 10-YEAR 8% BOND HEDGED WITH 9.5-YEAR LIBOR SWAP RECEIVING 8.25%)



The addition of the swap has reduced the volatility to 17.26% (0.332MM/1.192MM) of what it originally was. In other words, a volatility reduction of 82.74% has been achieved.

The same idea is applied to a prospective test in Figure 1. The frequency distribution of 2,000 quarterly changes (simulated in the manner described in the Appendix) in the value of a $100 million 8% 10-year corporate bond is overlaid with analogous information for the *combination* of the bond and an 8.25% LIBOR-based 9.5-year swap of the same notional amount. The graph shows a volatility reduction from $5.602 million to $0.625 million.

More formally, the volatility reduction measure is defined as:

VRM = 1 – [stdev (hedge package) / stdev (item being hedged)]

According to the example in Figure 1, the prospective volatility of the hedged item is expected to be reduced by 88.84% (1 – 0.625MM / 5.602MM).

Prospective and retrospective results for a given hedge can be combined into a single measure by properly weighting the inputs into the VRM formula. In this way, the entire behavior of the hedge is encapsulated in a single result.

The VRM approach is similar to the idea of variance reduction introduced by Ederington in 1979 for assessing hedging performance.[4] As implied by its name, Ederington's method measures volatility reduction from a ratio of variances. Our preference for using standard deviations is motivated (apart from the fact they tend to be more meaningful to management) by our desire to be in accord with the 80/125 rule. A VRM result of 80% is equivalent to variance reduction of 96%, which can be misleading if one is focused on an 80% threshold. Or, to put it differently, a variance reduction of 80% is equivalent to a VRM of 55%. Clearly, in the spirit of the 80/125 rule, this last case has failed.

A clearer way of confirming that the dimensionality of VRM is in line with the intent of the 80/125 rule is to observe what happens when VRM is applied to a single set of changes in value. If the changes in the value of the bond and swap are Δ_i and Δ_d, respectively, then VRM (assuming standard deviation is calculated using the idealized mean of zero) is calculated as follows:

$$\begin{aligned} \mathbf{VRM} &= 1 - ((\Delta_i + \Delta_d)/\Delta_i) \\ &= 1 - (1 + (\Delta_d/\Delta_i)) \\ &= -(\Delta_d/\Delta_i) \\ &= \text{80/125 Result} \end{aligned}$$

VRM also has the added advantage of having a common analytical framework with Value at Risk (VaR), a widely used risk measure representing dollar exposure. VaR reporting is mandatory for financial institutions and may soon be required of

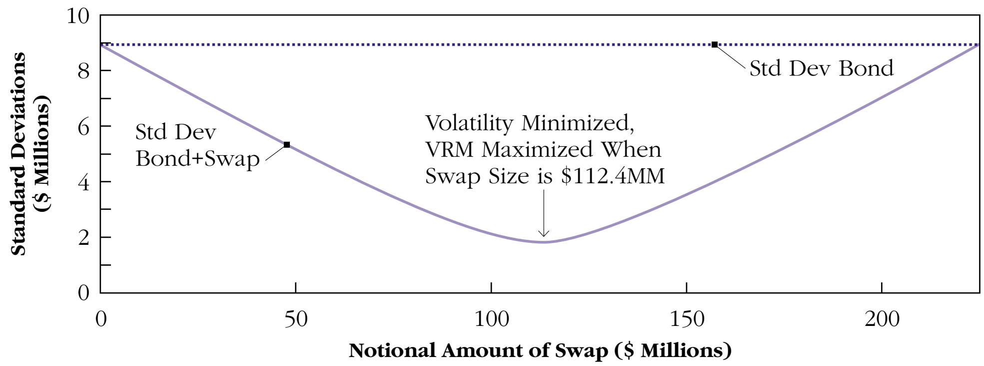
4. Louis H. Ederington, "The Hedging Performance of the New Futures Markets," *Journal of Finance* 34 (March 1979): 157-170. More recently, Kawaller and Koch mention variance reduction as a possible effectiveness test for FAS 133. (See Ira G. Kawaller and Paul D. Koch, "Meeting the 'Highly Effective Expectation' Criterion for Hedge Accounting," *Journal of Derivatives* 7 (Summer 2000)). For a review and analysis of various hedging effectiveness methods, see John D. Finnerty and Dwight Grant, "How to Test Hedge Effectiveness Under FAS 133" *Working Paper* (PricewaterhouseCoopers, LLP).

