Evaluation of Hedge Effectiveness Tests

Angelika C. Hailer, Siegfried M. Rump*

Abstract. According to IAS 39 or FAS 133 an a posteriori test for hedge effectiveness has to be implemented when using hedge accounting. Both standards do not regulate which numerical method has to be used.

A number of hedge effectiveness tests have been published recently. Such tests are of different quality, for example not all of them can deal with the problem of small numbers. This means a test might determine an effective hedge to be ineffective, a scenario which would increase the volatility in earnings. Therefore, it seems useful to have criteria at hand to discriminate and assess hedge effectiveness tests.

In this paper, we introduce such objective criteria, which we develop according to our understanding of minimum economic requirements. They are applicable to tests based on market values of two points in time as well as on tests based on time series of market values.

According to our criteria we compare common tests like the dollar offset ratio, regression analysis or volatility reduction, showing strengths and weaknesses. Finally we develop a new adjusted Hedge Interval test based on our previous one (2003). Our test does not show weaknesses of other effectiveness test.

Keywords: Hedge Accounting, Assessment of Effectiveness Tests, FAS 133, IAS 39


1 Introduction

The increasing importance of derivatives and hedges in today’s economy presents a number of challenges for accounting. Significantly, these challenges have not led to definitive guidelines directing accounting activities but merely to a regulatory framework defined in FAS 133 for US-GAAP (United States Generally Accepted Accounting Principles) and in IAS 39 for the International Accounting Standards.

Both require derivatives to be reported at fair value on the balance sheet. However, to avoid an increase in the volatility of earnings due to changes in their market values, they also allow for derivatives to be recognized in the reporting as part of a hedge.

Where the latter option is being chosen, a number of conditions have to be met. One of these conditions is an a posteriori test for hedge effectiveness. A

*Institute for Computer Science III, Hamburg University of Science and Technology, hailer@tu-harburg.de, rump@tu-harburg.de
Table 1: Definitions provided in IAS 39.

A **hedged item** is an asset, liability, firm commitment, highly probable forecast transaction or net investment in a foreign operation that (a) exposes the entity to risk of changes in fair value or future cash flows and (b) is designated as being hedged.

A **hedging instrument** is a designated derivative or (for a hedge of the risk of changes in foreign currency exchange rates only) a designated non-derivate financial asset or non-derivative financial liability whose fair value or cash flows are expected to offset changes in the fair value or cash flows of a designated hedge item.

**Hedge effectiveness** is the degree to which changes in fair value or cash flows of the hedged item that are attributable to a hedged risk are offset by changes in the fair value or cash flows of the hedging instrument.

A number of different tests are being used in practice; however, so far no criteria for assessing the quality of these tests have been established.

This paper starts by defining measurable criteria for the evaluation of effectiveness tests. These criteria are shown to be meaningful and most natural.

Existing tests can be broadly divided into those which are based on two points of time and those which are based on time series. We begin by examining a number of tests based on two points of time (dollar offset ratio, intuitive response to the small number problem, Lipp modulated dollar offset, Schleifer-Lipp modulated dollar offset, Gürtler effectiveness test, hedge interval). We then continue by looking at tests based on time series (expansion of test based on two dates, linear regression analysis, variability-reduction, and volatility reduction measure).

As we will demonstrate, the existing tests fail to meet the criteria developed in Section 3. However, by modifying the Hedge Interval Test discussed among others in Section 4, it is possible to obtain a test for hedge effectiveness which fulfills all of those criteria, as will be shown in Section 5.

## 2 Hedge Effectiveness according to IAS 39 and FAS 133

According to IAS we use the definitions of Table 1, i.e. in short a hedged item is the asset or liability responsible for the risk and a hedging instrument is the derivate to offset that risk. Additionally we consider a hedge position to be the added market value of hedged item and hedging instrument.

A hedging relationship qualifies for hedge accounting if and only if certain conditions defined in IAS 39 §88, or in FAS 133 §20, 21 resp. §28, 29 are met. One central condition part of both standards is the a posteriori assessment of hedge effectiveness as determined in IAS 39, §88 (e): “The hedge is assessed on an ongoing basis and determined actually to have been highly effective throughout the financial reporting periods for which the hedge was designated.”
Equivalently, retrospective evaluations are summarized in the Statement 133 Implementation Issue No. E7 as follows:

“At least quarterly, the hedging entity must determine whether the hedging relationship has been highly effective in having achieved offsetting changes in fair value or cash flows through the date of the periodic assessment.”

For considerations on the determination of market values and the use of marked-to-market values or clean values we refer to Coughlan, Kolb and Emery (2003). We concentrate on the problem of choosing an appropriate effectiveness test for a hedge position for which the fair values are already determined.

The dollar offset ratio is a common and probably the most simple method for asserting hedge effectiveness. It is explicitly explained in IAS 39 AG105. According to this measurement a hedge is regarded as effective if the quotient of changes of hedge item and hedging instrument is in the interval $\left[\frac{4}{5}, \frac{5}{4}\right]$.

Both standards explicitly state that other methods can be used as well: As part of FAS 133 §62 it is specified that the “appropriateness of a given method of assessing hedge effectiveness can depend on the nature of the risk being hedged and the type of hedging instrument used.” The equivalent formulation in IAS 39 AG107 is that this “Standard does not specify a single method for assessing hedge effectiveness. The method an entity adopts for assessing hedge effectiveness depends on its risk management strategy.” Ultimately the decision of which method is reasonable is left to the corporation’s auditors.

As no details for these methods are provided by the standards, a number of different tests have been published recently. To compare these hedge effectiveness tests, we develop assessment criteria in the following section.

3 Criteria for Hedge Effectiveness Tests

The purpose of the effectiveness test is to check whether the market development of hedged item and hedging instrument are almost “fully” offsetting each other. As stated in FAS 133 §62 this “Statement requires that an entity define at the time it designates a hedging relationship the method it will use to assess the hedge’s effectiveness in achieving offsetting changes in fair value or offsetting cash flow attributable to the risk being hedged.”

Even more explicitly this is addressed as part of FAS 133 §230: “A primary purpose of hedge accounting is to link items or transactions whose changes in fair values or cash flows are expected to offset each other. The Board therefore decided that one of the criteria for qualification for hedge accounting should focus on the extent to which offsetting changes in fair values or cash flows on the derivative and the hedged item or transaction during the term of the hedge are expected and ultimately achieved.”

So our first criterion should measure the degree of offsetting. This also implies that if concurrent market values in hedged item and hedged instrument are observed, a hedge should be regarded ineffective. As defined in both standards this means the relative deviation of the differences of changes of fair values to a perfect hedge should be limited.
Gürtler (2004) explains that the maximum possible gain or loss in the value of the hedge position should be limited. For example, a hedge where the difference in the market value of the hedging instrument is always \(-80\%\) of the difference in the hedged item, is regarded as effective under the dollar offset ratio. But for a large loss in the market value of the hedged item the reduction in the value of the hedge position would be significant. In other words this “problem of large numbers” has to be avoided, the second criterion.

The implementation of the dollar offset ratio frequently generates the difficulty known as the “problem of small numbers”. When there are just small changes in the market value of the hedged item, i.e. when the denominator gets small, the dollar offset ratio often indicates ineffectiveness although the tested hedge could be perfect. Therefore one third criterion should focus on this case: Equality of increase and decrease must not be strict, when nearly no changes in the market value of hedged item and hedging instrument are observed. In this case the hedge should be measured as effective.

Offsetting does mean that if one of the market values of hedged item or hedging instrument decreases the other will increase, and vice versa. This implies symmetry with respect to hedged item and hedging instrument, i.e. if for a hedge position the value of the hedged item increases about e.g. 100,000 US$ and the loss in the hedging instrument is 120,000 US$, then the same result should be obtained of the effectiveness test as it would be obtained for a gain of 100,000 US$ in the market value of the hedging instrument and a decrease of 120,000 US$ for the hedged item.

