CHAPTER 8

Model Risk

Any book on financial risk management needs to address the subject of *model risk*, the risk that theoretical models used in pricing, trading, hedging, and estimating risk will turn out to produce misleading results. This book, which emphasizes quantitative reasoning in risk management, pays particularly close attention to how models can be used and misused in the risk management process.

Since the publication of the first edition of this book, the financial risk management focus on model risk has intensified. In the wake of the 2007–2008 crisis, as we discuss in Sections 5.1 and 5.2.5.3, there have been accusations that model failure was one of the root causes of the meltdown. When a widely discussed article has the title "Recipe for Disaster: The Formula That Killed Wall Street" (Salmon 2009), it is clear that model risk needs to be addressed with a sense of urgency.

Fortunately, in addition to this sense of crisis surrounding model risk, the past several years have witnessed greater attention to analysis of how model risk can be controlled. Concise, excellent articles by Derman (2001) and Rebonato (2003) are now recognized as touchstones for the analysis of model risk. Morini (2011) is the first thorough book-length treatment of model risk. The Federal Reserve and Office of the Comptroller of the Currency joint document for "Supervisory Guidance on Model Risk Management," which I will reference as FRB (2011), and the Basel Committee on Banking Supervision's "Supervisory Guidance for Assessing Banks' Financial Instrument Fair Value Practices," which I will reference as Basel (2009b), provide regulatory responses to the lesson of the 2008 events for model risk. I find the joint Federal Reserve/Comptroller of the Currency document to be particularly thorough and persuasive in its analysis of the many aspects of model risk.

This chapter begins, in Section 8.1, with an overview focusing on the variety of opinions that have been expressed about the importance of models, or their unimportance, in managing financial risk. Section 8.2 examines the procedures that ought to be used for risk evaluation and control for models of all types. The following three sections give a more detailed analysis of model review standards, distinguishing among three types of models: those used for valuation and risk measurement of liquid instruments in Section 8.3, those used for valuation and risk measurement of illiquid instruments in Section 8.4, and those used for making trading decisions in Section 8.5.

8.1 HOW IMPORTANT IS MODEL RISK?

When examining model risk, one immediately encounters a very wide range of views on the role that models can play in controlling risk and creating new risks. These vary all the way from viewing model error as the primary cause of financial risk to viewing models as largely irrelevant to risk.

The view that models are largely irrelevant to risk can often be encountered among traders who view models as just convenient mathematical shorthand with no real meaning. All that really matters are the prices the shorthand stands for. A good example is the yield of bonds as calculated by Securities Industry Association standards. This includes many detailed calculations that have no theoretical justification, but can only be explained historically (for example, some parts of the calculation use linear approximations, which made sense before calculations were done on computers). No one would claim that this yield has a precise meaning—you don't necessarily prefer owning a bond yielding 7 percent to one yielding 6.90 percent. However, you can translate between yield and precise price given the industry standard rules. It is convenient shorthand to convey approximate values. The degree to which these calculations give misleading yields hurts intuitive understanding, but does not result in mispricing.

Those who view models as playing no real role in pricing and risk management view almost all models used in financial firms as playing a similar role to that of bond yield calculation. A typical claim would be that the Black-Scholes option model, probably the model most frequently used in the financial industry, is just a mathematical convenience that provides shorthand for quoting options prices as implied volatilities rather than as cash prices. In this view, implied volatilities are an attractive way of providing quotations, both because of common usage and because they provide more intuitive comparisons than a cash price, but they should not be regarded as having any meaning beyond representing the price that they translate to using the Black-Scholes formula.

If this viewpoint is correct, models would play an extremely minimal role in controlling risk, and model testing would consist of little more than rote checking to see if industry-standard formulas have been properly implemented. However, this extreme a view cannot explain all the ways in which trading firms use models such as Black-Scholes. The valuation of unquoted options is derived by interpolating the implied volatilities of quoted options. The Black-Scholes model is used to translate prices to implied volatilities for the quoted options and implied volatilities to prices for the unquoted options. The risk reports of position exposures use the Black-Scholes model to compute the expected impact of changes in underlying prices on option prices. Scenario analyses presented to senior management quantify the impact of changes to the implied volatility surface. For more details, see Chapter 11 on managing vanilla options risk. This behavior is inconsistent with a claim that the model is being used purely to provide convenient terminology. By contrast, the industry standard bond yield formulas are not used in comparable calculations—interpolations and risk reports are based on a more sophisticated model of separately discounting the individual cash flows that constitute a bond, with a different yield applied to each cash flow. In this computation, none of the linear approximations of the industry standard formulas are utilized. For more details on these calculations, see Chapter 10 on managing forward risk.

The view that models are the primary cause of financial risk is often encountered in articles describing major trading losses, which are frequently ascribed to the firm having the wrong model. What is often unclear in these claims is whether "having the wrong model" just means making incorrect forecasts about the future direction of market prices or if it means misleading the firm's traders and managers about the nature of positions being taken. A good illustration is the discussion in Section 4.2.1 of whether the reliance by Long-Term Capital Management (LTCM) on models should be viewed as a primary cause of the collapse of the fund. And after the 2007–2008 crisis in mortgage collateralized debt obligations (CDOs), one began to encounter claims such as "[David] Li's Gaussian copula formula will go down in history as instrumental in causing the unfathomable losses that brought the world financial system to its knees" (Salmon 2009).

Of course, once products start encountering losses, modelers who had been promoting a view of the importance of models may now wish to take the opposing view. Morini (2011, Preface) quotes modelers, speaking after the crisis, telling him "Models were not a problem. The problem was in the data and the parameters! The problem was in the application!" Morini's response is that "Models in finance are tools to quantify prices or risks. This includes mathematical relations, a way to use data or judgment to compute the parameters, and indications on how to apply them to practical issues.

Only by taking all these things together can we talk of 'a model.' Modellers should stay away from the temptation to reduce models to a set of mathematical functions that can be thought of separately from the way they are specified and from the way they are applied. If this were the case, models would really be only blank mathematical boxes and people would be right to consider them useless, when not outright dangerous." I would add that any modelers who want to separate their work from choices on data or parameters are basically saying that they are programmers. There's nothing wrong with being a programmer—it's a highly demanding profession. But with rare exceptions (a few people who are able to pioneer an extraordinary speedup of existing calculations), programmers are not compensated at the level modelers are and do not have the degree of influence in making decisions about innovations in products that modelers do.

In the final analysis, whether model builders take the responsibility or the traders and risk managers who use them take the responsibility, models play a key role in managing risk and we must develop clear guidelines to see that the role they play is to clarify issues rather than to obscure them. This is the task to which we now turn.

