DwightGrantConsulting 7332 Eads Ave. La Jolla CA 92037 415-509-3943

Dwight Grant, PhD
dwight@dwightgrantconsulting.com

Alternative Approaches to Testing Hedge Effectiveness Under SFAS 133

by

John Finnerty and Dwight Grant

Accounting Horizons 16 (June 2002), 94 – 108

ALTERNATIVE APPROACHES TO TESTING HEDGE EFFECTIVENESS UNDER SFAS 133

I. Introduction

Statement of Financial Accounting Standards No. 133 (SFAS 133), *Accounting for Derivative Instruments and Hedging Activities*, as amended by Statement of Financial Accounting Standards No. 138 (SFAS 138), *Accounting for Certain Derivative Instruments and Certain Hedging Activities – An Amendment of FASB Statement No. 133,* prescribes comprehensive new rules for accounting for derivative instruments. SFAS 133 standardizes the accounting treatment for derivative instruments by requiring all entities to report their derivatives as assets and liabilities on the balance sheet and to measure them at fair value (paragraph 3).

Reporting changes in the fair value of a derivative in earnings each quarter could create a matching problem. If the derivative is being used as an economic hedge, changes in the value of the derivative might increase (or decrease) reported earnings one period while the opposite change in the value of the hedged item affects earnings in a later period. To avoid distorting earnings, SFAS 133 permits firms to match the timing of the gains and losses of hedged items and their hedging derivatives, provided the derivative qualifies as a “highly effective” hedge (paragraphs 3, 20, and 28). For a fair value hedge, SFAS 133 permits the hedger to record the change in the fair value of the hedged item concurrently with the gain or loss on the hedging derivative (paragraph 22). In the case of a cash flow hedge, the effective portion of any changes in the hedging derivative’s fair value is recorded in other comprehensive income until the change in the value of the hedged item is recognized in earnings (paragraph 30).[1]

[1] SFAS 133 defines three classes of hedges, fair-value hedges, cash-flow hedges, and foreign-currency hedges. This paper does not address how to test the effectiveness of hedges of net investments in foreign operations.

In principle, a hedge is highly effective if the changes in fair value or cash flow of the hedged item and the hedging derivative offset each other to a significant extent. *Hedged item* refers to an asset or a liability or a prospective cash inflow or outflow. *Derivative* refers to any derivative or combination of derivatives used to hedge changes in fair value or cash flow. *Hedged position* refers to the combined hedged item and derivative. The hedged item can be a designated portion of an asset or liability or a designated expected future cash flow that is attributable to a particular risk (SFAS 133, paragraphs 21 and 28, and Derivatives Implementation Group (DIG) Issue E10). To qualify a derivative position for hedge accounting, the hedging entity must specify the hedged item, identify the hedging strategy and the derivative, and document by statistical or other means the basis for expecting the hedge to be 'highly effective' in offsetting the designated risk exposure. The documentation step is called *prospective testing*, and it must be done before entering into the hedge and on an ongoing basis to justify continuing hedge accounting. The hedger must also regularly perform *retrospective testing* to determine how effective the hedging relationship has been in actually achieving offsetting fair values or cash flows.

This article examines the test choices that firms must make. Section 2 of Appendix A of SFAS 133 requires the use of statistical or other numerical tests for hedge effectiveness, unless a specific exception applies, but SFAS 133 does not endorse any specific testing methodology. The hedger must select the methodology, such as regression analysis, choose the measurement period, and specify an appropriate test statistic like adjusted R^2 along with the critical value to distinguish a 'highly effective' hedge from one that is not; for example, an adjusted R^2 above 0.8.

The article describes four existing methods of testing hedge effectiveness: (1) the Dollar-Offset Method, (2) the Relative-Difference Method, (3) the Variability-Reduction Method, and

(4) the Regression Method. We refine the Variability-Reduction Method and the Regression Method to correct their potential to accept poorly performing hedges. Our exposition identifies challenges that firms may confront and choices that firms must make when testing hedge effectiveness. These descriptions and illustrations reflect our best understanding of this subject. This paper does not make any representation as to the acceptability of any of these tests by a hedger's auditors. We illustrate each of the methods with a hedging example.

II. Methods of Testing Hedge Effectiveness

A hedge is "highly effective" only when the change in the fair value of the derivative substantially offsets the change in the fair value of the hedged item or cash flows attributable to the risk being hedged. The seminal papers on measuring economic effectiveness of hedges include Ederington (1979) and Franckle (1980). An entity need not perform a detailed analysis of hedge effectiveness when the critical terms of the hedging derivative and the hedged item are the same because the hedger can reasonably expect complete offset of the risk being hedged (SFAS 133, paragraph 65, and DIG Issue G9).[2] We discuss methods of testing the effectiveness of forwards, futures and swaps as hedges when the critical terms of the hedging derivative and the hedged item are not identical. We do not evaluate the hedge effectiveness of options.[3]

