TechHackers'
QuantNews

September 14, 2000

Welcome to the sixth issue of QuantNews.

Effectiveness testing for FAS 133

by Andrew Kalotay and Leslie Abreo, Andrew Kalotay Associates

Background Corporations routinely employ derivative products, such as interest rate swaps, to hedge business risk. For example, a financial institution with floating rate assets might synthetically convert its fixed coupon bond obligation into a floating rate obligation by entering into an appropriate interest rate swap.

But corporations can also use derivative products for speculation, with unpredictable and possibly disastrous consequences, such as in the notorious case of the Procter & Gamble fiasco. The problem for shareholders is that this kind of exposure typically gets buried in footnotes to the corporation's financial statements even if it has the potential (realized in the case of P&G) to have a material effect on reported earnings.

In an attempt to make a corporation's exposure to its derivative positions more transparent, the Financial Accounting Standards Board (FASB) has issued Statement 133 (later amended by Statement 138). This statement requires full disclosure of the derivative products employed by the corporation. In addition, derivatives must be marked to the market and the changes in their market value reported in the income statement. However, to mitigate the resulting volatility of earnings (anathema to Wall Street's equity analysts), FAS 133 allows the corporation to mark the hedged liability or asset to market as well. In order to qualify for this favorable treatment, it must first be shown that the hedge is highly effective.

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

The basic motivation for hedging is to eliminate *unpredictability* due to market changes. In the real world of imperfect hedges, practitioners tend to be concerned only with large deviations relative to the *scale of the hedged item*.

In principle, a hedge is effective if the price movements of the hedged item and the hedging vehicle roughly offset each other so the net change of the *package* is negligible relative to the hedged item.

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

FASB Guidelines The FASB has suggested two approaches to effectiveness testing. One is the so-called "80/120 Rule". A hedge is deemed effective if the ratio of the change in value of the derivative to that of the hedged item is between 80% and 120%.

An unintended and unfortunate consequence of this test is that during periods of market stability virtually any hedge is likely to fail. Consider, for example, a $100 million bond hedged with an interest rate swap. A $10,000 change in the value of the bond and a $4,000 opposite change in the value of the swap results in a ratio 0.4. Hence the hedge will be deemed *ineffective* under the 80/120 Rule, even though the net change of $6,000 is a miniscule 0.006% of the face amount. The fundamental shortcoming of the 80/120 Rule is that it takes no account of scale. As a consequence, the test is susceptible to "false positives" in the absence of meaningful information.

The second approach suggested by the FASB is loosely based on the correlation of the changes in value of the hedged item and that of the derivative. Roughly speaking, a hedge is deemed effective if the R-squared[1] of the regression line explaining the data is sufficiently high, say 80%.

But a high R-squared alone is not a reliable indicator of effectiveness. In addition, the changes in value should be roughly offsetting, i.e. the *slope* of the regression line should be close to 1, a consideration not explicitly referred to in FAS133. A related issue not addressed is the intercept of the regression line. Should the intercept be constrained to 0 or should the "best fit" regression determine it? Constraints on the regression process, it should be noted, lower the R-squared, increasing the likelihood that the hedge will fail the effectiveness test.

As indicated, FASB has provided only broad guidelines to effectiveness testing. In the absence of definitive guidance, corporations are expected to devise, apply and defend their own tests. In that vein, we propose the following approach.

The Volatility Reduction Measure (VRM)[2] Traders and portfolio managers, whose compensation is affected by the actual performance of hedges, judge the effectiveness of a hedge in terms of *volatility reduction*. The volatility of the item being hedged in the *absence* of a hedge is the obvious point of reference against which this reduction should be measured. In contrast, the FASB guidelines fo-

[1] In an unconstrained bivariate regression, R-squared is the square of the correlation between the two variables.

[2] The Volatility Reduction Measure (patent pending) for hedge effectiveness testing was invented by Andrew Kalotay Associates, Inc.

cus on *pair-wise* (date-by-date) comparison of changes in value, rather than on overall volatility with and without the hedge. The VRM method described below captures the significance of hedging to practitioners while retaining the basic intent of FASB.

The following example graphically conveys the essence of the VRM approach. Figure 1 shows the frequency distribution of 1,406 quarterly changes in value of a $100 million 8% 10-year LIBOR-credit corporate bond, as well as the analogous information for the combination of the bond and an 8% LIBOR-based 9-year swap of like notional amount. The standard deviation (volatility) of the bond alone is $3.651, while the bond / swap combination has a volatility of only $0.275 million.