8.2 MODEL RISK EVALUATION AND CONTROL

In this section, we look at those procedures that ought to be used for risk evaluation and control for all types of models—those used for making trading decisions as well as those used for valuation and risk measurement. In Section 8.2.1, we discuss the scope of model review and in 8.2.2 the proper roles and responsibilities that need to be established around model review and control. In Section 8.2.3, we look at those procedures that check whether the model selected has been correctly implemented—whether the model actually performs as specified; Morini (2011) calls this *model verification*. In Sections 8.2.4 and 8.2.5, we examine two particularly important pieces of model verification, the verification that contractual arrangements have been correctly specified in the model and the evaluation of approximations. In Section 8.2.6, we turn to procedures that check whether the model selected is appropriate for the product or trading strategy being modeled; Morini (2011) calls this *model validation*.

The procedures in Sections 8.2.3, 8.2.4, 8.2.5, and 8.2.6 are primarily designed for the initial evaluation of models leading up to the decision whether the model should be approved for use, and what restrictions, if any, should be placed on its use. In 8.2.7 and 8.2.8, we look at those aspects of model evaluation and control that should take place continuously or periodically during the life of the model's use to see if any new information is available to change the initial conclusions about the model approval or to suggest model modification or replacement.

8.2.1 Scope of Model Review and Control

The first point that needs to be established is what determines that something is a model that requires review and control. FRB (2011, Section III) casts the net very wide, stating that "For the purposes of this document, the term *model* refers to a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into quantitative estimates. . . . Models meeting this definition might be used for analyzing business strategies, informing business decisions, identifying and measuring risks, valuing exposures, instruments or positions, conducting stress testing, assessing adequacy of capital, managing client assets, measuring compliance with internal limits, maintaining the formal control apparatus of the bank, or meeting financial or regulatory reporting requirements and issuing public disclosures."

It is important that a definition this broad be used. A computation may be made by a simple formula in a spreadsheet and still give rise to as great a danger of incorrect estimation as a computation requiring a complex mathematical derivation and a supercomputer churning away for hours to produce the result. Simply averaging observed two- and three-year interest rates to obtain a two-and-a-half-year interest rate already entails an assumption that requires review and control (we'll discuss this example further in Section 8.2.6.1). The mental image many of us have of a model as a complex piece of mathematics and computer engineering can create blinders when we are looking for potential sources of model risk.

A second point made in FRB (2011, Section V) is that "Vendor products should nevertheless be incorporated into a bank's broader model risk management framework following the same principle as applied to in-house models, although the process may be somewhat modified." Whether a model has been created in-house or by a vendor, the consequences of the model being incorrect still affect the profit and loss (P&L) of the firm using the model, so there should be no variation in the standards applied for model review and control. The Federal Reserve goes on to point out the challenges of reviewing vendor models since they "may not allow full access to computer coding and implementation detail" and there is a need for "contingency plans for instances when the vendor model is no longer available or cannot be supported by the vendor." The model review procedures of this chapter can be used for vendor models, but I have encountered instances where a vendor model is so opaque that I have needed to insist that it be replaced by an in-house model or by another vendor model that permitted more transparency.

After establishing the scope of the definition of a model, the next step is to agree on what needs to be included in a review. Some key points from FRB (2011, Section III):

- "Models are of necessity simplified representations of real-world relationships and so can never be perfect."
- As a result, model use invariably results in model risk, which can be defined as "financial loss, poor business and strategic decision making, or damage to a bank's reputation" based on "incorrect or misused model outputs and reports."
- Model risk can result from either fundamental errors in the model or inappropriate use of a model, particularly the use of a model outside the environment for which it was designed.
- "Model risk should be managed like other types of risk. Banks should identify the source of risk and assess the magnitude. Model risk increases with greater model complexity, higher uncertainty about inputs and assumptions, broader use, and larger potential impact." The intensity and rigor of model reviews need to be matched to the degree of model risk identified.
- Model risk cannot be eliminated, so it needs to be controlled through limits on model use, monitoring of model performance, adjusting or revising models over time, and informed conservatism in inputs, design, and outputs. But while conservatism "can be an effective tool" it cannot be "an excuse to avoid improving models."

8.2.2 Roles and Responsibilities for Model Review and Control

I have been involved with the design and approval of several firmwide model review policies. In every case, I have insisted on a prominent statement that "Risk management serves as a second set of eyes for model review." This means that the business unit that develops and utilizes the model has first responsibility for reviewing the model and assessing its risks. The role of the risk management function is very important in ensuring that an independent unit without insider incentives reviews the model and in creating a uniform model review environment throughout the firm. But the knowledge that an independent review will be performed by risk management cannot be used by the business unit as an excuse for not performing its own thorough review. Having two sets of eyes reviewing the model is important both for providing an extra layer of security and in obtaining the benefit of insider product expertise to complement outsider independence and model review process expertise.

FRB (2011) supports this viewpoint on business unit responsibility. In Section VI it states: "Business units are generally responsible for the model risk associated with their business strategies. The role of model owner involves ultimate accountability for model use and performance. . . . Model owners should be responsible for ensuring that models are properly developed, implemented, and used . . . [and] have undergone appropriate validation and approval processes."

Just stating that business units have responsibility for model review is only the first step. Incentives need to be properly aligned to make sure this responsibility is taken seriously. Steps to assure this include:

- The business unit responsibility for model review is just as much about clear communication as it is about clear thinking. As Morini (2011, Section 1.4.1) emphasizes strongly: "The choice of a valuation model must be based on an analysis . . . reported to senior management in an *aggregated and understandable form*." "Quants, traders, and other technically strong practitioners" must find ways to communicate technical ideas in nontechnical language "comprehensible for senior management." Technically strong practitioners who have difficulty in doing this should seek help from colleagues who have stronger communication skills or from corporate risk management personnel who have more experience in this aspect of model review. But no one should be under the illusion that they will escape responsibility for consequences because "the senior guys just weren't capable of understanding what we were doing." If you truly can't get senior managers to understand the potential consequences, even with renewed effort at clever communications, then this is a product your business unit should *not* be trading.
- A clear distinction should be made between losses due to market uncertainties that were clearly identified and advertised as part of the business unit's model review and losses due to market uncertainties that were ignored in the model review. The latter should have more serious consequences for performance review and compensation than the former, and this policy should be widely advertised within the firm.
- To make sure that the policy in the previous bullet point is successfully implemented, an analysis of significant trading losses needs to be conducted by control personnel independent of the business unit to determine how losses are related to model reviews.
- The independent model review conducted by the risk management area should include identification of weaknesses in the business unit

model review. Patterns of weaknesses need to be addressed by corrective action, as well as consequences for performance review and compensation.

I have not addressed here the issue of how business unit model responsibility should be divided between model builders and traders. This will have different solutions for different business units and different models. I will just note again my comments in Section 8.1 that any function seeking to shun responsibility for model error should accept that reduced responsibility and reduced compensation opportunities go hand in hand.