While the illustrations in this article are taken from the realm of fixed income, the VRM approach is applicable to hedges of most kinds of financial instruments and exposures. Moreover, VRM provides an analytical framework for accounting and risk management that is consistent with VaR's, and that has direct applicability to optimizing hedges.

TABLE 2
CHANGES IN VALUE OF $100 MILLION SWAP AND $100 MILLION BOND (IN $ MILLIONS)

Δ_d (Swap)	Δ_i (Bond)
5.0	(6.0)
2.0	(2.2)
(8.0)	7.0
9.0	(8.0)
11.0	(15.0)
(6.0)	8.0
σ_d (Swap): 7.782	σ_i (Bond): 8.934
ρ_{id} = –97.886%	

FIGURE 2
VOLATILITY REDUCTION OF $100MM BOND ISSUE



other corporations. VaR computations, like VRM, use standard deviation-based formulas that are applied to historical or simulated changes in the value of the corporation's positions.

OPTIMAL HEDGE RATIO AND VRM

As mentioned earlier, the R-squared test without reference to slope is not meaningful. Since R-squared (same as correlation squared in an unconstrained bivariate regression) is independent of scale, the size of the derivative does not affect it. In contrast, there is only *one size* that actually maximizes VRM.

Consider a hedged item and a derivative with known standard deviations and correlation. The standard deviation of the hedge package (item + derivative) is minimized, and therefore VRM is maximized[5] when the derivative position is scaled so that:

$$\sigma_d = -\rho_{id} . \sigma_i \qquad \text{(Equation 1)}$$

where σ_i, σ_d, and ρ_{id} are the standard deviations of the hedged item and the derivative, and their correlation, respectively.

The maximum volatility reduction[6] is given by:

$$\text{Max VRM} = 1 - (1 - \rho_{id}^2)^{1/2} \qquad \text{(Equation 2)}$$

As an illustration, say a $100 million bond issue is hedged with a $100 million swap. Based on six data points (for simplicity), we see in Table 2 that the standard deviation of the changes in value of the swap is $7.782 million; that of the bond is $8.934 million, and the correlation between them is –97.886%. Figure 2 demonstrates how changing the size of the swap affects the performance of the hedge. The maximum volatility reduction is achieved when the notional amount of the swap is $112.4 million ($-\rho_{id} . \sigma_i / \sigma_d = 0.97886 \times 8.934 / 7.782 = 1.124$ from Equation 1 above). This maximum is 79.545% (consistent with Equation 2 above).

5. To maximize VRM, find the weight of the derivative that minimizes the standard deviation of the hedge package (σ_p), given the standard deviations of the hedged item and the derivative and the correlation between them (σ_i, σ_d, and ρ_{id} respectively). Solve by differentiating $\sigma_p = (\sigma_i^2 + 2\rho_{id} + \sigma_d^2)^{1/2}$ with respect to σ_d and setting the result equal to 0.

6. Substitute $\sigma_d = -\rho_{id} . \sigma_i$ in the VRM formula:
$(1 - (\sigma_i^2 + 2\rho_{id} + \sigma_d^2)^{1/2} / \sigma_i)$.

This method of optimizing the hedge ratio is useful when originating a hedge. However, when attempting to qualify an existing hedge for hedge accounting treatment, the size of the swap is given and must be marked to market in full. On the other hand, it is possible to claim that *only a certain percent of the bond's face amount is being hedged.* Similar techniques can be employed to determine the size of the bond position that maximizes VRM. This could make the difference between the hedge passing and failing.

IN CLOSING

While the foregoing discussion and examples are in the realm of fixed income, the VRM test can be applied to hedges in general, be they currency or commodity related. The VRM test can also be directly applied to portfolio-based hedging (i.e., a portfolio of assets or liabilities being hedged with a portfolio of derivatives), even though FASB (inexplicably) does not accord favorable treatment to such at present.