The retrospective testing and the update of the prospective testing should be performed at least each quarter or each time a financial statement or earnings are reported until the hedge is unwound (SFAS 133, paragraph 20, and DIG Issue E7). Data used in retrospective testing must include the actual results since the inception of the hedge and may include additional historical data. Tests used to document hedge effectiveness must be consistent with the hedger's stated

[2] There is also a 'shortcut method' that is available for interest-rate swaps and recognized interest-bearing assets or liabilities. See SFAS 133, paragraphs 68-70, 114, and 132, and DIG Issues E4, E10, E12, E14, E15, and E16.
[3] Testing option hedge effectiveness is more complex because the hedge can be either dynamic or static and you can test using the option's market value, intrinsic value or minimum value (SFAS 133, paragraph 63).

approach to risk management, and the hedger must use the same method to test the effectiveness of similar hedges unless different methods are explicitly justified.[4] In addition, SFAS 133 requires the hedger to recognize in current earnings any ineffective portion of the hedge (paragraphs 22 and 30). Thus, even when a hedge is determined to be highly effective, there is an impact on current earnings when there is not an exact offset of the hedged risk.[5]

Defining and testing a measure of hedge effectiveness are important and potentially challenging aspects of hedge accounting. Failure to execute these aspects well may introduce substantial volatility in reported earnings. Several statistical tests are available for testing hedge effectiveness (Kawaller and Koch, 2000). Because SFAS 133 does not specify a bright line test to distinguish highly effective hedges from less effective or ineffective hedges, the interpretation of "highly effective" is a matter of judgment. The high-effectiveness requirement is intended to have the same meaning as the 'high correlation' requirement of SFAS 80 (SFAS 133, paragraph 389), which has been interpreted to mean either that the cumulative changes in the hedging derivative should offset between 80 percent and 125 percent of the cumulative changes in the fair value or cash flows of the hedged item (Swad, 1995) or that the regression of changes in the hedged item on changes in the derivative should have an adjusted R^2 of at least 80 percent (Lipe, 1996).

The firm must choose the frequency of data observation and the time span of observation to use in the effectiveness testing. SFAS 133 and SFAS 138 provide flexibility in both choices. When discussing risk-management issues, some authors argue that the testing interval should

[4] The hedger must specify the method of retrospective testing and the length of the testing period, as well as perform the initial prospective testing, before implementing the hedge (SFAS 133, paragraph 62). However, SFAS 133 does permit the hedger to use different methods for prospective and retrospective testing provided it documents the choices before implementing the hedge.
[5] The amount of hedge ineffectiveness for accounting recognition is determined by directly comparing the difference between changes in fair values of the hedging derivative and the hedged item in the case of a fair-value hedge and by directly comparing the change in the fair value of the hedging derivative and the change in present value of the cash flows of the hedged item in the

match the hedge's time horizon (Ederington, 1979, and Figlewski, 1984). For example, use annual data to evaluate whether a particular derivative is an effective hedge of a 12-month exposure. However, matching exposure and measurement periods may limit the number of independent observations available for statistical testing. For this example, obtaining 30 independent data points to test a 12-month hedge requires 30 years of historical data.[6] This may not be feasible given available data, or it may not be appropriate because the market changed substantially over that period of time. Our compromise is to use more frequent observations of the data over a shorter historical time period. Therefore, to illustrate the methods of testing hedge effectiveness with a 12-month exposure, we use monthly changes in the values of the hedged item and the derivative.[7]

To illustrate the methods described here, we adopt the perspective of a U. S. company that is considering, on December 31, 1999, hedging its purchase of aluminum that is expected to occur on December 31, 2000.[8] The derivative is a long forward contract on the London Metals Exchange (LME).[9] On December 31, 1999, the cash price for delivery to the company's plant is $1,712.30 per metric ton, compared to the cash and one-year forward prices on the LME of $1,630.50 and $1,641.00, respectively. For the previous two years, on average the company's cost of delivered aluminum was 7.2% higher than the LME cash price, and the LME forward price was at a 4.0% annualized premium to the LME cash price. The latter relationship is important because the hedger can choose to exclude the forward premium or discount from the

case of a cash-flow hedge and determining the extent to which exact offset is not achieved (SFAS 133, paragraphs 22, 30, and 63, and DIG Issue E7.)

[6] SFAS 133 allows the use of simulated data but provides little guidance for setting up the simulation. This paper uses historical data and leaves the discussion of simulation to future research.