This emphasis on business unit accountability for model review is consistent with placing the main responsibility for model development with the business unit and allowing them as much freedom in structuring models as possible. Models need to take advantage of as much inside information, in the form of trader beliefs about the future, as possible. Firms must try to be open to as many trading ideas as possible and not dismiss ideas on the grounds that they do not line up with some approved theory (for example, rational expectations or marketplace efficiency). However, a culling process must also be available for measuring the success of trading ideas and eliminating those ideas that are not proving successful. Insiders should be given latitude in the theories used in deciding how to trade, but not in the theories used in deciding when to recognize P&L. Profits should not be booked and bonuses not be paid out until the forecasts of the trading models have proven correct.

For decisions on when to book P&L, it is better to rely on outsiders to avoid bias. You may lose accuracy by not having access to the insiders' market knowledge, but this will only result in delays in recognizing earnings, which is not as serious a problem as taking the wrong positions. Insiders may object that this delay in recognizing P&L will cause them to turn away good business, but they have two alternatives: find others in the market who share their opinions and sell off the risk recognizing the profits, or, if they are sufficiently confident, wait to recognize the P&L until after the risk position has matured.

The risk management units that are part of control functions and that constitute the independent "second set of eyes" in model review do not have the business units' incentive issues. It is very clear to them that unidentified model risks that lead to trading losses will cost them in compensation and may cost them their jobs and even their careers. The incentive problem here runs in the opposite direction: the negative consequences of approving a model that later proves defective are so clear that there is a danger of playing it safe by creating unreasonably high barriers to model approval. After all, a rejected model will never have a chance to show how it would have performed, so there might appear to be little danger in being proved wrong by being overly cautious. Fortunately, in most firms, business units will press their case with sufficient passion and sound analysis to overcome such unreasonable barriers to new business. But managers of independent model reviews need to be alert to this temptation toward caution, and need to be constantly challenging model reviewers to make sure the right balance is being struck.

The problems for independent risk managers are more likely to rest with issues of expertise and access than they are with issues of incentive. As outsiders, they have less chance to build up the thorough knowledge of models and markets that business units possess. This can sometimes be addressed by employing former model builders as independent reviewers, but this is a career move that appeals to only a small subset of model builders. Other techniques for trying to overcome this gap in expertise will be addressed throughout the remainder of this chapter and this book. For issues of access, rules need to be put in place and enforced to see that business units are forced to share model code, documentation, and supporting data with independent reviewers. Claims of need for secrecy to protect proprietary model features must be viewed with suspicion—these are often just excuses to try to avoid independent scrutiny. When found legitimate, such claims need to be addressed by controls that restrict access to only those actually directly involved in the independent review; they must *never* be used as a reason to limit the scope of the independent review.

Independent model reviewers need to clearly identify steps that need to be taken when they find issues with models. As FRB (2011, Section VI) states, "Control staff should have the authority to restrict the use of models and monitor any limits on model usage. While they may grant exceptions to typical procedures of model validation on a temporary basis, that authority should be subject to other control mechanisms, such as timelines for completing validation work and limits on model use." In all cases where follow-up action is called for, there should be definite dates for further review established and a well-organized procedure for making certain that a follow-up review is performed evaluating these follow-up actions.

The role just specified for independent model review is consistent with the guiding principle of FRB (2011, Section III) of "effective challenge" of models—"critical analysis by objective, informed parties who can identify model limitations and assumptions and produce appropriate changes." Requirements for effective challenge outlined by the Federal Reserve are separation of the challenge from the model development process, knowledge and modeling skills adequate to conduct appropriate analysis and critique, and sufficient "influence to ensure that actions are taken to address model issues."

To what extent should external resources (i.e., consultants) be used as part of the independent model review process? There are many reasons for wanting to minimize the use of external resources: the desire for confidentiality of proprietary models, the desire to build up in-house expertise through the experience gained by conducting model reviews, and the fears of discontinuity if an external resource becomes unavailable or proves unsatisfactory. There still may be times when use of external resources is desirable, either because of a lack of in-house expertise within the independent model review group or because manpower available is not sufficient to meet the demand of new models needing review. When external resources are utilized, care should be taken that a designated in-house reviewer becomes as familiar as possible with the work of the consultant. This serves the function of acquiring some in-house expertise that can be utilized in subsequent model reviews, as well as having someone in-house who can monitor and coordinate the work of the consultant, provide a point of contact for subsequent discussion of the consultant's work, and be able to ramp up involvement in case of discontinuity. (Compare the discussion in this paragraph with the segment headed "External Resources" in FRB [2011, Section VI]).

Finally, all model review activities, both those of business units and those of independent reviewers, must be properly documented. This covers both documentation of the model itself and of the model review process. Model developers are often anxious to get on to the next project, and business units are anxious to develop the next model; model documentation can get shortchanged in the process. Poorly documented models are likely to cost money in the long run, by making model revisions more difficult and time-consuming and by increasing the likelihood that model errors will be missed. If necessary, business units may need to establish separate model documentation teams to complete documentation based on interviews with model developers.

Standards for documentation of models and model reviews should include the following:

- A review of the adequacy of business unit documentation by the independent model reviewer should be included, with recommendations for gaps that need to be remedied.
- "Documentation and tracking of activities surrounding model development, implementation, use and validation are needed to provide a record that makes compliance with process transparent" (FRB 2011, Section VI).
- An inventory of all models that require review should be maintained, both by business unit and firmwide. This inventory can serve as a central control point for scheduling model reviews, keeping track of

documentation, providing information on contacts, and scheduling updates of model reviews (FRB 2011, Section VI, "Model Inventory").
- Documentation is also required for the policies governing model review and the roles and responsibilities of business units and independent reviewers (FRB 2011, Section VI, "Policies and Procedures").

What is the role for the senior management of the firm in model review? FRB (2011, Section VI) emphasizes senior management responsibility for assuring that a process is in place that meets the standards outlined in this section and in Section 8.2.1. This certainly includes providing adequate funding for these functions. Basel (2009b) has a similar statement in its Principle 1 and, in Principle 2, emphasizes that the review capacity has to be adequate to handle conditions of stress. This is obviously a response to the stressed conditions of the 2008 crisis.

In addition to these more formal requirements, senior management must be prepared to understand the aggregate level of model risk that the firm is exposed to and to set limits on this risk. Having a requirement that business units communicate model risk in an aggregated and comprehensible form to senior managers, as we have at the start of this section, entails a corresponding responsibility of senior managers to make use of this information. There will, of course, be cases of conflicting presentation to senior management—often a more sanguine view of risk from the business unit and a more cautious view from the risk management groups. Senior managers must insist that both sides get a fair hearing, preferably in the same room at the same time, and that arguments be presented in a comprehensible manner. At the end of the process, it is senior management that owns the risk and must reach a decision.

Boards of directors in principle exercise an oversight role over senior management in controlling model risk, as in all other critical aspects of the business. In practice, it is very difficult for directors, whose involvement is only for a small part of each month, to have much impact, particularly since their access to information is often tightly controlled by senior management. The one course of action I would recommend is that directors on the risk committee of the board insist on private meetings with senior risk management personnel. This will at least provide a forum for concerns to be expressed that could give directors enough information to pose questions to senior management.