In addition, significantly over- or under-hedged positions should not be regarded as effective, which means no bias with respect to gain or loss in the hedge position should occur. The effectiveness test should have the same result when applied to differences in market values of both hedged item and hedging instrument as when applied to their negative differences, i.e. in the above example a decrease in the value of the hedged item about e.g. 100,000 US$ and an increase of 120,000 US$ in the value of the hedging instrument should lead to the same result.

Furthermore, a test should be expected to be scalable, so that if a hedge relationship is effective, then using the same percentage of both hedged item and hedging instrument should result in an effective hedge as well. Analogously this should hold for an ineffective hedge. Consequently, the amount of the hedging position should not influence the result of the test. If the test is not scalable, separating a hedge position in two parts could lead to one effective and one ineffective part, which would not be reasonable.

Figure 1 on page 5 contains an adjustment of our example (Hailer and Rump, 2003), which was expanded by Gürtler (2004). It contains three main stages. From \(t_0\) to \(t_3\) the hedge is obviously effective, from \(t_3\) to \(t_5\) it illustrates the problem of extreme losses as mentioned by Gürtler (2004) and from \(t_5\) on the development of the market values is concurrent, i.e. we have an increase instead of an offsetting of the risk, which means the hedge is strongly ineffective.

We investigate a number of tests in the following sections and summarize the results for this example in Tables 2 and 4. Deviations from the expected results
Figure 1: Illustration of a sample market value development of hedged item and hedging instrument which belongs to a perfect hedge for balance sheet dates $t_0$ to $t_3$. From $t_3$ to $t_5$ an probably rather theoretical extreme movement in the market value can be observed and from $t_5$ the market values are concurrent which obviously implies the hedge to be ineffective.
are emphasized. A cursive no implies to regard an effective hedge as ineffective and therefore cause the hedge position to be dissolved. So the volatility in earnings would be increased. Even worse it a cursive yes, which allows an ineffective hedge to qualify for hedge accounting. In this case, earnings or losses can be hidden in the hedge position.

For developing comparable criteria according to the objectives described above, we distinguish between effectiveness tests based on the market value on two points of time and tests based on time series of market values.

3.1 Tests Based on Two Points of Time

Let $GG_t$ denote the market value of the hedged item at date $t$ and let $\Delta GG$ denote the market value difference in the hedged item, let $SG_t$ and $\Delta SG$ be defined analogously for the hedging instrument and $GP_t$ and $\Delta GP$ analogously for the hedge position, i.e. $GP_t = GG_t + SG_t$.

As mentioned in Statement 133 Implementation Issue No. E8 (2000) the difference $\Delta GG$ and respectively $\Delta SG$ can be calculated at date $t$ on a period-by-period approach as $\Delta GG = GG_t - GG_{t-1}$, or cumulatively as $\Delta GG = GG_t - GG_0$:

“In periodically (that is, at least quarterly) assessing retrospectively the effectiveness of a fair value hedge (or a cash flow hedge) in having achieved offsetting changes in fair values (or cash flows) under a dollar-offset approach, Statement 133 permits an entity to use either a period-by-period approach or a cumulative approach on individual fair value hedges (or cash flow hedges).”

This citation “relates to an entity’s periodic retrospective assessment and determining whether a hedging relationship continues to qualify for hedge accounting”. As already stated we consider the a posteriori test of hedge effectiveness and do not refer to the measurement of actual ineffectiveness that has to be reflected in earnings according to FAS 133 §22 or §30. For these, the Standard requires calculations on a cumulative basis.

In general it is not advisable to use only local information for a global measurement. In this case, when relying on period-by-period information, small, slow changes of the market value of the hedge position, which are not recognized as significant, may sum over time. In accordance with Coughlan, Kolb and Emery (2003) and Finnerty and Grand (2002) the following considerations for measurements relying on data of two points of time are based on the use of cumulative differences.

In addition to the general consideration for effectiveness tests, we expect a hedge effectiveness test based on two points of time to be continuous in the sense that there are no unnatural limits for the transition from effectiveness to ineffectiveness. For example, if for a decrease in the hedged item of 100.01 US$ a hedge is regarded as effective for an increase in the market value of the hedging instrument of 80.01 US$ and as not effective for and increase of 80.00 US$, then we cannot understand a hedge with a decrease in the hedge item of 100.00 US$
to be effective for a range of changes in the market value of the hedging instrument from -100.00 US$ to 100.00 US$. We assume these marginal values to be unnatural, and therefore should be avoided. So we expect a smooth transition from effectiveness to ineffectiveness.

The degree of offsetting can be geometrically interpreted when plotting $\Delta SG$ against $\Delta GG$ as shown in Figure 2.

The objective of measuring offsetting is then expressed by the relation

$$\Delta SG \approx -\Delta GG .$$

Therefore a hedge is highly effective if the point with the coordinates $\Delta GG$ and $\Delta SG$ is on or near the northwest-southeast diagonal, the bisecting line of the second and fourth quadrant. To determine the degree of “nearness” is the necessary task of a hedge effectiveness test.

All known tests based on market values of just two dates can be illustrated in the plane spanned by the coordinates $\Delta GG$ and $\Delta SG$ as shown in Figure 2.

The area where a hedge is regarded effective can be defined by bounding functions

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \text{and} \quad \bar{f} : \mathbb{R} \rightarrow \mathbb{R} .$$

They define a hedge to be effective if

$$f(\Delta GG) \leq \Delta SG \leq \bar{f}(\Delta GG) .$$

The bounding functions of some of the known measurements contain constants that depend on the initial value of the hedge position $GP_0 = GG_0 + SG_0$. When we need to indicate this parameterization for the underlying hedge position $GP_0$ we write $f_{GP_0}$ and $\bar{f}_{GP_0}$ as well. In the figures the effective area is marked in grey.

For example, using the dollar offset ratio a hedge is effective if the point with the coordinates $\Delta GG$ and $\Delta SG$ is part of two cones, see gray area in Figure 4.

![Figure 2: The plane spanned by $\Delta GG$ and $\Delta SG$ used for geometrical interpretation of effectiveness tests. The coordinates of a perfect hedge are on or near the dashed line.](image-url)
on page 13. In this case the problem of small numbers is visible by the tolerance being very small close to the origin, the vertex of the cones.

From this illustration and the general considerations on the objectives of an effectiveness test at the beginning of this section, we deduce the following measurable criteria:

**Criterion 1** We assume an effectiveness test to comply with the following requirements:

(i) **Offsetting:** The surface indicating effectiveness should contain all effective hedges which are represented as part of the northwest-southeast diagonal. The relative deviation of this line should be limited for all $\Delta GG \in \mathbb{R}$.

(ii) **Large numbers:** The maximum gain or loss in the hedge position should be limited. So the absolute deviation of the northwest-southeast diagonal should be limited for all $\Delta GG \in \mathbb{R}$.

(iii) **Small numbers:** To avoid numerical problems the area should at no point have a vanishing “diameter”, i.e. for arbitrary $\Delta GG$

$$|\overline{\mathcal{F}}(\Delta GG) - \underline{\mathcal{F}}(\Delta GG)| > \delta$$

for fixed $\delta \geq 0$.

(iv) **Symmetry:** The surface should be symmetric to the northwest-southeast diagonal for symmetry in gain and loss of the hedge position and to the southwest-northeast diagonal to guarantee symmetry in hedged item and hedging instrument.

This means, assumed the functions $\mathcal{F}$ and $f$ are invertible the equations

$$f(x) = f^{-1}(x) \quad \text{and} \quad \overline{\mathcal{F}}(-f(x)) = -x$$

should hold true for all $x \in \mathbb{R}$.