More formally, the volatility reduction measure is defined as:

$$\begin{aligned} VRM &= 1 - \text{stdev (hedge package)} \\ &\quad /\text{stdev (item being hedged)} \end{aligned}$$

In the above example, the volatility of the hedged item was reduced by 92.46% (1 − 0.275MM/3.651MM).

Optimal Hedge Ratio and VRM The VRM methodology can also be used to determine the size the derivative position to use as a hedge. Consider a hedged item and a derivative with known standard deviations and correlation. A straightforward application of calculus can be used to show that the standard deviation of the hedge package (item + derivative) is minimized, and therefore VRM is maximized, when the derivative position is scaled so that

$$\begin{aligned} &\text{stdev (hedge)} = \\ &\qquad -\rho \cdot \text{stdev (item being hedged)} \end{aligned}$$

where ρ is the correlation between the hedged item and the derivative.

As an illustration, say a $100MM bond issue is to be hedged with a swap—assumed for now to be $100MM notional as well. Based on the following six data points (for simplicity),

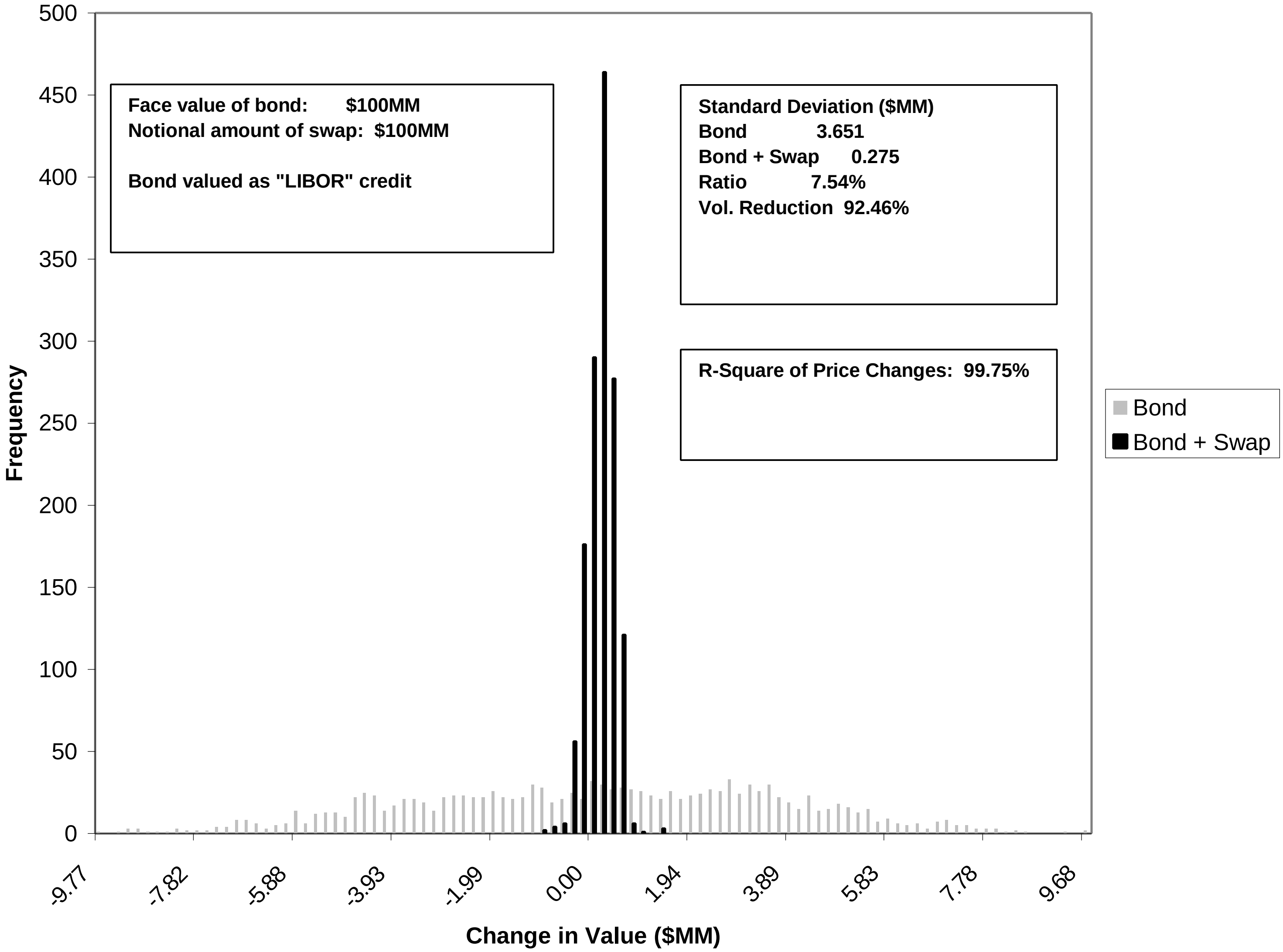
Change in value (swap, $ millions)	Change in Value (bond, $ millions)
5.0	-6.0
2.0	-2.2
-8.0	7.0
9.0	8.0
11.0	-15.0
-6.0	8.0