The VRM approach is superior in its simplicity as well as its rigor and defensibility. Standard deviation is the accepted measure of volatility. When expressed in dollar terms, standard deviation reflects actual business risk, and unlike arcane statistics such as R-squared, it is familiar to higher management. Last but not least, the VRM method is consistent with Value-at-Risk (VaR), a widely used measure in risk management.

■ ANDREW KALOTAY

is founder and owner of Andrew Kalotay Associates, a New-York based firm specializing in risk management and capital markets advisory. Dr. Kalotay is a member of the Fixed Income Analysts' Society Hall of Fame.

■ LESLIE ABREO

is a senior financial analyst at Kalotay Associates. He has an M.S. in financial engineering and has co-authored papers on various aspects of corporate debt management.

APPENDIX ■ CONSIDERATIONS FOR FIXED INCOME SECURITIES

Determining the inputs into the VRM process is a challenge in itself when it comes to fixed income securities because their values must be derived from appropriate yield curves. In retrospective tests, the yield curves prevailing at the end of every quarter that the hedge has been in existence provide a reasonable starting point. But before calculating the changes in value of a bond, one must recognize that a component of the change is attributable to the aging of the bond: its price approaches par as it matures, regardless of interest rates.

This problem comes out in striking relief when a bond with an above-market coupon is hedged with an at-market interest rate swap. Clearly, the swap has an initial and terminal value of zero. In other words, the cumulative changes in value add up to zero over its lifetime. The bond, however, has a starting value of, say 114 (percent of face) and an ending value at maturity of 100. This hedge is going to appear highly ineffective unless the aging component of the change in value of the bond is isolated away from the analysis.

The theoretically rigorous way to accomplish this is to calculate the change in value at the end of the current quarter relative to what this value was projected to be (using forward curves) three months earlier.

A second element of contention is the treatment of accrued interest. In fair value hedges of fixed coupon bonds with fixed for floating swaps, the difference in accrued interest can be a significant component of the value of the swap. However, accrued interest is already being passed to earnings as interest expense. Since the FASB's intent is to expose volatility that was previously hidden, our view is that accrued interest should be excluded from valuation and testing for FAS 133 purposes.

We should recognize that hedges of bonds with interest rate swaps can be imperfect for a couple of reasons, even if all the payment dates line up well. One source of potential trouble is the difference between the bond coupon and the effective fixed rate on the swap,[1] which causes a duration mismatch. A second problem occurs in the terminal coupon payment period. By that time, the swap's floating rate is fixed, causing the bond and the swap to move in the same direction with changes in interest rates. Thus even hedges that qualify for the short-cut method arguably fail for one or both of these reasons.

Simulation of Yield Curves

The simulation of interest rates for prospective testing is an area that also deserves some attention. Given the limited availability of data, one must be careful to avoid the autocorrelation effects of using historical rate information.

Our recommended solution to this problem is as follows:

- From the available set of historical yield curves, calculate daily term-by-term changes in rates as "vectors" of ratios. For example, from Day 1 to Day 2, the 6-month rate has moved from 5.5% to 5.7%, …, the 10-year rate has moved from 5.8% to 5.9%, … , and the 30-year rate has moved from 5.3% to 5.4%. Express this change as:

[5.7/5.5, … , 5.9/5.8, … , 5.4/5.3]

- Repeat this process for consecutive days until the available historical yield curves are exhausted. Note that each ratio vector represents a daily yield curve transformation, i.e., it describes how each point on the yield curve (6 months out to 30 years) changed on a given day.
- Randomly select 62 of these ratio vectors and multiply them term-by-term to produce one quarterly yield curve transformation vector (a quarter is assumed to contain around 62 trading days).
- Repeat the previous step until an adequate number of quarterly yield curve transformations are obtained. We believe 200 to be a sufficient number to avoid sample size inadequacy.
- Multiply each of the quarterly yield curve transformations by the curve prevailing on the date of the test to produce the Monte Carlo set of curves for prospective valuation.

1. The effective fixed rate on a swap that pays say 8% and receives LIBOR + 65 basis points is 7.35% if the spread of 65 basis points is netted out on both sides of the swap.