(v) **Scalability:** For all percentages $\alpha \in (0, 1]$ we should obtain

$$\underline{\mathcal{F}}_{\alpha \cdot GP_0}(\alpha \cdot \Delta GG) = \alpha \cdot \underline{\mathcal{F}}_{GP_0}(\Delta GG)$$

and

$$\overline{\mathcal{F}}_{\alpha \cdot GP_0}(\alpha \cdot \Delta GG) = \alpha \cdot \overline{\mathcal{F}}_{GP_0}(\Delta GG)$$

for all $\Delta GG \in \mathbb{R}$. When regarding the functions $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$ mathematically correct as functions of two parameters $GP_0$ and $\Delta GG$, i.e. $\underline{\mathcal{F}}, \overline{\mathcal{F}} : \mathbb{R}^2 \rightarrow \mathbb{R}$, then $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$ are said to be homogeneous of degree 1.

(vi) **Smooth transition:** The transition between an effective and ineffective hedge should be natural, which implies the bounding functions $\underline{\mathcal{F}}$ and $\overline{\mathcal{F}}$ to be continuous.

These criteria are independent of each other, which means it is not possible to deduce one from the other. So investigating an effectiveness test, all of the criteria (i) to (vi) have to be considered. In Section 4.1 we apply these to the main effectiveness tests known.
3.2 Tests Based on Time Series

According to IAS 39, AG106 effectiveness “is assessed, at a minimum, at the time an entity prepares its annual or interim financial statements.” And more detailed as part of FAS 133, §20 (b) and §28 (b), an “assessment of effectiveness is required whenever financial statements or earnings are reported, and at least every three month.” Therefore, we focus on assessments on a quarterly basis.

The first time a hedge position fails the hedge effectiveness test, it has to be dissolved as determined as part of IAS 39, AG113: “If an entity does not meet hedge effectiveness criteria, the entity discontinues hedge accounting from the last date on which compliance with hedge effectiveness was demonstrated.” An equivalent explanation can be found in FAS.

Applying one of the methods based on two points of time quarterly indirectly includes historical data to the measurement. But this still incorporates only quarterly market values to the test. Thus, more detailed statistical approaches have been developed which are applicable to input data of time series of market values. The main idea is that the evaluation of the effectiveness of a hedge can be optimized when using as much information as available.

A problem appears in day to day business in the generation of these time series: IT-systems used for accounting purposes are often designed for punctual evaluations on reporting days. And even if accounting systems are able to deal with daily values, according to the securities used for the hedge daily market values may not be available.

Further on it seems to be common consent that statistical tests should be based at least on some 30 data points. In addition, for certain statistics the time intervals used should correspond to the hedged horizon as explained by Kawaller and Koch (2000):

“Unfortunately, the need to use either quarterly price changes or price changes measured over the same time frame as the hedged horizon is common to any method of statistical analysis.”

Using quarterly data this would imply historical data for at least seven years from inception of the hedge before a test could be applied, in contradiction to the assessment recommended on an ongoing basis by the standards. So the statistical requirements often cannot be fulfilled because of a lack of data.

Again the problem of small numbers may occur: For illustration we expand the example proposed by Kalotay and Abreo (2001): They consider a 100 US$ million bond hedged with an interest rate swap and a 10,000 US$ rise in the value of the bond as well as a fall of 4,000 US$ in the value of the swap. We assume the initial value of the hedging instrument to be zero at inception of the hedge. According to the dollar offset ratio this hedge is ineffective, which contradicts Kalotay and Abreo’s statement that “the net change of US$ 6,000 is a miniscule 0.006% of the face amount”.

We construct a time series of 61 dates, where the market values of \( t_0 \) and \( t_{60} \) are those suggested by Kalotay and Abreo. The other dates are interpolated
regarding a randomly perturbed logarithmic increase for the market value of the hedged item and decrease for the hedging instrument, as illustrated in Figure 3.

In day to day business this effect due to unchanged market values will occur less often when using daily market values for a larger time period, as they can be expected to have significant changes. We show in detail in the next section that a number of known measurements based on time series result in an ineffective hedge for this example. So the problem of small numbers is not necessarily avoided, although the probability of occurrence is reduced compared to the dollar offset ratio.

Common to all known statistical measurements based on time series is the fact that the influence of the offsetting ratio on one point of time is reduced. Therefore, the management decision, whether or not to regard a hedge as effective even if it gets ineffective for single points in time should be discussed in advance.

Assume we have $n$ dates where market values are available. For the dates $i = 1, \ldots, n$ let $\Delta GG_i$ denote the cumulative difference in the market value of

Figure 3: Illustration of the adjusted example of Kalotay and Abreo (2001) representing a market development for 60 days where nearly no changes can be observed. All tests evaluated in Section 4.2 result in an ineffective hedge. According to Kalotay and Abreo “the net change of $6,000 is a miniscule 0.006% of the face amount”, and therefore the hedge should be regarded as effective.
the hedged item i.e.
\[ \Delta GG_i = GG_i - GG_0 \]
or the period-by-period difference
\[ \Delta GG_i = GG_i - GG_{i-1} , \]
and for the hedging instrument \( \Delta SG_i = SG_i - SG_0 \) or \( \Delta SG_i = SG_i - SG_{i-1} \), respectively.

Let \( \vec{\Delta GG} \) denote the \( n \)-dimensional vector containing all \( \Delta GG_i \) and \( \vec{\Delta SG} \) the \( n \)-dimensional vector containing all \( \Delta SG_i \). For adjusting the criteria for two points to higher dimensions as necessary for time series, we introduce the effectiveness test function
\[
T : \mathbb{R}^{2n} \rightarrow \{0, 1\} : (\vec{\Delta GG}, \vec{\Delta SG}) \rightarrow \{\text{not effective}, \text{effective}\}.
\]

FAS 133 does not detail the requirements for statistical measurements, but the following warning is contained in Implementation Issue No. E7:

“The application of a regression or other statistical analysis approach to assessing effectiveness is complex. Those methodologies require appropriate interpretation and understanding of the statistical inferences.”

Independently of this point we suppose the general criteria we have described for an effectiveness test at the beginning of this section should be fulfilled, regardless of the complexity of the test.

So analog to the effectiveness tests based on two points we can formulate measurable criteria:

**Criterion 2** Suppose \( n \in \mathbb{N} \) and \( T : \mathbb{R}^{2n} \rightarrow \{0, 1\} \) is an effectiveness test for hedge accounting. Then \( T \) should have the following properties:

(i) **Offsetting:** The scatter plot of all points \((\Delta GG_i, \Delta SG_i)\) should be close to the northwest-southeast diagonal, i.e. the relative deviation of this line should be limited for all points.

(ii) **Large numbers:** For all points \((\Delta GG_i, \Delta SG_i)\) the maximum distance to the northwest-southeast diagonal, i.e. the absolute deviation, should be limited.

(iii) **Small numbers:** The problem of small numbers should be avoided. Therefore, if
\[ \max_{i \in \{1, \ldots, n\}} \{\max\{|\Delta GG_i|, |\Delta SG_i|\}\} \leq c \]
for a constant \( c \) which may depend on the initial value of the hedge position, i.e. \( c = c_{GP_0} \) the hedge should be regarded as effective.
(iv) **Symmetry:** Using the points with the coordinates \((\Delta SG_i, \Delta GG_i)\) for testing effectiveness should imply the same result as the use of \((\Delta GG_i, \Delta SG_i)\), i.e.

\[
T(\Delta GG, \Delta SG) = T(\Delta SG, \Delta GG),
\]

and symmetry with respect to gains and losses should be fulfilled, i.e.

\[
T(\Delta GG, \Delta SG) = T(-\Delta GG, -\Delta SG).
\]

(v) **Scalability:** Let \(\alpha \in (0, 1]\). Then the property

\[
T(\Delta GG, \Delta SG) = T(\alpha \cdot \Delta GG, \alpha \cdot \Delta SG)
\]

should hold true.

In Section 4.2 we apply these criteria to the main effectiveness tests known, which are based on times series of market values. The results for all tests investigated are summarized in Table 5. In this case we expect a hedge at time \(t_5\) to be not effective, as the decision is based on the time period from \(t_4\) to \(t_5\), which contains the problem of large numbers at the beginning.

According to standard statistical notation we use the following definitions: Let \(\bar{X} = \sum_{i=1}^{n} x_i\) denote the mean value of \(x_1, \ldots, x_n\).