we see that the swap standard deviation is $7.78M while the bond standard deviation is $8.93M; their correlation is −0.979. The maximum volatility reduction is achieved when the notional of the swap is $-(-0.979) \cdot \$8.93M \cdot (\$100MM/\$7.78M) = \$112MM$.

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

The VRM approach is superior in its simplicity. It is rigorous, defensible and reasonable. Standard deviation is the accepted measure of volatility. When expressed in dollar terms, standard deviation reflects actual business risk, and unlike arcane statistics such as R-squared, it is familiar to higher management.

Figure 1: Volatility Reduction Measure
10 Yr 8% Bond Hedged with 9 Yr 8% LIBOR Swap
Distribution of 1406 Quarterly Changes in Value Based on Rates from 5/94 to 12/99



Last but not least, the VRM method is consistent with Value at Risk (VaR), a widely used measure in risk management.

Automated Customer Decisioning: Detecting Money Laundering and other Frauds

by David Porter, Unisys Consulting

Making the Right Decision Everybody talks about "the right information to the right person at the right time." But it doesn't pass the "so what?" test. These days people at the front line are typically deluged with information: trends go up, trends go down, there are clusterings here and segments there, but what does it all mean, and more importantly, what should we do with it? A much better approach is to make the right decision at the right time.

This is particularly important when banks are dealing with customers. For those with small, localised customer bases, dedicated account managers can provide a degree of customised personal service. However for a typical high street bank with millions of customers and the desire to embrace 24 $\times$ 7 e-business it is not economical or indeed feasible to employ armies of people to carry out rapid, high quality, consistent decisioning. Automation of decisioning is therefore an inevitable development in the banking world, in other words, computers playing a key role in helping human decision makers make sense of all the data and the key inferences that need to be made in order to retain and grow profitable customers.

Customer Decisioning Pitfalls Imagine Ms. X, a long-standing customer of a large bank that offers a range of *bancassurance* products and services. The bank has fed Ms. X's profile through the mixer—a combination of data warehouses and various profitability, propensity and risk models—decided that she is a high value, profitable customer and has designed a range of customised offerings to meet her needs such as preferential loan rates and a more relaxed credit policy.

If we consider two different scenarios in which Ms. X might apply for a personal loan, the customer decisioning process can be seen at work. In a situation where Ms. X's requires a quick loan to fix an emergency, she is probably not sensitive to the interest rate, being more concerned about getting money in a hurry. Here there is no benefit to the bank giving away money by offering a preferential rate, when the business can be secured by simply loosening the credit policy. Alternatively, Ms. X might want to buy a car, so she might start trawling the Internet to get a shrewd deal on a personal loan. Because she is prepared to shop around she is likely to be price-sensitive. For the bank to win her business it may be worth surrendering a small element of profit by offering a preferential rate.

As these examples show, customers may have radically different needs at different times and from a bank's point of view the challenge is to make the right decision at the right time. The aim is to move from product-oriented to relationship-oriented interaction, or "selling more our services and selling our services for more." The relationship Ms. X has with the bank, and the need to anticipate and satisfy her needs, is therefore paramount.

However things are not always as clear-cut as this. Consider the scenario where, several months later, an official contacts a bank saying: "Ms. X has just been arrested on fraud and money laundering charges. You have been handling her financial transactions…" Someone the Bank once viewed as a valued customer is now quite the opposite.

There is no doubt that the vast amounts of sheer change caused by the growth of e-commerce is being exploited by the criminal element. In the plastic card industry the increase in e-commerce transactions has led to a resurgence in "card-not-present" fraud, latest industry surveys revealing that this form of fraud is growing at double the overall plastic fraud growth. No-one knows exactly how much criminal money is laundered annually. The US Drug Enforcement Administration has talked about a global figure of close to $1 trillion annually as a direct result of international drug trafficking. The last available figures for the UK economy demonstrated a conservative estimate of approximately £2.4 billion of criminal money annually passing through the City of London. The simple fact is that not all customers, or e-business scenarios, are valuable.