8.2.3 Model Verification

Most model problems are related to the fit between the product or trading strategy being modeled and the model selected. This issue of model validation we will address in Section 8.2.6. Here we deal with the simpler question of whether the model selected has actually been properly implemented—does the model actually do what it claims to do? This can be controlled by adequate model documentation and thorough checking by competent reviewers before the model is put into production. Here are a few rules that should be borne in mind to make sure this gets done properly (more detail on these points can be found in FRB [2011, Section V] and Morini [2011, Section 1.5.1, "Model Verification"]):

- Thorough documentation of what the model is trying to achieve, model assumptions, and derivation of formulas must be insisted on. All formula derivations should receive an independent check. Useful advice from Morini (2011, Section 1.5.1) is: "When a model is used for the first time the passages from dynamics to closed-form formulas or the other way around should be verified. This should be the case for any new model developed by a front office quant, also for any new model that simply appears on the Internet—or in a journal. There are errors even in published literature, never be too trusty."
- Systems implementation of the model should be subject to rigorous standards of documentation, change control procedures, and systems testing.
- The best check on an implementation is to perform an independent implementation and see if the results agree. It is tempting to cut costs by confining checking to having an independent analyst read through the documentation, equations, and code of the model builder and confirm it is correct. But it is much easier to miss an error in reading through someone else's equations or programming code; it is much more unlikely for two analysts working independently of one another to make the same error.
- Whenever possible, the independent implementation used as a check should employ a different solution methodology than the implementation being tested. For example, if the implementation being tested has used a Monte Carlo simulation, the test should be made solving backwards on a tree, where this is feasible. Using different implementation methodologies reduces further the chances that the two implementations will have the same flaw.
- Models should be tested on degenerate cases that have known solutions. For example, a down-and-out call with a barrier of zero is equivalent to a vanilla call, so setting the barrier to zero in a down-and-out call model should produce the standard Black-Scholes result. Other examples would be: (1) to always check that put-call parity for European-style options is preserved for any model used to price options, and (2) to

always check exotic option models against known analytic solutions for flat volatility surfaces (see the introduction to Section 12.1 and Section 12.3.1).

- Models should be tested for their impact on VaR and stress test calculations as well as on valuation and limit calculations.
- Models should be tested on extreme inputs to see that they handle these cases properly. For example, interest rates much lower and much higher than have recently been experienced should be input to see that the model can produce reasonable results. For more details on this model stress testing, see Morini (2011, Section 3.1).
- Produce graphs of model output plotted against model inputs and explore any instances where they do not make intuitive sense. This is another good check on the model's ability to handle extreme inputs, as per the previous bullet point. The impact of varying several inputs simultaneously should be compared to the sum of the individual impacts to check if the interactions of variable changes produce reasonable results.
- When a new model is replacing an existing model, a thorough benchmarking process should be used to compare results of the two models for an identical set of inputs. Model differences should be checked for reasonableness and unreasonable differences investigated. The same benchmarking standards should be utilized whenever one systems implementation of an existing model is being replaced or supplemented by another systems implementation of that model.
- Model error due to incorrect representation of transactions is just as worrisome as model error due to incorrect equations. This will be addressed in Section 8.2.4.
- A particular point of concern is approximation error introduced by the need for fast response time in a production environment. This will be addressed in Section 8.2.5.

Be careful about the degree of complexity introduced into models. Is there sufficient gain in accuracy to justify the reduction in intuitive understanding that results from added complexity? To illustrate with an example from my own experience:

I had recently taken a new job and found that my most pressing problem was widespread user dissatisfaction with a model upgrade that had recently been introduced. The old model had been easy for traders and risk managers to understand; the new one was supposed to be more accurate, but could be understood only by the model development group. My initial examination showed that, on theoretical grounds, the difference between answers from the two models should be too small to make an actual difference to decision making, so I tried to persuade the model builder to switch back to the original, simpler model. Finding him adamant on the need for what he viewed as theoretical correctness, I examined the new model more closely and found a major implementation error—a factor of 2 had been dropped in the equation derivation. This is the sort of mechanical error that would certainly have been picked up as soon as a formal model review was performed. But a similar error in a less complex model would have been caught long before, by the people using it on a day-to-day basis.

8.2.4 Model Verification of Deal Representation

Verification of transaction details that serve as input to models can be just as important in avoiding valuation and risk measurement errors as verification of the model itself.

The quote from FRB (2011) in Section 8.2.1 that "Models are of necessity simplified representations of real-world relationships and so can never be perfect" applies just as much to the representation of transactions in models as it does to the models themselves. A single transaction confirmation document often runs to tens of pages and its representation in the model is just a few numbers, so inevitably some simplification and approximation are being utilized.

In some ways, this is a more difficult issue to deal with than verification of the model itself, because of the large number of transactions that are often input to a single model. Reconciling transaction details between confirmations and position entries to models is an important middle-office control function, as emphasized in Sections 3.1.1 and 3.1.2, and will certainly be expected to catch numerical errors in data entry and details such as correct day count convention. But middle-office personnel lack the intimate knowledge of the model that might allow them to identify an important contract detail that is not being captured in the way the transaction is being represented in the model. Some steps that should be taken to control this risk are:

- While model builders and independent model reviewers who do have intimate model knowledge won't have the time to review every transaction confirmation, they should review all of the very largest transactions and a sample of the remainder, to look for both individual errors and patterns of errors. Samples should be selected at random, but with some weighting scheme that makes it more probable that large-impact transactions have more of a chance of being part of the reviewed sample than those of lesser impact. Review should consist of a thorough reading of the confirmation, comparison with how the confirmation has been represented in the model input, and consideration of any possible gaps in the representation.

- Middle-office personnel should be strongly encouraged to immediately raise any question they have about adequate representation with an independent model reviewer.
- In the daily P&L verification process, discussed in Section 8.2.7.1, any transaction that makes a payment significantly out of line with the payment projected by the model should be investigated. This may uncover an outright error in data entry, but may also identify a facet of the contract that has not been adequately represented in the model input.
- Some wording differences between different variants of contracts may be very subtle (for example, see the discussion of legal basis risk on credit default swaps in Section 3.2 and Section 13.1.1.2). The best approach in this type of case may not be to try to capture all these variants in the model input. It may be better to have a separate offline calculation of the risks arising from wording differences and to establish limits and reserves against this risk on the basis of this offline calculation. This can be regarded as a type of liquid proxy, per our discussion in Section 8.4, with the most common contract type serving as the liquid proxy in all standard risk calculations, such as VaR and stress tests, but with the separate offline calculation capturing the nonliquid risk.