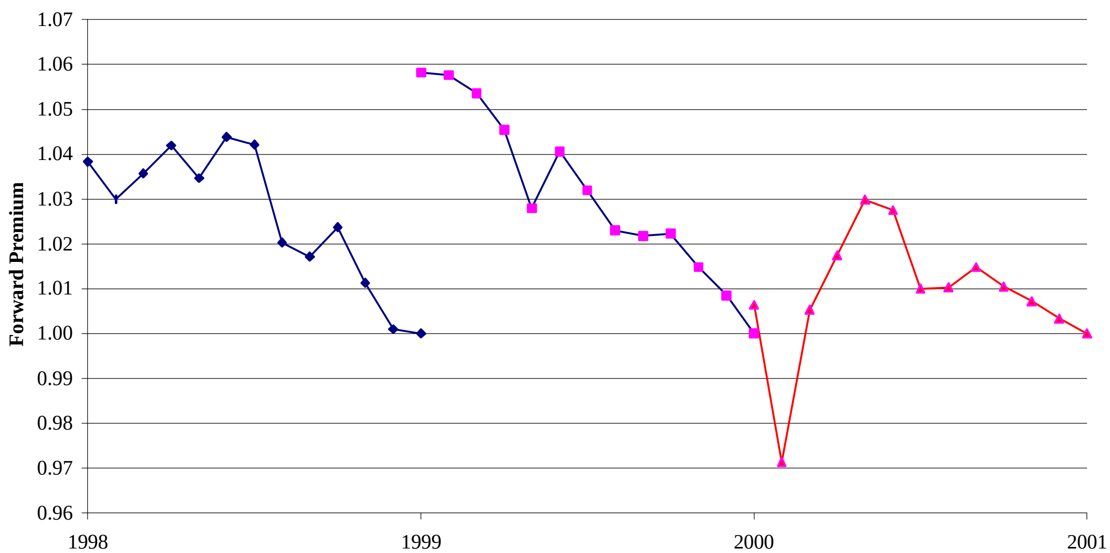
[7] Whether the hedging effectiveness test results are sensitive to the measurement interval is a complex issue. The answer may vary depending on the characteristics of the data, including whether the price series exhibit drift or autocorrelation, neither of which is present in the data used in this paper.

[8] All of the illustrations assume static hedging strategies. SFAS 133 also permits hedge accounting for dynamic hedging strategies that are properly specified, documented, and tested (paragraphs 85-87, and DIG Issue E11).

[9] The New York Mercantile Exchange did not begin to trade aluminum futures contracts until mid-1999.

calculation of hedge effectiveness. The forward premium is also important because variations in it may contribute to hedge ineffectiveness. Figure 1, which plots the dollar value of the forward premium for the years 1998 – 2000, indicates that there is a relatively low forward premium at the beginning of 2000 and a relatively high level of variation in the forward premium in the first quarter of 2000.

Figure 1
London Metals Exchange Aluminum Forward Premium:
Monthly Data 1998 - 2000[a]



[a] The figure plots the ratio of the forward price to the spot price at monthly intervals. The data display convergence of the spot and forward prices at year's end and illustrate the volatility of the premium, especially in the first half of 2000.

To determine whether the one-year forward contact is expected to provide a highly effective hedge, the company collected 1998 and 1999 monthly cash prices for delivery of aluminum to its plant and forward prices on the LME. The forward prices are the prices for delivery at the end of each year, respectively.[10] The prices do not exhibit seasonality.

[10] The London Metals Exchange provides daily quotations for the cash price and for 3-month and 15-month forward contracts. To calculate the prices for a forward contract for delivery at the end of the year, we use linear interpolation between the two prices that fall on either side of the delivery time.

All methods of prospectively testing hedge effectiveness require the firm to decide how many past periods to use in assessing whether a hedge will likely be highly effective. If the firm uses data to calculate a test statistic for only one prior period, the hedge either passes or fails the test. If the firm uses data to calculate test statistics for more than one prior period, there are two possible definitions of highly effective. The more stringent standard is that the hedge must satisfy the test in every period. The less stringent requirement is that the hedge must satisfy the test in a high proportion of the periods, perhaps 80% or 90%.[11] For retrospective purposes, the test can be based on a comparison of either (a) the fair value or cash flow changes that occurred during the period being assessed or (b) the cumulative fair value or cash flow changes to date since the inception of the hedge (DIG Issue E8). In either form, the hedge passes the test only if the ratio is within the critical range.

Our hedging example illustrates the tests of effectiveness for each method on a quarterly basis. For each quarter, we present the test results using (1) the data for just that quarter and (2) the data for the preceding 12 months. When testing hedge effectiveness, a firm can include or exclude the premium or discount on the derivative, which is the difference between the London Metals Exchange one-year forward price and the London Metals Exchange cash price. Therefore, we performed the calculations both with and without the premium or discount.

1. Dollar-Offset Method

The Dollar-Offset Method, which has some historical significance for the accounting profession (Kawaller and Koch, 2000, and DIG Issue E7), compares the changes in the fair value or cash flow of the hedged item and the derivative. The Dollar-Offset Method can be applied either period-by-period or cumulatively (DIG Issue E8). For a perfect hedge, the change in the

[11] To our knowledge the accounting profession has not adopted this less stringent standard.

value of the derivative exactly offsets the change in the value of the hedged item, and the negative of their ratio is 1.00. In its cumulative form this means

$$-\left(\sum_{i=1}^{n} X_i \Big/ \sum_{i=1}^{n} Y_i \right) = 1.0\,, \tag{1}$$

where $\sum_{i=1}^{n} X_i$ is the cumulative sum of the periodic changes in the value of the derivative and $\sum_{i=1}^{n} Y_i$ is the cumulative sum of the periodic changes in the value of the hedged item. The minus sign in front of the ratio adjusts for the two sums being opposite in sign in a hedging relationship.