Balanced Decisioning Automating decision making has steadily evolved over last ten years as evidenced by the increasing take-up of credit scoring, needs-based marketing and fraud detection applications. However the ad-hoc nature of this process has left a legacy of silo applications or *decision islands* which have either supplemented or replaced specific business experts in areas such as credit, marketing, pricing and financial valuation.

Forward-looking financial institutions are realising that customer relationship management requires knowledge and understanding of the "full picture" that is beyond the capabilities of any one of these individual experts. The field is just too large and complex to be modelled within a single system, if only because the maintenance of such a system would be extremely difficult to manage. Furthermore, banks are also realising that although e-business poses tremendous opportunities for exploiting customer relationship management technologies this must be balanced against managing the risk of exposure to individuals engaged in potentially harmful activities. In other words, "the e-business we don't want."

The art of automating the customer decisioning process lies in the ability to collate individual talents and opinions into a single process, rather like a committee of specialists brought together to solve a complex problem. This structure is based on the intuitive understanding that the participation of several experts can improve the quality of the outcome in complex situations. Although the participants may be working towards a common overall goal, e.g. maximising the net present value from a customer, it is natural for different experts to have different opinions on how best to achieve it.

In IT and business terms this has led to the concept of *balanced decisioning* in which a high level horizontal mechanism arbitrates across different vertical experts (or decision making systems) within the organisation in order to arrive at a rational, profit-maximising decision. Such experts will typically include areas such as marketing, operations, financial valuation, pricing, risk and compliance.

The aim is to resolve conflict, facilitate co-ordination between the experts and achieve a level of optimisation based on overall customer value rather than independent product criteria. The experts themselves retain their original roles as functional specialists who make recommendations using the best available customer data. The development of an arbitration mechanism will evaluate the overall value of a customer rather than being influenced by an individual agenda.

Modelling is Everything So how do we build this kind of system? At the heart of this kind of decisioning system is a body of knowledge that will form the basis for making optimised customer decisions. Such a knowledge base must be ultimately encoded and executed on a computer. There is no reason why conventional programming languages such as C++ cannot be used for this purpose. However, before we start coding, we need additional techniques to cater for the elicitation and analysis of tacit knowledge-based thought processes. In other words, bridge the gap between human beings and adding machines.

Modelling is the key to this and the first step is to understand that human decisioning is generic in nature. In other words, a doctor and a car mechanic both perform the same generic decisioning task known as *systematic diagnosis*, even though they differ in terms of their specific domain knowledge content. This and other generic decisioning tasks can be arranged in a *decisioning family tree* that can be divided at the highest level into analytic and synthetic tasks.

An example of an analytic task is heuristic classification, a sub-class of diagnosis. The most powerful and most general inferences are those that are made at a high level of abstraction, since typically these lead to a shorter solution path. Heuristic classification seeks to take advantage of this by applying high level heuristics (or "rules of thumb") to map from a problem space to a solution space. Systematic diagnosis is another sub-class of diagnosis. Whereas heuristic classification uses non-systematic rules of thumb, systematic diagnosis is used when there is sufficient structure in the domain knowledge for us to be able to represent a model of the domain object being diagnosed. Diagnosis proceeds by systematically moving through the structured model as information about the fault becomes available.

Decisioning Road Maps For each leaf node task in the decisioning family tree it is possible to define a generic, skeletal model (or road map) that describes how the decisioning task is achieved in terms of basic inferences. The notation used is intended to be more formal than unstructured linguistic statements about how decisioning is achieved but more abstract than a technical computer programming language.

The model therefore acts as a blueprint for driving the decisioning analysis process, providing vital clues about what sort of knowledge is used in a particular decisioning process and how that knowledge is applied. This approach facilitates decisioning system development since it promotes a top down approach involving the reuse and refinement of existing "library" material. In the event that a model is not available to fit a particular decisioning application then a new model can be constructed in a bottom up manner using primitive elements.

Four Layers of Knowledge The decisioning inference structure described above is just one layer of a four layered knowledge modelling approach. Key to this approach is the principle that different types of knowledge—domain, inference, task and strategy—can be identified, each knowledge type playing a particular role in the decisioning process and each having a different structuring principle.

A knowledge model fulfills two key roles. Firstly, it formally documents a knowledge-based or decisioning process so it can be readily understood by both business people during the analysis phase and those ultimately responsible for providing operational support. Secondly, it acts as an architectural blueprint for an automated software solution for supporting or even performing the process in question.