For \(n\) observation dates let \(x_i\) and \(y_i\) denote market values or changes in market values. Then the empirical variance \(\sigma^2_x\) and the empirical covariance \(\sigma_{xy}\) are defined as

\[
\sigma^2_x = \frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \bar{X})^2 \quad \text{and} \quad \sigma_{xy} = \frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \bar{X}) \cdot (y_i - \bar{Y}),
\]

and the standard deviation can be estimated as \(\sqrt{\sigma^2_x}\).

### 4 Evaluation of Common Hedge Effectiveness Tests

#### 4.1 Tests Based on Two Points of Time

##### 4.1.1 Dollar Offset Ratio

The dollar offset ratio is defined in the following effectiveness test.

**Test 1** A hedge is regarded effective if the quotient of changes of hedge item and hedging instrument is part of the interval \([80\%, 125\%]\), i.e. if

\[
\frac{\Delta SG}{\Delta GG} \in \left[\frac{4}{5}, \frac{5}{4}\right].
\]

We summarize the results of this test for the example presented in Figure 1 in Table 2 on page 15. Geometrically a hedge is effective if the coordinates given by the values of \(\Delta GG\) and \(\Delta SG\) fall in the cones being spanned by the lines \(\Delta SG = -\frac{4}{5} \Delta GG\) and \(\Delta SG = -\frac{5}{4} \Delta GG\), the gray area in Figure 4.
Figure 4: Comparison of the geometrical interpretation of the dollar offset ratio, the intuitive response and the Lipp Modulated dollar offset ratio. A hedge is effective if the coordinates of the changes of hedged item and hedging instrument are part of the grey area.
Figure 5: Comparison of the geometrical interpretation of the Schleifer-Lipp Modulated dollar offset ratio, the effectiveness test proposed by Gürtler and the hedge interval. A hedge is effective if the coordinates of the changes of hedged item and hedging instrument are part of the grey area. In particular regarding small or medium scale all points are effective when using the test of Gürtler.
Table 2: Results of the application of the hedge effectiveness tests based on two points of time for the hedge introduced in Figure 1 on page 5. Deviations from the expected results are emphasized.

<table>
<thead>
<tr>
<th>Expected to be effective</th>
<th>Dollar Offset Ratio effective</th>
<th>Intuitive Response effective</th>
<th>Lipp effective</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$ yes</td>
<td>100.02% yes</td>
<td>100.02% yes</td>
<td>100.02% yes</td>
</tr>
<tr>
<td>$t_2$ yes</td>
<td>100.00% yes</td>
<td>100.00% yes</td>
<td>100.00% yes</td>
</tr>
<tr>
<td>$t_3$ yes</td>
<td>70.00% no</td>
<td>yes</td>
<td>99.70% yes</td>
</tr>
<tr>
<td>$t_4$ no</td>
<td>112.50% yes</td>
<td>112.50% yes</td>
<td>112.50% yes</td>
</tr>
<tr>
<td>$t_5$ yes</td>
<td>n/a</td>
<td>yes</td>
<td>100.00% yes</td>
</tr>
<tr>
<td>$t_6$ no</td>
<td>-99.98% no</td>
<td>-99.98% no</td>
<td>no</td>
</tr>
<tr>
<td>$t_7$ no</td>
<td>-99.99% no</td>
<td>-99.99% no</td>
<td>no</td>
</tr>
<tr>
<td>$t_8$ no</td>
<td>-100.00% no</td>
<td>-100.00% no</td>
<td>no</td>
</tr>
</tbody>
</table>

| Schleifer-Lipp $S_{LP} = 0.6$ effective | Gürtler’s Test $GP_T$ effective | Hedge Interval based on $[\frac{4}{19}, \frac{5}{19}]$ effective $x$ | $|x| \leq 9$ | Hedge Interval based on $[\frac{9}{19}, \frac{10}{19}]$ effective $x$ | $|x| \leq 19$ |
|----------------------------------------|---------------------------------|------------------------------------------------|-------------|------------------------------------------------|-------------|
| $t_1$ 100.02% yes                      | 100.00% yes                     | 0.99 yes                                     | 0.97 yes    |                                                |             |
| $t_2$ 100.00% yes                      | 100.00% yes                     | -1.00 yes                                    | -1.00 yes   |                                                |             |
| $t_3$ 99.98% yes                       | 100.00% yes                     | -0.04 yes                                    | -0.17 yes   |                                                |             |
| $t_4$ 112.50% yes                      | 50.00% no                       | -4.00 yes                                    | -21.50 no   |                                                |             |
| $t_5$ 100.00% yes                      | 100.00% yes                     | 0.00 yes                                     | 0.00 yes    |                                                |             |
| $t_6$ no                              | 91.67% yes                      | -80.99 no                                    | -360.95 no  |                                                |             |
| $t_7$ no                              | 83.33% yes                      | -80.99 no                                    | -360.98 no  |                                                |             |
| $t_8$ no                              | 75.00% yes                      | -81.00 no                                    | -361.00 no  |                                                |             |
All criteria proposed in Section 3 except (ii), the maximum deviation for large values of $\Delta GG$, and except (iii) are fulfilled:

As one can easily see the surface is symmetric, and the functions $f$ and $\overline{f}$ are continuous. This is also true in the origin, even though at this point they are not differentiable. Scalability is fulfilled as any percentage would be canceled in the fraction.

As already mentioned as “problem of small numbers” for arbitrary $\delta > 0$ and $\Delta GG = 0$ we obtain

$$|\overline{f}(\Delta GG) - f(\Delta GG)| = 0 < \delta,$$

so criterion (iii) is not fulfilled. And as explained by Gürtler (2004) the maximum loss of the hedge position is not limited as the distance of the bounding lines of the cones is arbitrarily large for large absolute values of $\Delta GG$ and $\Delta SG$.

Out of these reasons modifications of the dollar offset ratio have been developed, which are investigated in the following sections.

### 4.1.2 Intuitive Response to the Small Number Problem

In the transition from accounting corresponding to the German accounting standards according to HGB (Handelsgesetzbuch) to IAS or US-GAAP, German companies encountered the problem of small numbers when implementing the dollar offset ratio.

One intuitive way to respond to this problem is to implement a fixed maximum value for changes in the market development of hedged item and hedging instrument, up to that a hedge is considered effective without further test. As far as we know this was proposed and accepted by one of the leading auditing firms.

**Test 2** A hedge is effective

$$\begin{cases} \text{without test} & \text{for } \max\{|\Delta SG|, |\Delta GG|\} \leq c \\ \text{if } -\Delta SG \overline{\Delta GG} \in [\frac{4}{5}, \frac{5}{4}] & \text{else.} \end{cases}$$

This test is only scalable, if the value of $c$ is dependent on the value of the hedge position at inception of the hedge, in our example we take a value of $1\%$, i.e.

$$c_{GP_0} = 0.001(GG_0 + SG_0) = 100.$$

As shown in Figure 4 on page 13 the functions $f$ and $\overline{f}$ are not continuous, resulting in this example in an unexplainable transition of effectiveness for values of $\Delta GG = 100$ or $\Delta SG = 100$.

The problem of maximum deviation remains unsolved as already explained for the dollar offset ratio. But except criteria (ii) and (v), all others are fulfilled.
### 4.1.3 Lipp Modulated Dollar Offset

Further modification of the dollar offset are published by Schleifer (2001). These add basic values for the consideration of relative changes and discuss different values $M_p$ for these. The Lipp modulated dollar offset is one measurement that is defined by the following test.

**Test 3** *A hedge is regarded as effective if*

\[
\text{sgn}(\Delta SG) = -\text{sgn}(\Delta GG) \quad \text{and} \quad \frac{\left|\Delta SG\right| + NT_A}{\left|\Delta GG\right| + NT_A} \in \left[\frac{4}{5}, \frac{5}{4}\right],
\]

*where $NT_A$ is the absolute value of a noise threshold.*

Schleifer (2001) proposes a definition of $NT_A = M_p \frac{NT_N}{1000}$, where $NT_N$ is a user-defined noise threshold and $M_p$ depends on the hedged item’s cash flows, in first-order approximation the present-value of one leg of the hedging instrument.