Of course, perfection is not necessary to qualify for hedge accounting. In a speech at the SEC's 1995 Annual Accounting Conference, a member of the SEC's Office of the Chief Accountant articulated an 80/125 standard for hedge effectiveness (Swad, 1995). This became a guideline for assessing the hedge effectiveness of futures contracts under SFAS 80, and it carries over to testing the effectiveness of hedges under SFAS 133. In the context of the Dollar-Offset Method, the derivative's change in value should offset at least 80% and not more than 125% of the value change of the hedged item. The logic underlying 80/125 is that the standard is independent of the arbitrary choice of numerator and denominator because 80% = 4/5 and 125% = 5/4. The formal expression of the test is

$$0.8 \le -\left(\sum_{i=1}^{n} X_i \Big/ \sum_{i=1}^{n} Y_i \right) \;\le\; 1.25\,. \tag{2}$$

Anyone choosing this test should be aware that researchers question its reliability (Canabarro, 1999, Kawaller and Koch, 2000, and Althoff and Finnerty, 2001).[12] We emphasize that the test's result is very sensitive to small changes in the value of the hedged item or the

[12] A study by Canabarro (1999) finds that the Dollar-Offset Method rejects hedge accounting for a large number of hedges even when changes in the fair value or cash flow of the derivative and the hedged item are highly correlated.

derivative. For example, suppose that the hedged item is inventory valued at $1,000,000 and the hedge is a short position in a futures contract. At the end of the quarter, suppose that the value of the inventory increased by some small amount, say $10,000, or 1%. The short futures position will decrease in value by $10,000, offsetting the change in the value of the underlying asset, plus or minus the change in the futures' basis. If the change in the basis is as little as 0.33% of the notional value (+$3,333 or –$3,333), then the Dollar-Offset Method implies that the hedge is ineffective because the short futures' value change is either 33% greater or 33% less than the inventory's value change.

Table 1 reports the four sets of Dollar-Offset test ratios. The results for 1998 and 1999 are prospective tests and those for 2000 are retrospective tests. We highlight the ratios that fall outside the 80/125 critical range. When we use the price changes for one quarter and include the forward premium, the hedge fails the test in five of the 12 quarters. Excluding the forward premium, the hedge again fails the test in five of the 12 quarters. When we base the hedge effectiveness test on the preceding 12 months of price changes and include the forward premium, five of the nine ratios[13] fall outside the acceptable range. But when we use 12 months of data and exclude the forward premium, only one of the nine ratios falls outside the acceptable range.

We noted earlier that when the changes in the values of either the hedged item or the derivative are close to zero, small dollar amounts of ineffectiveness can produce extreme ratios. To check whether these indications of ineffectiveness are attributable to that phenomenon, we rank the observations for each of the four cases, from smallest to largest, by the absolute value of the change in value of the hedged item. We number the observations one to 12 or one to nine as appropriate. When we use the price changes for one quarter and include the forward premium,

[13] We have only nine ratios because we did not use the 1997 data necessary to measure the 12-month ratios for each of the first three quarters of 1998.

the ranks of the five cases of ineffectiveness are 1, 2, 3, 4, and 5. Excluding the forward premium, the ranks of the five cases of ineffectiveness are 1, 2, 4, 5, and 8. When we base the hedge effectiveness test on the preceding 12 months of price changes and include the forward premium, the ranks of the five cases of ineffectiveness are 1, 2, 3, 4, and 9. Lastly, when we use 12 months of data and exclude the forward premium, the rank of the one case of ineffectiveness is 2. The preponderance of low rankings among the tests indicating ineffectiveness suggests that the indication of ineffectiveness is often an artifact of this ratio measure.

Table 1
Measuring the Hedging Effectiveness of an Aluminum Forward Hedge
Using the Dollar-Offset Method, 1998 – 2000[a]

Year	1998	1998	1998	1998	1999	1999	1999	1999	2000	2000	2000	2000
Quarter	1	2	3	4	1	2	3	4	1	2	3	4
Current Quarter's Data												
Including forward premium	0.85	1.05	**0.32**	1.20	**4.32**	0.92	0.95	0.85	1.12	**0.56**	**2.64**	**0.77**
Excluding forward premium	**0.71**	0.95	0.89	1.04	**2.57**	1.03	1.12	0.99	1.09	**0.60**	**3.08**	**0.71**
Preceding 12 Months' Data												
Including forward premium				1.09	**1.36**	**0.00**	**0.57**	0.82	0.85	**0.73**	**0.62**	1.19
Excluding forward premium				0.90	1.07	**0.79**	0.99	1.00	1.03	0.92	0.84	1.04

[a] The test statistic for the Dollar-Offset Method is $-\left(\sum X_i / \sum Y_i\right)$, the ratio of the sum of the changes in the value of the derivative to the sum of the changes in the value of the hedged item. For a highly effective hedge, $0.8 \leq -\left(\sum X_i / \sum Y_i\right) \leq 1.25$. Test values that fall outside the critical range are highlighted in bold.