The domain layer contains knowledge describing a *declarative theory* of the domain in the form of static concepts, relations, structures and rules. The inference layer contains an inference structure that makes explicit the different types of inference that can be made on the domain layer. Inference knowledge therefore uses and reasons about domain knowledge. It contains the key rules that underpin a particular decisioning application, or the system's basic "competency".

The task layer contains knowledge that describes how the rules in the inference layer might be used. It acts as a kind of "procedural overlay" that shows the sequence of invocation of each basic inference. Finally, the strategy layer controls and coordinates the different tasks defined in the task layer. The knowledge in this layer is concerned with higher level strategic issues, enabling the system to "reflect" about its current approach and dynamically re-plan its overall problem solving strategy in the event of impasses, deadlocks, etc. A complex, dynamic strategy layer may even be a decisioning system in its own right, for example a planning system whose output alters dynamically according to the performance of the cooperating tasks (experts) involved in the decisioning process. Now we have a good foundation for the "balanced decisioning" architecture discussed above.

Let's Get Formal If the conceptual modelling approach is not formal enough for your taste there is no reason why we cannot go one step further and embrace more formal methods. The different layers of a conceptual knowledge model can be directly mapped into an equivalent structure of logical theories in a formal model.

For example, we can use logic to represent declarative domain knowledge, i.e. facts and rules that are true in the domain and are represented independently from how they are to be used. One approach is to use first order predicate calculus with two extensions. Firstly theories can be expressed in a modular way, i.e. axioms can be divided into sub-theories which can then be combined using simple meta-theoretic operators. This enables formalised domain knowledge to be structured in the same way as its counterpart in the conceptual model, i.e. the construction of different "viewpoints" on the basic domain theory. Secondly, an order sorted language is used in which all variables and constants have associated sorts (types) organised in a sub-sort hierarchy. Predicates and functions are also typed according to their arguments and results respectively.

In the inference layer, primitive inferences and their input/outputs can be mapped onto equivalent formal constructs. Given that the inference knowledge in the conceptual model is a theory about the use of the domain knowledge, formal inference knowledge is defined as a meta-layer of the domain knowledge. Meta-logic is used to implement this metarelation, i.e. the use of meta-theories (expressed using order-sorted first order predicate calculus) whose terms refer to an object theory.

The purpose of task knowledge is to impose procedural control over the declarative inference steps defined in the inference knowledge. In order to express this knowledge we can use quantified dynamic logic. This is an extension of first order logic specifically designed for reasoning about the properties of programs (i.e. state and sequence). The language provides predicate logic, atomic programs and syntactic constructions which can be used to express traditional programming language constructs such as sequence, iteration, selection, etc. Formal task knowledge therefore uses the declarative primitive inferences as programs that can be executed.

Conclusions The availability of robust knowledge engineering techniques such as those described above means that automated decisioning is now a relatively mature discipline. Indeed many of the supporting tools and techniques date back many decades, and in some cases many centuries. Regardless of the underlying approach used, be it statistical, rule-based or some form of automated self-learning (a hybrid approach often being the most sensible option), the aim is to create models of behaviour against which transactions can be analysed and the appropriate decisions made.

Over the past decade automated decisioning has tended to polarise into two key application areas: finding the "good guys", i.e. profitable, or potentially profitable, customers in order to increase profits; finding the "bad guys", i.e. bad debtors, abusers, fraudsters and money launderers in order to reduce losses and protect reputation. However, there is now a trend towards coordinating the activities of formerly disparate decisioning systems. In both cases we are looking for either profitable needles or bad needles in a mass market haystack. At a generic level there are valuable lessons to be learnt about how we should best find the needles or cut away the straw. The fact that the phrase "know your customer" is being uttered by both marketing and compliance officers indicates that there are advantages to be gained by better coordination.

Indeed, there is no reason why the "identify, segment, target, predict, influence" approach used by marketeers cannot be adopted by those involved in fraud and compliance. There is also potential for extending customer profiling so that riskier customers are better defined (the boundaries between bad debt and deliberate fraud being very blurred) and re-appraising existing profiling and segmentation schemes, in particular the definition of "valuable".

As the development of e-business continues to advance, this important principle of balance is certain to feature strongly in the "second wave" of customer relationship management solutions. Three effective design lessons are already emerging from this field of work. First, remember that users have their day jobs to do. Second, be innovative rather than pioneering. Third, above all, keep it simple.

In Conclusion...

We hope you've enjoyed issue 6 of Quant*News*. Comments, submissions, and other requests should be sent to your editor at steven.janowsky@unisys.com.