8.2.5 Model Verification of Approximations

The quote from FRB (2011) in Section 8.2.1 that "Models are of necessity simplified representations of real-world relationships and so can never be perfect" could be paraphrased as saying that "All models are approximations." But for model review it is very important to distinguish between two different types of approximations:

1. Approximations in which some source of risk or driver of value has been omitted to reduce model complexity.
2. Approximations in which a computational approximation is being used in order to speed calculations and reduce cost.

Approximations involving the omission of a risk factor pose greater challenges for model review, since it can be very difficult to estimate the potential impact on earnings and risk. This issue dominates our discussion of model validation in Section 8.2.6 and in Sections 8.3, 8.4, and 8.5. Approximations involving computational approximation are much easier to control, since the model review process can create a detailed comparison between the production model and a more thorough model that is run less often or only on a selected sample of transactions. This section focuses on techniques for dealing with computational approximations.

We begin by looking at a set of suggested controls for computational approximation and then illustrate with a detailed example. Suggested controls are:

- Model reviews should explicitly recognize the trade-offs between model accuracy and investment of resources. Models used in production must be sufficiently fast to produce answers within the time frame required for providing quotes to customers and providing risk analysis to the trading desk and senior management or they will prove useless. Their development cost must be reasonably related to the revenue that can be realized on the products they support. The time required for development must be consistent with overall business plans.
- Evaluations of the inaccuracy of a production model need to be made by comparison to a more thorough model. Since model testing can be performed over a period of days or weeks, as compared to the minutes or seconds required of a production model, there is ample room to develop much more thorough models in testing environments. Comparison of results to the production model will show just how much accuracy is being lost.
- Comparisons of the production model to a more thorough model need to be performed not just for current market conditions. Tests should be performed to anticipate the impact on approximation of potential future market conditions.
- Where this test shows significant loss of accuracy, this identifies a good target for improved approximations. Until such improvements can be implemented, remedies can include valuation reserves against inaccuracy along with periodic revaluations with a slower but more accurate model, as well as traders exercising a degree of conservatism in pricing and hedging.
- Improved approximations can be achieved by "throwing more money" at the problem—buying more hardware to increase the number of calculations that can be performed in a given period of time. But more can usually be accomplished by the design of clever approximation algorithms. Indeed, one of the dirty little secrets of industry quants is just how much of the effort of people with PhDs goes into applying advanced mathematics to creating better approximation algorithms, rather than to the creation of new ideas for financial modeling. For example, see Section 13.3.3.
- In some cases, the thorough model can be so computationally intensive that it can be evaluated on only a sample of transactions. This is clearly a less desirable test of accuracy than one that looks at the full portfolio, but when this is necessary the reviewed sample should include all of the

very largest transactions and a random selection of the remainder. The random selection should be chosen with some weighting scheme that makes it more probable that large-impact transactions have more of a chance of being part of the reviewed sample than those of lesser impact.

- Approximations need to be reevaluated periodically, since approximation inaccuracy can be strongly related to portfolio size and composition. This point will be discussed in more detail in Section 8.2.8.1.
- A clear distinction needs to be made between the degree of accuracy needed to specify initial market conditions and the degree of accuracy needed to specify the evolution of market conditions. For example, as we will see in Section 12.5.2, a multifactor model for the evolution of interest rates does not offer much added accuracy over a single-factor model for the valuation of Bermudan swaptions. But a specification of the shape of the initial yield curve that does not utilize a full set of liquid points on the yield curve can have a very significant impact on this valuation.

As an illustrative example, I want to consider a situation I was involved with as a model reviewer. It involved a large portfolio of illiquid interest rate derivatives, all of which required Monte Carlo simulation for valuation calculations and for calculations of risk parameters. The number of simulation paths being run had been very clearly selected by the traders as the number that would allow all of the needed calculations to be performed between the close of business and the opening of trading the next day.

To test for the impact of this choice on accuracy, I first set up a simulation for the whole portfolio that would run continuously for about a week in an offline environment. The results from this very much larger number of runs allowed me to estimate the number of runs needed to determine accurate valuation to within the tolerance required for financial significance (see, for example, Hull 2012, Section 20.6, "Number of Trials"). I was also able to see how much variance from this accurate valuation resulted from the smaller sample being used in production runs.

I next used the larger sample to estimate the impact of the selection of different sets of Monte Carlo simulation paths on the production run. I determined that there was reasonable stability over a sufficiently small time frame; a set of paths that produced accurate values when comparing results from the smaller number of paths in the production run to the larger number of paths used in the offline run would also be fairly accurate over the next few days. But as the time period lengthened from days to weeks, a set of paths that had previously produced accurate values lost accuracy, both as a result of shifts in composition of the portfolio as new transactions were added and older transactions had less time left to expiry, and as a result of changes in market parameters.

I therefore set up the following process. Once a month, a new offline run would take place and would be used to determine the set of paths that was going to be used in the production runs for the next month. These production runs determined valuations and sensitivities reported for P&L and risk management purposes. Each month, the shift from one set of paths used in the production runs to a new set of paths would cause a change in P&L. These changes were randomly distributed, as likely to be increases in P&L as decreases, with a standard deviation that could be estimated from comparisons between valuations of different subsets of paths in the offline run.

I argued that while these changes were randomly distributed, the firm should have a reserve against negative changes so that we were reporting to shareholders only valuations we could be fairly sure were actually achievable in the long run (since the portfolio consisted of illiquid positions, it was not possible to realize a current value just by liquidation—we were committed to holding the portfolio for a longer time period). Therefore, each time new transactions were added, the reserve would be increased to reflect added uncertainty, while as older transactions got closer to expiry, the reserve would be reduced to reflect less uncertainty. Each month, when the change in paths took place, the resulting gain or loss would impact the reserve and would not impact P&L (the reserve would have had to be exhausted to impact P&L, but this never occurred in practice). I was able to get agreement from the firm's finance function and accounting firm to make this process part of the official books and records of the firm, and not just the risk management reports.

Note that this process had a built-in set of controls against changes in the market environment or portfolio composition, since the monthly offline run would automatically pick up any such impacts; the size of required reserve was recomputed each month by recomputation of the standard deviation between valuations of different sets of paths.

The issue of computational approximations is particularly important for credit portfolio models and the closely related collateralized debt obligation (CDO) models. These will be covered in some detail in Sections 13.3.3 and 13.4.2. Another area in which computational approximation plays a major role is in multifactor interest rate models. This is discussed in great depth in Morini (2011, Section 6.2), which includes extensive examination of the accuracy of these approximations.

8.2.6 Model Validation

We must now move beyond the tests of internal model consistency we have focused on in Sections 8.2.3, 8.2.4, and 8.2.5 to look at the fit between the model and the product or trading strategy being modeled, what Morini

(2011) calls *model validation.* It is not surprising that this more challenging task does not lend itself to the easy consensus and process-oriented approach of these last three sections. We will distinguish between three basic approaches to model validation: one focused on interpolation, one focused on the long-term cost of hedging, and one focused on discovering the prevailing market model. These are not completely competitive approaches—there is some overlap among them—but they do have distinctly different emphases. As we will see, the appropriate approach has much to do with the purpose the model will be used for and the liquidity of the product being modeled. We will also look at the issue of how to deal with risks that may not be evident to model reviewers. I would strongly recommend comparing my approach in this section to Morini (2011, Sections 1.5.1 and 1.5.2).