In the plane spanned by $\Delta GG$ and $\Delta SG$ we get for $\Delta GG \geq 0$ the inequalities

\[
-\frac{NT_A}{4} - \frac{5}{4} \Delta GG \leq \Delta SG \leq -\frac{NT_A}{5} - \frac{4}{5} \Delta GG \quad \text{and} \quad \Delta SG \leq 0,
\]

and for $\Delta GG \leq 0$

\[
-\frac{NT_A}{5} - \frac{4}{5} \Delta GG \leq \Delta SG \leq -\frac{NT_A}{4} - \frac{5}{4} \Delta GG \quad \text{and} \quad \Delta SG \geq 0.
\]

In our calculation for Table 2 and Figure 5 we use a value of $NT_A = 10$.

One problem concerning the vertex of the cone remains unsolved: Nearly no changes in the market value of hedged item and hedging instrument could imply an ineffective hedge relationship. For example, a raise of both values of $\frac{1}{100}$ basis point can be interpreted as noise in the data.

Further on this test is only scalable if $NT_A$ is a fixed percentage of $GR_0$. Obviously the bounding functions $f$ and $\overline{f}$ are not continuous. So criteria (ii), (iii) and (v) are not fulfilled.

### 4.1.4 Schleifer-Lipp Modulated Dollar Offset

One suggested modification of this measurement is the Schleifer-Lipp modulated offset.

**Test 4** *A hedge is effective, if*

\[
\text{sgn}(\Delta SG) = -\text{sgn}(\Delta GG)
\]

*and for $S_T > -1*

\[
\frac{\left|\Delta SG\right| \left(\frac{\sqrt{\left|\Delta SG\right|^2 + \left|\Delta GG\right|^2}}{NT_A}\right)^{S_T} + NT_A}{\left|\Delta GG\right| \left(\frac{\sqrt{\left|\Delta SG\right|^2 + \left|\Delta GG\right|^2}}{NT_A}\right)^{S_T} + NT_A} \in \left[\frac{4}{5}, \frac{5}{4}\right].
\]
For a parameter of $S_T = 0$ the Lipp modulated dollar offset is the same as the Schleifer-Lipp modulated dollar offset. Again we use a value of $NT_A = 10$ in our calculations.

This test is only scalable if $NT_A$ is a fixed percentage of $GP_0$. It is similar to the Lipp modulated dollar offset and does not satisfy criteria (ii), (iii) and (v).

### 4.1.5 Gürtler Effectiveness Test

Gürtler (2004) develops a test from a risk-theoretical basis as he minimizes the maximum possible loss of the hedge position:

**Test 5** A hedge position is regarded as effective if and only if

$$1 - \frac{\alpha}{a} \leq \frac{GP_t}{GP_0} \leq 1 + \frac{\alpha}{a},$$

where Gürtler suggests to use a value of $\frac{\alpha}{a} = 25\%$.

As Gürtler shows this is equivalent to a hedge being effective if

$$1 - \frac{\alpha}{a} \frac{GP_0}{|\Delta GG|} \leq -\frac{\Delta SG}{\Delta GG} \leq 1 + \frac{\alpha}{a} \frac{GP_0}{|\Delta GG|}$$

or equivalently

$$-\Delta GG - \frac{\alpha}{a} GP_0 \leq \Delta SG \leq -\Delta GG + \frac{\alpha}{a} GP_0.$$

Geometrically these inequalities represent a fixed band around the northwest-southeast diagonal. The width of this band depends on the constant $\frac{\alpha}{a}$ and on the value of hedge position at inception of the hedge. For illustration see Figure 5 on page 14.

Up to a maximum loss of 25\% of the hedge position $GP_0$, Gürtler’s test always results in an effective hedge. Normally one would not observe such extreme movements in the market values as in our example in periods $t_3$ to $t_5$. Therefore, with this measurement a lot more hedges are qualifying for hedge accounting than when applying dollar offset ratio.

So this test satisfies all criteria but the first, which must not only be considered as the most important one according to the standards but also describes the main objective for implementing an effectiveness test.

This measurement is geometrically equivalent to the relative-difference test described by Finnerty and Grand (2002), according to which a hedge is effective if

$$\left| \frac{\Delta SG + \Delta GG}{GG_0} \right| \leq 3\%.$$  

This method is investigated by Finnerty and Grand (2002) with a suggested bandwidth of $\sqrt{2} \cdot 0.03 \cdot GG_0$, whereas Gürtler proposes a bandwidth of $\sqrt{2} \cdot 0.25 \cdot (GG_0 + SG_0)$. In our example we obtain a bandwidth of 4,242.64 US$
in the first and of 35,355.34 US$ in the second case. With this narrower band the expected results would be obtained in our example, but the first criterion denoting the relative deviation is not met. This first criterium corresponds to the definition of offsetting in both Standards.

4.1.6 Hedge Interval

We presented a hedge interval with the following properties (2003):

- On a large scale the interval is essentially identical to the known dollar offset ratio.
- For small numbers the intersection of the cones of the dollar offset ratio is broadened.
- The transition from large to small is continuous.

The measurement is defined as follows:

**Test 6** A hedge is to be regarded as effective if and only if

\[
\left| \frac{40 \Delta SG + 41 \Delta GG}{\sqrt{\Delta GG^2 + c}} \right| \leq 9. \quad (\ast)
\]

The lower and upper bounding functions \( f \) and \( \bar{f} \) are approximating the lines

\[
\Delta SG = -\frac{5}{4} \Delta GG \quad \text{and} \quad \Delta SG = -\frac{4}{5} \Delta GG
\]

for larger values and broadening the intersection of the two cones. The parameter \( c \) in (\ast) determines the distance of the approximating function to the cones in the origin.

The test is not sensitive to changes in the parameter \( c \), so \( c \) may vary within about an order of magnitude without causing too much change of behavior. If all balance sheet items regarded for hedging are about the same order of size the parameter \( c \) could be determined as a constant. According to the suggestions of Gürtler and to guarantee scalability we advice to introduce a dependency of \( c \) on the squared of the initial hedge position. For example, a value of \( c = 10^{-7} \cdot GP_0^2 \) seems appropriate in all real cases. For our calculations of the example described in Figure 1 we used this value of \( c = 10^{-7} \cdot GP_0^2 = 1000 \).

Interpreting the hedge interval from the view of numerical mathematics we look at relative error for large \( \Delta \) and at absolute error for small \( \Delta \). The transition is smooth.

So all criteria except (ii), the maximum deviation for large values of \( \Delta GG \), are satisfied. As Gürtler stated it is common practice to use an interval for the dollar offset ratio of \( \left[ \frac{3}{10}, \frac{4}{5} \right] \) instead of \( \left[ \frac{4}{5}, \frac{3}{2} \right] \). At the end of this paper we present a generalization of this hedge interval to arbitrary underlying dollar offset intervals and adjust it to fulfill criterion (ii) as well.

19
4.2 Tests Based on Time Series

4.2.1 Expansion of Tests based on Two Dates

First of all a simple test which is easy to implement can be deduced from all of the measurements presented for two points of time as follows: A hedge is regarded as highly effective if and only if the coordinates of the differences $\Delta GG_i$ and $\Delta SG_i$ are part of the effective area for all dates $i$.

This is probably the strictest of the effectiveness test based on time series as it does not allow the hedge to get ineffective at one single point in the sense defined in the underlying two point criterion.

One modification of this measurement is presented by Coughlan, Kolb and Emery (2003), answering the question whether or not a hedge has to be effective on all dates where market values are available. They introduce a compliance level as

$$\text{Compliance level} = \frac{\text{Number of compliant results}}{\text{Number of data points}},$$

and suggest a threshold of 80%, i.e. to regard a hedge effective if the compliance level has a value greater than 80%.