One of these cases highlights the small-number problem. The aggregate change in price of a short forward position for the last six months of 1998 and first six months of 1999 is –$0.03. Given that price change, the hedge fails the Dollar-Offset test for any change in the spot price, measured to penny accuracy, other than $0.03. These observations caution us against relying exclusively on the Dollar-Offset Method because of its sensitivity to small changes in value.

2. Relative-Difference Method

Because the Dollar-Offset Method may produce a false negative by incorrectly signaling ineffectiveness for small changes in value, we may be able to modify the Dollar-Offset Method to avoid this problem. One approach is to reduce the influence of small dollar changes in value by using percentage changes to test for effectiveness (Kawaller and Koch, 2000). Known as the Relative-Difference Method, this approach defines the measure of hedge effectiveness as

$$RD_n = \frac{\sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i}{V_0}, \tag{3}$$

where $\sum_{i=1}^{n} X_i$ and $\sum_{i=1}^{n} Y_i$ have the same meaning as in equation (1), and the combined sum is expressed as a proportion of the initial value of the hedged item, V_0. For a perfect hedge, RD_n equals zero. A hedge can be accepted as highly effective if RD_n is sufficiently close to zero. But before implementing the hedge, a firm must set a "sufficiently close" critical value. Of course, its auditors must concur with the method of testing and the choice of critical value. To illustrate the method, we assume a critical value of 3%, meaning that a hedge is effective if $-3\% \le RD_n \le +3\%$. This is an arbitrary choice; we are not aware of any statistical test based on RD_n that has a critical range consistent with the Dollar-Offset Method's 80/125 standard.

Review of the earlier example that violates the Dollar-Offset Method illustrates the potential merit of the Relative-Difference Method. In the earlier example, the hedged item is inventory valued at \$1,000,000. The inventory increases by \$10,000 in a quarter and the futures hedge decreases by \$13,333. The Dollar-Offset Method signals ineffectiveness because \$13,333/\$10,000 is greater than 1.25. The Relative-Difference Method does not signal ineffectiveness because (–$13,333 + $10,000)/$1,000,000 is only –0.33%, a relatively small deviation from a perfect offset and far less than the critical value –3%.

Table 2 presents test results similar to those in Table 1. The 1998 and 1999 results represent prospective tests and the 2000 results represent retrospective tests. We measure RD_n using three months of data for each of the 12 quarters, and we also measure the 12-month cumulative value of RD_n for each quarter beginning with the last quarter of 1998. We calculate each of these measures with and without the forward premium. None of the 24 quarterly absolute values of RD_n exceeds the critical value of 3%. Likewise, when we exclude the forward premium, none of the 12-month measures exceeds the critical value. However, when we include the forward premium, the 12-month relative differences are relatively large with six of the nine values exceeding the critical value of 3%. The ineffectiveness measured is attributable to the substantial changes in the forward premium during this period.

Table 2
Measuring the Hedging Effectiveness of an Aluminum Forward Hedge
Using the Relative-Difference Method, 1998 – 2000[a]

Year	1998	1998	1998	1998	1999	1999	1999	1999	2000	2000	2000	2000
Quarter	1	2	3	4	1	2	3	4	1	2	3	4
Current Quarter's Data												
Including forward premium	1.0%	-0.5%	-1.1%	-1.2%	-2.3%	-0.9%	-0.3%	-1.3%	-0.6%	-1.6%	0.6%	0.7%
Excluding forward premium	1.9%	0.5%	-0.2%	-0.2%	-1.1%	0.3%	0.9%	-0.1%	-0.5%	-1.4%	0.8%	0.9%
Preceding 12 Months' Data												
Including forward premium				-1.8%	**-5.1%**	**-5.5%**	**-4.6%**	**-4.8%**	**-3.2%**	**-3.9%**	-2.9%	-0.9%
Excluding forward premium				2.0%	-1.0%	-1.2%	-0.1%	0.0%	0.6%	-1.2%	-1.2%	-0.2%

[a] The test statistic for the Relative-Difference Method is $RD_n = \left(\sum X_i + \sum Y_i\right) / V_0$, the combined sum of the changes in the value of the derivative and the hedged item divided by the initial value of the hedged item. In the table, a hedging relationship is treated as highly effective if $-3\% \le RD_n \le +3\%$. Test values that fall outside the critical range are highlighted in bold.

If the Relative-Difference Method is defined in the case of forwards or futures hedges to exclude the forward premium, it may hold promise as a means of overcoming the problems the Dollar-Offset Method has with producing false negative signals for small changes in values. Its disadvantages are that it is relatively complex to implement successfully because of the need to adjust for the forward premium and the lack of a well-established critical value.