8.2.6.1 Interpolation Approach In the interpolation approach, models are viewed as primarily serving as interpolation tools from observable to unobservable prices. This is closely related to the view discussed in Section 8.1 that downplays the importance of models. Viewing models as interpolation tools provides valuable insight into why certain models have been able to achieve a high degree of acceptance in financial management. It is much easier to agree on an interpolation methodology than it is to agree on a fundamental method for pricing an instrument. The danger is that this view leads to unwarranted complacency, since model builders often regard interpolation as being a mathematically trivial or uninteresting task. The result can be uncritical acceptance of what seems a plausible interpolation method or a view that the choice of interpolation methods is somehow a matter of taste.

A closer examination will show that every choice of interpolation method entails significant financial assumptions. The interpolation of an unobservable price based on a set of observable prices amounts to the theory that the instrument with the unobservable price can be well hedged by the set of instruments with observable prices. As with any theory, this should be subjected to empirical testing and competition with alternative hedging proposals. Even the simplest-sounding interpolation proposal (for example, calculating the two-and-a-half-year rate as a 50–50 average of the two- and three-year rates) should be regarded as a model subject to the same tests as more mathematically complex models. We examine this in more detail in Sections 8.3 and 10.2.1. Models rarely cost firms money because modelers have made an error in complex mathematics; they frequently cost firms money because they embody financial assumptions that are not borne out by future events.

8.2.6.2 Cost of Hedging Approach Basing model validation on an examination of the possible costs of hedging a transaction over the long term is closely related to the approach advocated in Section 6.1.2 of establishing a liquid proxy for an illiquid instrument and then simulating the difference between the liquid proxy and the actual trade. This viewpoint has been laid out very eloquently in Derman (2001):

> *It's never clear what profit and loss will result from hedging a derivative security to its expiration. Markets will move in unexpected ways, sometimes intensifying transactions costs and often dismantling what seemed a reasonable hedging strategy. These effects are rarely captured by the conventional models used in front-office valuation systems. . . .*
>
> *Therefore, for illiquid positions, it is important to estimate the adjustments to conventional marked values that can occur as a result of long-term hedging. One should build Monte Carlo models that simulate both underlyer behavior and a trader's hedging strategy to create distributions of the resultant profit or loss of the whole portfolio. These distributions can be used to determine a realistic adjustment to the trading desk's conventional marks that can be withheld until the trade is unwound and their realized profit or loss determined. . . . Monte Carlo analysis provides a good sense of the variation in portfolio value that will be exhibited over the life of the trade due to transactions costs, hedging error and model risk. Ultimately, such analyses should be part of the desk's own front-office valuation system.*

Note: There is a rough equivalence between Derman's use of "underlyer" liquid instruments being used to hedge the illiquid instrument and my emphasis on a representative liquid hedge. As Derman says: "Derivative models work best when they use as their constituents underlying securities that are one level simpler and one level more liquid than the derivative itself."

8.2.6.3 Prevailing Market Model Approach Rebonato (2003) emphasizes model validation based on the model that prevails in the marketplace and anticipation of directions in which the prevailing market model might evolve:

"Model risk is the risk of occurrence of a significant difference between the mark-to-model value of a complex and/or illiquid instrument, and the price at which the same instrument is revealed to have traded in the market."

- "From the perspective of the risk manager the first and foremost task in model risk management is the identification of the model ('right' or 'wrong' as it may be) currently used by the market to arrive at traded prices."

- "[M]arket intelligence and contacts with the trader community at other institutions are invaluable."
- Requires a variety of models to reverse-engineer observed prices.
- Requires information about as many observed prices as possible.
- "No matter how good or convincing a theoretical model might be, few states of affairs should worry a risk manager more than the trader who, using this model, consistently beats all competing banks in a competitive-tender situation."

- "The next important task of the risk manager is to surmise how today's accepted pricing methodology might change in the future" (including changes to model, changes to calibration, and changes to numerical implementation). "Being aware of the latest market developments, and of academic papers can be very useful in guessing which direction the market might evolve tomorrow."
- "To a large extent, the model risk management task can be described as an interpolation and extrapolation exercise that simply cannot be carried out in an informational vacuum . . . without at least some anchor points of solid knowledge about the levels and nature of actual market transactions."

Hull and Suo (2001) present an approach to model validation closely related to Rebonato's. They quantify the risk of a model being used by a trading desk by estimating how much of a loss the trading desk would suffer if a different model turned out to be correct.

8.2.6.4 Matching Model Validation to Model Purpose There are two dimensions to matching model validation to model purpose. The first relates to the degree of liquidity of the instrument being modeled. The second relates to differences between models being used for managing the firm's overall risk, the models of valuation and position measurement described in Section 6.1, and models being used to make trading decisions.

The key to designing a proper model valuation procedure for models being used to manage the firm's risk is to fit the model review to the degree of liquidity of the instrument for which the model is being used. Essentially, we need to work out the model review implications of the liquidity differences discussed in Section 6.1.1.

Please note that the distinction here is between liquid and illiquid *instruments*, not liquid and illiquid *positions*. If a position in a liquid instrument is so large as to create an illiquid position, this needs to be dealt with by modifying VaR and stress test calculations, as discussed in Section 6.1.4, but does not require a different model or model review than would be needed for a smaller position in the same instrument.

The interpolation approach to model validation is usually very reasonable for liquid instruments. We will look at the details of applying the interpolation approach to liquid instruments in Section 8.3. For illiquid instruments, the interpolation approach has little relevance; the kind of frequent checks of interpolation methodology with the market recommended in Section 8.3 are not possible because of illiquidity. Given that interpolation will be of little use, model verification for illiquid instruments must therefore rely on either the cost of hedging approach or the prevailing market model approach or some combination of the two. We explore this issue in detail in Section 8.4. For models used in making trading decisions, discussed further in Section 8.5, the prevailing market model approach is the most salient, as the next example illustrates.

To better understand the implications of different types of model use, consider the case of the 1998 breakdown of the historical relationship between the pricing of interest rate caps and interest rate swaptions, discussed in detail in Morini (2011, Sections 11.3 and 11.4). For models utilized for managing the firm's overall risk on liquid instruments, this breakdown was probably a nonevent. Since both caps and swaptions had adequate liquidity in external price quotes, most firms would be using some form of interpolation model to value and compute risk statistics for their caps just using market cap prices, and a separate interpolation model to value and compute risk statistics for their swaptions just using market swaption prices. There would have been no interaction between the two models. (Some possible exceptions where liquid prices in one market would be allowed to override liquid prices in another market will be examined in Section 8.3.)