These tests fulfill all of the corresponding criteria to their underlying method for two points of time, except the first two concerning offsetting: When a compliance level lower than 100% is used, the maximum relative and absolute deviation is not limited all of the time.

4.2.2 Linear Regression Analysis

In risk management calculations focusing on hedging strategies the hedge effectiveness can be tested using a linear regression on the ratio of the differences (Hull, 2003), even if Kalotay and Abreo (2001) refer to it as “arcane statistics such as R-squared”.

This method is explicitly mentioned in IAS 39 F.4.4, where its application is detailed as follows: “If regression analysis is used, the entity’s documented policies for assessing effectiveness must specify how the results of the regression will be assessed.”

The linear regression is based on the equations

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i.$$

We refer to the version of Coughlan, Kolb and Emery (2003), where the independent variable $x$ refer to the hedged item and the dependent variable $y$ to the hedging instrument. This does not correspond to Kawaller and Koch (2000), who interchange the variables $x$ and $y$.

For obtaining offsetting in differences of market value developments, i.e. $x_i = \Delta GG_i$ and $y_i = \Delta SG_i$, the value of the slope $\hat{\beta}_1$ should be close to $-1$ and of the intercept $\hat{\beta}_0$ close to 0. Further on, generally the adjusted R-squared, $R^2$, is determined and it seems to be common consent that it has to have a value greater than 80% for a hedge to be effective.
As stated in Finnerty and Grand (2002) “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.” Naturally we assume that for hedge effectiveness certain requirements for $\hat{\beta}_0$ and $\hat{\beta}_1$ have to be fulfilled. Otherwise, with a value of $\hat{\beta}_1 = 1$ a perfectly ineffective hedge would be regarded effective, or with $\hat{\beta}_0 \neq 0$ over- or under-hedging could be accepted as effective.

Coughlan, Kolb and Emery (2003) mention that it is important when using a linear regression to validate the statistical significance with a t-test, and suggest to use for this t-test a confidence level of 95%.

The standard t-test for a linear regression tests the hypotheses $H_0$, that the parameters $\beta_0$ and $\beta_1$ are equal to zero, to determine if the influence of these is statistically significant. If a probability less than 5 % is obtained than $H_0$ can be rejected. Coughlan, Kolb and Emery (2003) provide a sample output of a statistical tool for regression analysis. According to the data this standard t-test seems to have been used.

In the case of an effective hedge we expect to obtain values for $\beta_0$ close to 0 and for $\beta_1$ close to $-1$. So probably we will obtain the result that we can reject $H_0 : \beta_1 = 0$ and not reject $H_0 : \beta_0 = 0$. In our calculations for the example described in Figure 1 on page 5 we included this test for $\beta_1$ and were always able to reject $H_0 : \beta_1 \neq 0$.

We suppose in this case other hypotheses like $H_0 : \beta_1 \neq -1$ could be more appropriate. In addition, we would suggest to a priori set the intercept to zero in a regression based on changes of market values for a hedge position.

Nevertheless for the comparison of the tests we focus on the evaluation of $\hat{\beta}_0$, $\hat{\beta}_1$ and R-squared. Coughlan, Kolb and Emery (2003) use for the retrospective regression analysis an effectiveness threshold of $-80\%$ for the correlation and $-0.80$ to $-1.25$ for the slope. According to Kalotay and Abreo (2001) and Kawaller (2002) in our examples we regard a threshold of 80% for the R-squared.

The parameter for $\hat{\beta}_0$ and $\hat{\beta}_1$ are determined with standard statistic as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{X})^2} = \frac{\sigma_{xy}}{\sigma_x^2} \quad \text{and} \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}.$$ 

The R-squared can be determined with the empirical variances and covariance as

$$R^2 = \frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2}.$$ 

For the linear regression different dependent and independent variables are discussed by Kawaller and Koch (2000) when investigating a priori hedge effectiveness tests. The main concern is whether regression should be applied to data on price levels or on price changes. No method is directly recommended by Kawaller and Koch (2000):

“This discussion might suggest that the appropriate indicator of hedge effectiveness should be the correlation of price levels, as opposed to price changes, but this conclusion is similarly flawed. The
<table>
<thead>
<tr>
<th>Period</th>
<th>Fair Value: ( t_0 - t_3 )</th>
<th>Fair Value: ( t_3 - t_5 )</th>
<th>Fair Value: ( t_5 - t_8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>96,000</td>
<td>85,000</td>
<td>800</td>
</tr>
<tr>
<td></td>
<td>100,000</td>
<td>100,000</td>
<td>1,500</td>
</tr>
<tr>
<td></td>
<td>-2,000</td>
<td>-10,000</td>
<td>-800</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2,000</td>
<td>0</td>
<td>800</td>
</tr>
<tr>
<td></td>
<td>4,000</td>
<td>-1,500</td>
<td>-1,500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Period-by-Period: ( t_0 - t_3 )</th>
<th>Period-by-Period: ( t_3 - t_5 )</th>
<th>Period-by-Period: ( t_5 - t_8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4,000</td>
<td>-200,000</td>
<td>85,000</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-4,000</td>
<td>0</td>
<td>-10,000</td>
</tr>
<tr>
<td>0</td>
<td>-400,000</td>
<td>0</td>
</tr>
<tr>
<td>-400,000</td>
<td>-1,500</td>
<td>-1,500</td>
</tr>
</tbody>
</table>

Figure 6: Scatter plots for the linear regression for the example introduced in Figure 1 for the different dependent and independent variables.
statement that two price levels are highly correlated does not necessarily imply a reliable relationship between their price changes over a particular hedge horizon, which is the issue of concern for the FASB.”

For price changes it has to be further distinguished between price changes on a period-by-period assessment or cumulative, i.e. calculating all differences to the inception of the hedge. We do not consider price changes based overlapping periods for the evaluation of the effectiveness tests based on time series. For a discussion we refer to Kawaller and Koch (2000).

In summary, we use the following simplified criteria, even if no explicit thresholds for $\beta_0$ are included in this test.

**Test 7** A hedge is regarded effective, if and only if a linear regression, which can be executed on fair values, cumulative or period-by-period changes results in a value

\[ \hat{\beta}_1 \in \left[-\frac{4}{5}, \frac{5}{4}\right], \quad \text{and} \quad R - \text{squared} \geq 80\% , \]

and if $\hat{\beta}_0$ is sufficiently small for regression based on changes and close to the value of the initial hedge position for regressions based on fair values.

In Table 4, we compare the results of these three types of regression analysis, and in Figure 6 we illustrate the difference in the data for our example introduced in Figure 1 with scatter plots.

All these types of regression measure offsetting but do not meet the criteria of large numbers. A perfect linear dependency of the data with a value $R^2 = 100\%$, a slope $\beta_1 = -4/5$, and additionally an intercept of zero for market value changes is under the regression test an effective hedge, but the loss of the hedge position is not limited.

The problem of “small numbers” is not avoided as well: When there are nearly no changes in market value of hedged item and hedging instrument the coordinates for the points are all close to each other. So a line approximating these must not necessarily exist, even if the hedge is almost perfectly effective. For the adjusted example of Kalotay and Abreo (2001) we obtained ineffectiveness, the detailed results are summarized in Table 3.

The measurements based on regression analysis as defined in Test 7 are not symmetric but scalable. So only the first and last criterium are fulfilled.

Table 3: Results of the application of linear regression analysis to the adjusted example of Kalotay and Abreo.

<table>
<thead>
<tr>
<th></th>
<th>$R^2$ [%]</th>
<th>(\hat{\beta}_0)</th>
<th>$\hat{\beta}_1$</th>
<th>effective</th>
</tr>
</thead>
<tbody>
<tr>
<td>fair value</td>
<td>94.57</td>
<td>99,998,433.08</td>
<td>-2.59</td>
<td>no</td>
</tr>
<tr>
<td>cumulative</td>
<td>95.01</td>
<td>-1,713.80</td>
<td>-2.65</td>
<td>no</td>
</tr>
<tr>
<td>period-by-period</td>
<td>4.75</td>
<td>149.93</td>
<td>-0.25</td>
<td>no</td>
</tr>
</tbody>
</table>
Table 4: Results of the application of the hedge effectiveness tests based on time series for the hedge data of Figure 1 on page 5. For each time interval 60 data points were created according to the market value development indicated in Figure 1.