3. Variability-Reduction Method

For a perfect hedge, the change in the value of the derivative exactly offsets the change in the value of the hedged item, $X_i + Y_i = 0.0$. The Variability-Reduction Method compares the variability of the fair value or cash flow of the hedged (combined) position to the variability of the fair value or cash flow of the hedged item alone. This method places greater weight on larger deviations than smaller ones by using the squared changes in value to measure ineffectiveness. Other authors (Kalotay and Abreo, 2001, and Althoff and Finnerty, 2001) suggest a ratio of standard deviations or variances. The test statistic we advocate for this method is the proportion of the hedged item's mean-squared deviation[14] from zero that the hedge eliminates:

$$VR = 1 - \frac{\sum_{i=1}^{n} \left(X_i + Y_i \right)^2}{\sum_{i=1}^{n} Y_i^2} \tag{4}$$

We use the mean-squared deviation from zero because the variance ignores certain types of ineffectiveness. For example, suppose that the change in the value of the hedged position is always –\$0.20, $D_i = X_i + Y_i = -\$0.20$. If we use the variance of D_i in the numerator, the test statistic is 1.0 because the variance measures the variability around the mean of –\$0.20.

[14] Mean-squared deviation from zero is often used as a measure of forecast error when the target error is zero. The objective of the hedge is to eliminate all variability in the value of the hedged item. Therefore, the target for $X_i + Y_i$ and its mean is 0.0, and any deviation from 0.0 represents ineffectiveness in the hedge, which the test of hedge effectiveness should detect.

However, since D_i = –\$0.20 in every period in this example, the offset is not perfect. By using mean-squared deviations, the test statistic reflects the lack of offset in the means.

The critical value for determining how large a reduction in variability is sufficient to demonstrate hedge effectiveness must be specified in order for this measure to be useful. Studies of futures hedging in the risk-management literature, where the hedged item and the derivative are very similar assets, document variance reduction of 80% to 100% (Ederington, 1979, Figlewski, 1984, and Malliaris and Urrutia, 1991). This paper uses 80% as the minimum acceptable reduction in the mean-squared deviation for a hedge to be accepted as highly effective.

Before reporting our results, we should comment on how this 80% critical value relates to the 80/125 standard of the Dollar-Offset Method. The simple case in which $X_i = -1.26Y_i$ violates the 80/125 standard of the Dollar-Offset Method, but the variability reduction is 93% ($D_i = X_i + Y_i = -0.26Y_i$ and $VR = 1 - (0.26)^2 = 0.93$). This value is well above our critical value of 80%, and seems to indicate a highly effective hedge. Such an illustration indicates that the critical value standards articulated for each method may not provide consistent results. Specifically, a minimum reduction of 80% in the mean-squared deviation from zero may be a less demanding test of effectiveness than requiring an 80/125 dollar offset. A thorough analysis of this important issue is beyond the scope of this paper.

Table 3 reports the results for the Variability-Reduction Method. The results with and without the forward premium are essentially identical. When we use the price changes for one quarter, the hedge fails the test in three of the 12 quarters, including the first quarter of 2000, when the forward premium was volatile. But when we base the hedge effectiveness test on the preceding 12 months of price changes, all observations are above the critical value, ranging from 81% to 97%.

Table 3
Measuring the Hedging Effectiveness of an Aluminum Forward Hedge Using the Variability-Reduction Method, 1998 – 2000[a]

Year	1998	1998	1998	1998	1999	1999	1999	1999	2000	2000	2000	2000
Quarter	1	2	3	4	1	2	3	4	1	2	3	4
Current Quarter's Data												
Including forward premium	92%	95%	**72%**	95%	**34%**	98%	96%	99%	**65%**	93%	83%	99%
Excluding forward premium	89%	95%	**75%**	98%	**50%**	98%	95%	100%	**65%**	93%	81%	99%
Preceding 12 Months' Data												
Including forward premium				91%	86%	94%	96%	96%	88%	82%	81%	83%
Excluding forward premium				91%	88%	95%	96%	97%	88%	83%	82%	83%

[a] The test statistic for the Variability-Reduction Method is $VR = 1 - \left(\sum \left(X_i + Y_i \right)^2 \Big/ \sum Y_i^2 \right)$, the proportion of the mean-squared deviations from zero of the hedged item that is eliminated by the hedge. For a highly effective hedge, $VR \geq 80\%$. Test values that fall outside the critical range are highlighted in bold.

4. Regression Method

The Regression Method uses regression analysis to identify the size of hedge to implement and to test hedge effectiveness (see Althoff and Finnerty, 2001, and Royall, 2001). Applying the principles of modern portfolio theory, Ederington (1979) formulates the hedger's problem and derives a formula for the hedge ratio that minimizes the variance of the price changes of the hedged position. Ederington shows that the variance-minimizing hedge ratio is the estimated slope coefficient of a regression in which the change in the value of the hedged item is the dependent variable and the change in the value of the derivative is the independent variable,

$$Y_i = \hat{a} + \hat{b}(-X_i) + e_i , \tag{5}$$

where $\hat{a}$ is the estimated intercept term, $\hat{b}$ is the estimated slope coefficient, and e_i is the error term. Given our definitions of *X* and *Y*, the slope of this regression equation should be positive and close to 1.0.