There may have been a different story for models used in managing the firm's overall risk on illiquid products. Some products, such as Bermudan swaptions, knock-out caps, and forward-start interest rate options (see Section 12.5.2 for details), may have been priced using models that incorporated both market cap prices and market swaption prices as inputs. Would the breakdown in the historical relationship have caused problems for these models? Using the approach we discuss in Section 8.4, there should be a very significant degree of conservatism relative to historical relationships built into the reserves kept against model risk and, in any case, the relationship that is important is what will happen over the lives of the deals. Could a temporary period of breakdown in historical relationships be enough to call into question the adequacy of these reserves? We discuss this further in Section 8.4.

But models being used by the trading desk to determine trading strategies involving caps and swaptions would definitely have been impacted. Here's the description by Morini (2011): "No matter whether or not the long term equilibrium was going to come back, the market had gone too far from it for too long for a bank or a fund with a risk management unit to stand it." Here it is clearly the case that the prevailing market model approach must govern.

8.2.6.5 Capturing Risks That Are Difficult to Identify I once heard a senior risk manager for an investment bank say, "I don't stay awake at night worrying about the risks I know, but about the risks I don't know." This is a sentiment with which I could readily identify. In performing model validation, the great fear is that there will be some exposure that is not being captured by the model and that you, as the independent model reviewer, don't even know about. For example, there might be some potential piece of legislation or judicial decision that would have a big impact of the transaction being modeled, but it has never been discussed in the literature you have access to.

This type of potential exposure is probably known to the front-office personnel who specialize in the product. Ideally, they should consider it in their internal model review and share their concerns with the independent model reviewer. But the usual moral hazard concerns come into play, with the incentives discussed in Section 2.1, motivating front-office personnel to be reluctant to share information that might lead to tightened controls.

JPMorgan in the late 1990s instituted an internal system called Risk Identification for Large Exposures (RIFLE) to try to address this issue. It is still in operation, as can be seen from the following quote from JPMorgan's 2010 annual report:

> *Individuals who manage risk positions in the Investment Bank are responsible for identifying potential losses that could result from specific, unusual events, such as a potential change in tax legislation, or a particular combination of unusual market events. This information is aggregated centrally for the Investment Bank. Trading businesses are responsible for RIFLEs, thereby permitting the Firm to monitor further earnings vulnerability not adequately covered by standard risk measures. (p. 145)*

But even with a mechanism like this in place, the incentive issue remains. Traders who do an honest job of reporting these risks may thereby lower their return on risk measures and attract added scrutiny of position sizes. To attempt to overcome this, a firm needs to make clear distinctions in performance evaluation between losses that occurred due to an event that the traders had made certain received adequate firmwide attention in advance of the trade being approved and losses that occurred due to an event where this type of advance notice was not provided.

8.2.7 Continuous Review

FRB (2011, Section V) calls for "validation activities [that] continue on an ongoing basis after a model goes into use." Three major components of ongoing validation activities are (1) daily P&L reconciliation for models being used for valuation and risk-reporting purposes, (2) back-testing for statistical forecasting models, and (3) analysis of overrides for cases where model output needs to be altered based on the expert judgment of model users. We consider each in turn.

8.2.7.1 Daily P&L Reconciliation In Sections 3.1.1 and 3.1.2 we stressed the importance of a daily explanation of P&L produced by independent support staff as a control measure against both fraud and nondeliberate incorrect information. Here, we want to stress its equal importance as a tool for identifying model weaknesses.

The basic approach of P&L reconciliation is to take position reports from the close of business (COB) of the prior day and combine them with actual market movements from the previous day's COB to the current day's COB to estimate P&L for the day. The idea is that since position reports show sensitivities to changes in market variables (e.g., option "greeks"), multiplication of these sensitivities by actual price changes should produce a reasonable estimate of P&L. Of course, this applies only to those positions that were in place as of the COB of the previous day, so it is first necessary to identify and segregate the P&L due to trades booked during the day (this includes hedges that may have been put on during the day in response to market moves). This segregation of P&L between that due to previous COB positions and that due to trades booked during the day is already valuable as a tool in avoiding inadvertent Ponzi schemes in which profits on newly booked trades cover up hedge slippage on existing trades (compare with the discussion in Section 2.2).

If the initial estimate of P&L is significantly different than actual P&L, what can be the possible causes? Incorrectly recorded positions are certainly a possibility; this is why P&L reconciliation is a valuable tool in uncovering fraud and incorrect information. It's true that an incorrectly reported position might impact both the COB position report and the daily P&L record, but at some point there will be an actual payment on the position, and at this point, when the payment becomes part of the daily P&L, a discrepancy between projected P&L and actual P&L will show up.

Another possibility is that the COB position has not been reported with sufficient detail. For example, an option position might be reported using just the first-order greeks, such as delta and vega, and not the second-order greeks, such as DdelV and the price-vol matrix (see Section 11.4 for a detailed discussion of option sensitivity measures and how they can be used in P&L reconciliation). In this case, the P&L reconciliation will identify the need for more detail in the position report, which will enhance management's ability to accurately measure exposure.

A third possibility is a deficiency in the model. For example, it could be that even though the contract details of a position are correctly entered, the model is not using these details correctly in computing P&L or in computing the position. Again, we might worry that the model flaw will impact both P&L and position reporting and so will not be spotted in reconciliation. But when a payment date is reached and actual payment becomes part of the daily P&L, a discrepancy should appear. More likely, the model is missing or mishandling some key factor of risk, and a position that appears hedged is actually suffering some hedge slippage. An example might be a Bermudan swaption model that fails to identify some circumstances in which early exercise becomes more profitable for the counterparty. So daily P&L reconciliation should form an important part of identifying model problems that may have been missed in initial model validation.

8.2.7.2 Back-Testing Any statistical forecasting model needs to have continuous monitoring of actual performance. The best-known example in risk management is back-testing of VaR models, discussed in detail in Section 7.1.2. Since VaR models produce statistical distributions of the size of losses that can be expected to occur at different percentiles (e.g., 1 percent of the time, 2 percent of the time), these projected distributions need to be continuously compared to actual experience. Statistical analysis should be applied to results, and when this analysis indicates a strong probability that the actual distribution differs from the projected distribution, corrections to the model need to be considered. Until corrections can be made, an extra layer of conservatism may be necessary in utilizing limits and reports based on the existing model.

Similar back-testing is called for in any statistical model being used to suggest a trading strategy. In hedge funds and on trading desks, one frequently finds trading strategies employed based on statistical studies of how the strategy would have performed historically (see Fabozzi, Focardi, and Kolm 2010, Chapter 7). Continuous back-testing is needed to update evaluation of the model's performance and identify changes in market environment that require alteration or even abandonment of the model.