<table>
<thead>
<tr>
<th></th>
<th>Expected to be effective</th>
<th>Linear Regression fair values</th>
<th>Linear Regression cumulative changes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$ [%]</td>
<td>$\beta_0$</td>
<td>$\beta_1$ eff.</td>
</tr>
<tr>
<td>$t_1$</td>
<td>yes</td>
<td>100.00</td>
<td>100.000 -1.00 yes</td>
</tr>
<tr>
<td>$t_2$</td>
<td>yes</td>
<td>100.00</td>
<td>100.000 -1.00 yes</td>
</tr>
<tr>
<td>$t_3$</td>
<td>yes</td>
<td>100.00</td>
<td>100.000 -1.00 yes</td>
</tr>
<tr>
<td>$t_4$</td>
<td>no</td>
<td>96.40</td>
<td>101,884 0.0013   yes</td>
</tr>
<tr>
<td>$t_5$</td>
<td>no</td>
<td>98.97</td>
<td>99,649 0.0013    yes</td>
</tr>
<tr>
<td>$t_6$</td>
<td>no</td>
<td>100.00</td>
<td>100,000 1.00    no</td>
</tr>
<tr>
<td>$t_7$</td>
<td>no</td>
<td>100.00</td>
<td>100,000 1.00    no</td>
</tr>
</tbody>
</table>

4.2.3 Variability-Reduction Measure

Finnerty and Grand (2002) develop a test based on the assumption, to regard a hedge as effective if and only if

$$RVR = 1 - \frac{\sum_{i=1}^{n} (-\hat{\beta}_1 \Delta SG_i + \Delta GG_i)^2}{\sum_{i=1}^{n} \Delta GG_i^2} \geq 80\%,$$

where $\hat{\beta}_1$ denotes the estimate obtained from the regression described above with opposed dependent and independent variables.

For the retrospective test instead of $\hat{\beta}_1$ the “actual hedge ratio the hedger implemented” should be used, which implies for a perfect hedge $\hat{\beta}_1 = -1$.

Then for period-by-period changes the variability-reduction is defined as

$$V R(\Delta GG, \Delta SG) = 1 - \frac{\sum_{i=1}^{n} (\Delta SG_i + \Delta GG_i)^2}{\sum_{i=1}^{n} \Delta GG_i^2}.$$
Test 8 A hedge is effective if the variability-reduction is at least 80%, i.e.
\[ VR \geq 80\% . \]

For evaluation of the criteria we regard the following example: Let a hedge have a constant decrease of 30,000 US $ in the market value of the hedged item and a constant increase of 20,000 US$ for the hedging instrument, i.e. the period-by-period differences are
\[ \Delta GG_i = -30,000 \text{ US$} \quad \text{and} \quad \Delta GG_i = 20,000 \text{ US$} \quad \text{for all dates } i. \]

Offsetting is measured by this test, but the problem of large numbers is not avoided, as in this example after \( n \) periods we obtain a loss in the hedge position of \( n \cdot 10,000 \text{ US$} \), but the variability-reduction has a constant value of \( VR = 88.89\% \). The adjusted example of Kalotay and Abreo (2001) illustrates that the problem of small numbers may occur in this test, as we obtain a value of \( VR = 0.0042\% \), which is obviously lower than 80%.

The symmetry is not fulfilled, as in the above example we obtain
\[ VR(\Delta GG, \Delta SG) = 88.89 \% \quad \text{and} \quad VR(\Delta SG, \Delta GG) = 75.00 \% , \]
which implies the hedge to be effective in the first and ineffective in the second case.

Scalability is fulfilled as one can easily see that a percentage will be canceled in the fraction. So again only the first and the last criteria are met.

4.2.4 Volatility Reduction Measure

One other approach to measure effectiveness is based on the idea of risk reduction, as stated by Hull (2003) “hedge effectiveness can be defined as the proportion of the variance that is eliminated by hedging.”

According to Coughlan, Kolb and Emery (2003) the relative risk reduction is defined as
\[ RRR = 1 - \frac{\text{risk of portfolio}}{\text{risk of underlying}} . \]

As possible risk measures they mention value-at-risk and the variance or volatility of changes in fair value. The latter is used by applying standard deviation of changes in fair value.

Kalotay and Abreo (2001) have developed their test based on 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.” This measurement is detailed in the following test.

Test 9 A hedge is effective, if the volatility reduction measure
\[ VRM = 1 - \frac{\sigma_{\Delta GP}}{\sigma_{\Delta GG}} = 1 - \frac{\sigma_{\Delta GG + \Delta SG}}{\sigma_{\Delta GG}} \]
is part of the interval \([80\%, 125\%]\).
Corresponding to the examples described by Coughlan, Kolb and Emery (2003) we use cumulative changes for the differences.

Coughlan, Kolb and Emery (2003) suggest other thresholds. They regard a hedge as effective, if $RRR \geq 40\%$, as they indicate that “a correlation of $-80\%$ corresponds to a level of risk reduction of approximately $40\%$”.

For evaluation of Test 9 we regard the following example: Let for date $i$ the cumulative differences in the market value of the hedging instrument be $\Delta SG_i = i \cdot 10,000$ US$ and let $\Delta GG_i = -5/4 \cdot \Delta SG_i$ for all dates $i$.

For any period this results in

$$V RM(\Delta GG, \Delta SG) = 80.00\%$$

which implies the hedge to be effective. Obviously this test measures offsetting, but the example illustrates that the maximum loss is not limited. The problem of small numbers may occur as well. Using the expansion of the example of Kalotay and Abreo illustrated in Figure 3, we obtain

$$V RM = 1 - \frac{1,742.14}{2,715.71} = 35.85\%,$$

which results in an ineffective hedge.

Further on this test is not symmetric, as changing $\Delta SG$ and $\Delta GG$ in the above example results in a value of $V RM = 75.00\%$ for all periods, which implies the hedge to be ineffective. As this test is scalable, it meets the first and last criterion.

When implementing this test further non-mathematical management considerations are necessary, as this method for determining effectiveness is subject of United States Patent Application 20020032624.

### 5 Adjusted Hedge Interval

In Section 3, we have formulated criteria which should naturally be met by an effectiveness test. As summarized in Table 5, none of the tests we presented fulfills all of these criteria. According to the geometrical interpretation of the criteria it seem obvious that the effective area has to be similar to the dollar offset ratio except for large and for small numbers: For large numbers it should be parallel to the northwest-southeast diagonal and for small numbers it should broaden the intersection of the two cones.

Suppose $h_1$ and $h_2$ are natural numbers with $h_1 < h_2$. A test which is in medium scale equivalent to the dollar offset ratio based on the interval $[\frac{h_1}{h_2}, \frac{h_2}{h_1}]$ can then be obtained with the following generalized hedge interval:

Let auxiliary functions $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ be defined with

$$f_1(\Delta GG) = \frac{h_2^2 + h_1^2}{2h_1h_2} \Delta GG \quad \text{and}$$

$$f_2(\Delta GG) = \frac{h_2^2 - h_1^2}{2h_1h_2} \sqrt{(\Delta GG)^2 + c}.$$
Table 5: Results of the evaluation of effectiveness test according to the criteria defined in Section 3.1 and Section 3.2. We indicate with the symbol (√), when a criterion is fulfilled only if particular constants are chosen for the measurement.