Three characteristics of the regression indicate the prospective effectiveness of a hedge: the intercept coefficient, the regression slope coefficient, and the adjusted R^2. When the intercept is 0.0, and the slope coefficient and the adjusted R^2 are both 1.0, the hedge is perfect. Therefore, there is prospective support for the effectiveness of the hedge to the extent that the intercept estimate is "close" to 0.0 and the slope coefficient and the adjusted R^2 are both "close" to 1.0. We can make the definition of "close" operational by calculating a prospective Regression Method measure of variability reduction, *RVR*. To do this, we calculate the mean-squared deviation that results when we implement the hedge based on the regression slope coefficient, $\hat{b}$. Because the deviation of the hedged position from the target of 0.0 is $\hat{b}X_i + Y_i$, the measure of hedge effectiveness is

$$RVR = 1 - \frac{\sum_{i=1}^{n}\left(\hat{b}X_i + Y_i\right)^2}{\sum_{i=1}^{n} Y_i^2} \tag{6}$$

RVR is a more reliable measure of hedge effectiveness than adjusted R^2 because it detects the effects of departures of the intercept from zero and the slope from one.

Because the Regression Method is similar to the Variability-Reduction Method, we again use a critical value of 80%. If $RVR \geq 0.80$, the aggregate ineffectiveness introduced by differences in the respective mean changes in values (the intercept), and the deviations of the correlation and the slope from 1.0 are sufficiently small for the hedge to be judged highly effective. The retrospective Regression Method test of effectiveness is also defined by equation (6), except that the retrospective test uses the actual hedge ratio the hedger implemented. In our example, we estimated the hedge ratio based on data for months 1 – 24, found it to be 1.0877, and implemented a hedge of that size for the year 2000.

Table 4 reports the results of the prospective Regression Method tests of hedge effectiveness based on 1998 and 1999 data and the results of the retrospective tests for 2000. When we use the price changes for one quarter and include the forward premium, the hedge fails the test in four of the 12 quarters. One of those failures is in the first quarter 2000, when the forward premium was very volatile. Excluding the forward premium, the hedge fails the test in only one of the 12 quarters, the first quarter of 2000. When we base the hedge effectiveness test on the preceding 12 months of price changes, none of the hedges fails the test, and the proportion of the variability eliminated by the hedge ranges from 84% to 99%. When we exclude the forward premium in this 12-month scenario, the proportion of the variability eliminated by the hedge ranges from 89% to 99%.

Table 4
Measuring the Hedging Effectiveness of an Aluminum Forward Hedge Using the Regression Method, 1998 – 2000[a]

Year	1998	1998	1998	1998	1999	1999	1999	1999	2000	2000	2000	2000
Quarter	1	2	3	4	1	2	3	4	1	2	3	4
Current Quarter's Data												
Including forward premium	94%	96%	**76%**	92%	**20%**	99%	94%	99%	**69%**	96%	**72%**	100%
Excluding forward premium	95%	99%	82%	93%	100%	100%	95%	100%	**76%**	98%	89%	99%
Preceding 12 Months' Data												
Including forward premium				92%	85%	95%	96%	97%	90%	85%	84%	85%
Excluding forward premium				95%	95%	98%	99%	99%	93%	90%	89%	89%

[a] The test statistic for the Regression Method is $RVR = 1 - \left(\Sigma \left(\hat{b} X_i + Y_i \right)^2 \Big/ \Sigma Y_i^2 \right)$, the proportion of the mean-squared deviations from zero of the hedged item that is eliminated by the regression-based hedge. For a highly effective hedge, $RVR \geq 80\%$. Test values that fall outside the critical range are highlighted in bold.

5. Comparison of the Methods

The Dollar-Offset Method is well-established with an articulated 80/125 standard for effectiveness. Firms adopting this method should be aware, though, that the test statistic is sensitive to observations with small changes in value. The Relative-Difference Method addresses the problem of the sensitivity to small changes in value. It is not a widely accepted method, and there is no articulated standard for defining high effectiveness. Any such standard may need to be defined in terms of the time interval over which changes in value are measured. Our results suggest that this method requires careful attention to the treatment of any forward premium. Excluding the forward premium not only complicates the calculations but also affects reported periodic income.

The Variability-Reduction Method and the Regression Method tests are both sensitive to the number of observations used in the tests. Hedgers should also be aware that the articulated standard we used for both the Variability-Reduction Method and the Regression Method, 80%, appears to be less stringent than the 80/125 standard used for the Dollar-Offset Method. Lastly, there is a tendency to interpret the Regression Method only by its adjusted R^2, although ineffectiveness can also appear in both the slope and intercept. One of the main contributions of this paper is to suggest that *RVR* given by equation (6) is a more reliable measure of hedge effectiveness than the adjusted R^2.