<table>
<thead>
<tr>
<th>Tests based on two points of data</th>
<th>(i) Offsetting</th>
<th>(ii) Large Numbers</th>
<th>(iii) Small Numbers</th>
<th>(iv) Symmetry</th>
<th>(v) Scalability</th>
<th>(vi) Smooth Transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Dollar offset ratio</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>2 Intuitive response</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>√</td>
<td>(√)</td>
<td>–</td>
</tr>
<tr>
<td>3 Lipp modulated dollar offset</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>√</td>
<td>(√)</td>
<td>–</td>
</tr>
<tr>
<td>4 Schleifer-Lipp modulated offset</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>√</td>
<td>(√)</td>
<td>–</td>
</tr>
<tr>
<td>5 Gürtler effectiveness test</td>
<td>–</td>
<td>√</td>
<td>√</td>
<td>(√)</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>6 Hedge interval</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>√</td>
<td>(√)</td>
<td>–</td>
</tr>
<tr>
<td>10 Adjusted hedge interval</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tests based on time series of data</th>
<th>(i) Linear regression (fair value)</th>
<th>(ii) Linear reg. (cumulative changes)</th>
<th>(iii) Linear reg. (period-by-period)</th>
<th>(iv) Variability-reduction measure</th>
<th>(v) Volatility reduction measure</th>
<th>(vi) Adjusted hedge interval based on 100 % compliance level</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 Linear regression (fair value)</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>8 Variability-reduction measure</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>9 Volatility reduction measure</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Adjusted hedge interval</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

We obtained the simple formula for our hedge interval (Test 6) from the geometrical interpretation by asserting

\[ f = -f_1 + f_2 \quad \text{and} \quad \bar{f} = -f_1 - f_2. \]

This is equivalent to a hedge being effective, if

\[ \frac{2h_1h_2 \Delta SG + (h_1^2 + h_2^2) \Delta GG}{\sqrt{\Delta GG^2 + c}} \leq h_2^2 - h_1^2. \]

For example, for the underlying interval \([\frac{9}{10}, \frac{10}{9}]\) this implies to regard a hedge effective if

\[ \frac{180 \Delta SG + 181 \Delta GG}{\sqrt{\Delta GG^2 + c}} \leq 19. \]

To adjust this interval to limit maximum loss or gain of the hedging position as part of criterion (i), we want to let the bounding functions of the effective area be parallel to the northwest-southeast diagonal. The maximum distance of this line should be a fixed percentage \(p\) of \(\sqrt{2} GP_0\) to ensure scalability of the test. In our examples we used a value of \(p = 25\%\). So for "large" values of \(\Delta SG\) we regard the bounding functions \(g(\Delta GG) = -\Delta GG - p GP_0\) and \(\bar{g}(\Delta GG) = -\Delta GG + p GP_0\) instead of \(f\) and \(\bar{f}\).
Let auxiliary terms \(d\) and \(e\) be defined with 
\[
ed = p GP_0 \quad \text{and} \quad e = \sqrt{d^2 h_2^2 h_1^4 + 2 d^2 h_2^2 h_1^4 + h_2^4 d^2 h_1^4 - h_1^2 h_2 c + 2 h_1^2 h_2^3 c - h_1 h_2^2 c}.
\]

Then the function \(f\) and \(g\) intersect at \(\Delta GG = x_1\) and \(\Delta GG = x_2\), where
\[
x_1 = \frac{-4 d h_2 h_1^2 + 4 h_2^2 d h_1 + 4 e}{8 h_1 h_2 (h_1 - h_2)}
\]
and
\[
x_2 = \frac{-4 d h_2 h_1^2 + 4 h_2^2 d h_1 - 4 e}{8 h_1 h_2 (h_1 - h_2)}.
\]

Analogously the function \(f\) and \(g\) intersect at \(\Delta GG = -x_2\) and \(\Delta GG = -x_1\).

To satisfy criterion (vi) we use these points as turning points. Let \(X_L\) and \(X_U\) be intervals defined with 
\[
X_L = [-x_2, x_2] \quad \text{and} \quad X_U = [-x_1, -x_1].
\]

Then, more precisely, we assume a hedge to be effective if and only if
\[
f(\Delta GG) \leq \Delta SG \leq f(\Delta GG)
\]
where
\[
f(\Delta GG) = \begin{cases} 
-f_1(\Delta GG) + f_2(\Delta GG) & \text{for } \Delta GG \in X_L \\
-\Delta GG - p GP_0 & \text{else}
\end{cases}
\]
and
\[
f(\Delta GG) = \begin{cases} 
-f_1(\Delta GG) - f_2(\Delta GG) & \text{for } \Delta GG \in X_U \\
-\Delta GG + p GP_0 & \text{else}.
\end{cases}
\]

This can simply and equivalently be formulated by the following criterion.

**Test 10 (AHCI – adjusted Hedge Interval)** Let \(h_1\) and \(h_2\) be natural numbers with \(h_1 \leq h_2\) representing an underlying dollar offset interval \([h_1, h_2]\). Let \(c\) be a fixed percentage of \(GP_0^2\), i.e. \(c = c_{GP_0^2}\).

A hedge is effective if and only if \(|GP_t - GP_0| \leq p GP_0\) and
\[
\left|\frac{2h_1 h_2 \Delta SG + (h_1^2 + h_2^2) \Delta GG}{\sqrt{\Delta GG^2 + c}}\right| \leq h_2^2 - h_1^2.
\]

The effective area is illustrated in Figure 7 with a value of \(p = 25\%\) and of \(c = 10^{-7} \cdot GP_0^2 = 1000\) for large scale. It is compared with different underlying dollar offset intervals in Figure 8 on page 30.

For the example introduced in Figure 1 we obtain the expected results when regarding just two points of time and when applying a 100 \% compliance level for the time series. For these we provide in Table 4 the maximum value of the fraction. For the periods \(t_3 - t_4, t_4 - t_5\) and \(t_7 - t_8\) the additional criterion \(|GP_t - GP_0| \leq p GP_0\) implies ineffectiveness as well. For the adjusted example of Kalotay and Abreo we obtain a maximum value of 7.5378, which implies the hedge to be effective as expected.
Figure 7: Illustration of the adjusted hedge interval, the points where the bounding functions $f$ and $\overline{f}$ change from determining cones to parallels to the northwest-southeast diagonal are marked with an ‘*’.
Figure 8: Illustration of the adjusted hedge interval for an underlying dollar offset interval of $[\frac{4}{5}, \frac{5}{4}]$ and $[\frac{9}{10}, \frac{10}{9}]$. 
Theorem 1  The adjusted hedge interval test meets criteria (i) to (vi).

Proof: Criteria (i) is fulfilled as the adjusted hedge interval approximates the dollar offset ratio in a medium scale, i.e. for most expected market developments. For large values the additional inequality \( |GP_t - GP_0| \leq p GP_0 \) limits the maximum possible gain or loss and for small values the area of effectiveness is enlarged.

The symmetry is guaranteed, as one can easily show that
\[
    f(x) = f^{-1}(x) \quad \text{and} \quad f(-x) = -f(x) \quad \text{for all} \quad x \in \mathbb{R}.
\]

For large values of \( \Delta GG \) the scalability is obvious and for small values the factor \( \alpha \) can be canceled in the fraction
\[
    \frac{2h_1h_2 \alpha \Delta SG + (h_2^2 + h_2^2) \alpha \Delta GG}{\sqrt{(\alpha \Delta GG)^2 + \alpha^2 GP_0^2}}.
\]

The last criterium is a direct consequence of the construction of the functions \( f \) and \( \overline{f} \), guaranteeing that they are continuous.

In summary, according to the criteria defined in Section 3 and to get a test as simple as possible, the adjusted hedge interval approach presents a geometrically natural enhancement of the dollar offset ratio. Both problems of the dollar offset ratio concerning small and large numbers are avoided. Currently no limitations are known.

FAS 133 §230 (Appendix “Background Information and Basis for Conclusions”) contains the following statement: “Because hedge accounting is elective and relies on management’s intent, it should be limited to transactions that meet reasonable criteria.”

While there appears to be no definitive answer to the question of what is or is not reasonable, the measurable criteria proposed in this paper may form at least parts of one.

References


Market values used for the adjusted example of Kalotay and Abreo.
Market values used for the hedged item of the example illustrated in Figure 1.