In general, one cannot evaluate the effectiveness of a hedging relationship strictly on the basis of measuring whether the hedge is statistically effective or ineffective. There must be a plausible economic rationale for expecting the hedge to be effective. It is easy to construct examples in which a derivative position not expected to produce an 80/125 dollar offset over a wide range of variation in the hedged item nevertheless appears highly effective based on tests performed on historical data. For this paper, we believe the hedging relationship we chose was likely to produce an effective hedge, because of the anticipated co-movement of the hedged item, aluminum, and the derivative, an LME forward contract for aluminum. Given that perspective, we reproduce in Table 5 the results for all four methods for the case that includes one year of data and excludes the forward premium from the calculations. All but one of the test statistics indicate a highly effective hedge. We believe the consistency reflected in Table 5 suggests the merit of the following:

- Use at least a moderately large number of data points, 12 in this case.
- Include a moderately long period of analysis, the 12 months of the hedge in this case.
- Exclude the forward premium.

Table 5
Summary of Test Method Results for 1998 – 2000[a]

Year	1998	1999	1999	1999	1999	2000	2000	2000	2000
Quarter	4	1	2	3	4	1	2	3	4
Dollar-Offset	0.90	1.07	**0.79**	0.99	1.00	1.03	0.92	0.84	1.04
Relative-Difference	2.0%	-1.0%	-1.2%	-0.1%	0.0%	0.6%	-1.2%	-1.2%	-0.2%
Variability-Reduction: One-to-One Hedging	91%	88%	95%	96%	97%	88%	83%	82%	83%
Variability-Reduction: Regression Hedging	95%	95%	98%	99%	99%	93%	90%	89%	89%

[a] The test statistics are based on the previous 12 months of data excluding the forward premium. The test value that falls outside the critical range for the Dollar-Offset Method is highlighted in bold. All the other test values fall within the respective critical ranges.

III. Conclusion

SFAS 133 does not specify a bright line test to distinguish highly effective hedges from less effective hedges. We do not presume to do what SFAS 133 does not do. Rather, we discuss what we believe are potentially relevant testing methods and illustrate the choices firms must make when implementing them. The illustrations are asset and time specific. Firms should therefore be careful to understand the characteristics of the data on which they base tests of hedge effectiveness. Lastly, in this article we do not address a number of important and more complex issues, including dynamic hedging, option hedging, and the use of multiple derivatives to hedge. These more complex topics deserve their own articles.

References

Accounting for Certain Derivative Instruments and Certain Hedging Activities — An Amendment of FASB Statement No. 133, Statement of Financial Accounting Standards No. 138, Financial Accounting Standards Board, Norwalk, CT, June 2000.

Accounting for Derivative Instruments and Hedging Activities, Statement of Financial Accounting Standards No. 133, Financial Accounting Standards Board, Norwalk, CT, June 1998.

John M. Althoff and John D. Finnerty, "Testing Hedge Effectiveness," *FAS 133 and the New Derivatives Accounting Landscape*, Institutional Investor, New York, NY, Fall 2001: 44 – 51.

Eduardo Canabarro, "A Note on the Assessment of Hedge Effectiveness Using the Dollar Offset Ratio Under FAS 133," Research Report, Goldman Sachs & Co., New York, NY, June 1999.

Derivatives Implementation Group, Financial Accounting Standards Board, Norwalk, CT, Issues Nos. E4, E7, E8, E9, E10, E11, E12, E14, E15, E16, E19, and G9.

Louis H. Ederington, "The Hedging Performance of the New Futures Markets," *Journal of Finance* 34 (March 1979): 157-170.

Stephen Figlewski, "Hedging Performance and Basis Risk in Stock Index Futures," *Journal of Finance* 39 (July 1984): 657-669.

Charles T. Franckle, "The Hedging Performance of the New Futures Markets: Comment," *Journal of Finance* 35 (December 1980): 1273-1279.

Andrew Kalotay and Leslie Abreo, "Testing Hedge Effectiveness for FAS 133: The Volatility Reduction Measure," *Journal of Applied Corporate Finance* 13 (Winter 2001): 93-99.

Ira G. Kawaller and Paul D. Koch, "Meeting the 'Highly Effective Expectation' Criterion for Hedge Accounting," *Journal of Derivatives* 7 (Summer 2000): 79-87.

Robert C. Lipe, "Current Accounting Projects," presentation at the 1996 Twenty-Fourth Annual National Conference on Current SEC Developments, Office of the Chief Accountant, U.S. Securities and Exchange Commission, Washington, DC, December 10, 1996.

A.G. Malliaris and Jorge L. Urrutia, "The Impact of the Lengths of Estimation Periods and Hedging Horizons on the Effectiveness of a Hedge: Evidence from Foreign Currency Futures," *Journal of Futures Markets* 11 (no.3, 1991): 271-289.

Robert L. Royall, II, "Use of Regression in Assessing Hedge Effectiveness," *FAS 133 and the New Derivatives Accounting Landscape*, Institutional Investor, New York, NY, Fall 2001:52 – 61.

Stephen M. Swad, "Accounting and Disclosures for Derivatives," presentation at the 1995 Twenty-Second Annual National Conference on Current SEC Developments, Office of the Chief Accountant, U.S. Securities and Exchange Commission, Washington, DC, January 11, 1995.