8.2.7.3 Analysis of Overrides When hedge funds and trading desks employ statistical models for trading, there will occasionally be the need for a trader to override the trading strategy recommended by the model because of economic insights that cause the trader to doubt the advisability of the model's recommendation. It is important that all such overrides be recorded and analyzed, with performance of the model strategy not pursued compared to success of the override, to spot possible needs for model modification.

It is less common for models used for valuation and risk reporting to be overridden, but there are examples. One typical case is the representation of binary options in option greeks. When trading desks do not use the liquid proxy representation of binary options by a call spread recommended in Section 12.1.4, it often happens that the trading desk must come to the risk managers and ask for an override on the large delta and gamma positions produced by a binary option nearing expiry at a price close to the strike. Risk managers will be willing to grant these requests, since the trading action that would be required to get back within the limit would be a foolish position to take, as discussed in Section 12.1.4. But any such overrides should be recorded and analyzed. It is just such analysis that has persuaded some firms to move to the use of the liquid proxy representation, since a model that is producing position reports that would be foolish to act on is clearly flawed.

8.2.8 Periodic Review

Once a model has been approved and is in production, ongoing validation is still required. FRB (2011, Section V) reiterates what has been a long-standing regulatory requirement that existing models should be reviewed at least once a year, but also calls for continuous monitoring of model performance. We'll examine periodic review in this section and ongoing monitoring in the next section.

To be productive, the periodic (generally, annual) review of existing models must be carefully designed. Merely replicating previous tests is likely to be both unlikely to produce new insights and wasteful of resources. Reviews need to be focused on changes to the environment in which the model is being used that should trigger new testing and possibly new conclusions. We'll focus on four types of environment changes that should be investigated: (1) changes in the population of transactions the model is being applied to, (2) changes in the market environment, (3) changes in the academic literature or in market practices, and (4) changes in technology. In addition, the periodic review should examine any patterns that have been revealed by the ongoing monitoring we describe in the next section.

8.2.8.1 Changes in the Population of Transactions Consider the Kidder Peabody disaster, discussed in Section 4.1.2, as an illustration. Whatever your opinion about whether Joe Jett was deliberately gaming the system, there is no doubt that the firm was ill-served by having a model that computed the value of forward transactions without proper discounting. But Kidder Peabody was hardly alone in the industry in using a model that omitted discounting of forwards. This is not due to a widespread ignorance of this fundamental principle of finance. What does often happen—and this is a pattern I have seen over and over again—is that a sensible decision is made at the time a model is built but is not subjected to adequate review as circumstances change.

So a model might be set up for valuing forwards that at the time of implementation is being used to evaluate trades that are of moderate size and no more than a few days forward. The added accuracy that comes from correct forward discounting might be quite small and thus easily justify a decision not to devote the added programming time and computational resources to include this factor. The situation changes as larger transactions with longer forward periods are added. As the situation changes through time, there comes a point at which the proper decision would be to change to a more accurate model. But the decision to invest the resources needed to improve accuracy can be a difficult one, involving considerable expense, diversion of resources from important new ventures, and perhaps a limitation on trading volume until the change is made. The environment may be changing gradually, so that no single point in time stands out as the time at which to switch.

This is the kind of situation in which a periodic review of the impact of changes in the population of transactions being valued by a model can be of tremendous value. Note that the change in population of transactions could be due to changes in number of transactions (an approximation that had little impact when the model was being used to value just a few deals causes more concern when the model is being used to value many deals); size of transactions (maybe the model is being used to value just a few deals, but the average size of deals has grown to the point that approximations have become worrisome); terms of transactions (for example, the large increase in the length of the forward period in the case discussed previously, or an increase in time to maturity, or the more frequent use of features that are difficult to evaluate); or a combination of all three.

Morini (2011, Section 1.4.1) explains the reasoning behind this policy very well:

> *A bank cannot expend big resources for a small exposure; and additionally banks and traders learn by trial and error, a new model needs to be tested for a while to really know its risks. When the exposure starts growing, a previous model validation must not automatically be considered valid: a surplus of effort can be spent on the model used, an effort that was not economically meaningful in the past but is crucial in the face of the increased exposure.*

When faced with a large change in deal population, an independent model review group must think very carefully about what is behind the change. Is it just a new market taking off to meet a customer need, or is it a structuring group looking to arbitrage a deficiency in the way a transaction is being valued or a risk is being measured? If traders and structurers are hired because of their skills in uncovering complex arbitrage opportunities in markets, one shouldn't be too shocked if they sometimes use the same skill set to try to find arbitrage opportunities in regulations, whether external (government) regulations or internal (risk management) regulations. When independent reviewers see signs that such an opportunity is possibly being exploited, they must expend extra effort on trying to uncover the motivation and possible consequences.

If a reviewer does spot a loophole being exploited, there should be no hesitancy in quickly improving the valuation procedure or risk measure. There will inevitably be cries of foul play from structurers who can no longer take advantage of the old system and complaints that "The rules of the game have been changed without warning." Those complaining need to be reminded that risk management is not a game but a serious endeavor to protect shareholders, depositors, and taxpayers. Keeping a fixed set of rules and allowing structurers to experiment and see where the weaknesses are is a recipe for disaster, as the rating agencies amply demonstrated by publishing fixed models for evaluating the risk of CDO tranches and letting bank structurers play with the models until they had designed trades that optimized the degree to which the models underreported deal risk (see Section 5.2.3 for further discussion of this example).

8.2.8.2 Changes in the Market Environment An example of a change in market environment would be a risk factor that previously could not be priced based on market observation but now has liquid prices available. This could change previous conclusions about which model inputs need to be derived from market prices. In the other direction, deterioration in the liquidity of a pricing source might prompt the need for new reserves or limits.

Another example would be changes in levels of prevailing market prices that might prompt reruns of sensitivity analyses and model stress tests. FRB (2011, Section V) states, "Sensitivity analysis and other checks for robustness and stability should likewise be repeated periodically. . . . If models only work well for certain ranges of input values, market conditions, or other factors, they should be monitored to identify situations where these constraints are approached or exceeded."

Another instance of change in market environment would be rapid growth in the size of a market. This should prompt reexamination of the relevance of historical data, since rapid growth may be the result of major changes in the nature of the market. A recent example was the explosive growth in U.S. subprime mortgages. The use of historical data on default rates for subprime mortgages should have then been treated with extreme caution. As it soon became clear, underwriting standards for approving these mortgages had become drastically more lax than in previous eras, which contributed to steeply rising default rates (see Section 5.2.1). A prior example was the precipitous growth in non-investment-grade bonds in the late 1970s and early 1980s. This growth was largely due to the efforts of Michael Milken at Drexel Burnham Lambert, and a major contributor was studies by Milken and others showing the very favorable historical returns of these bonds after adjusting for default losses. But as soon as there was a large increase in the issuance of these bonds, it should have been suspected (as turned out to be the case) that the growth was largely being fueled by types of transactions that had rarely been done previously and for which the historical data was of dubious relevance. Bruck (1988) is a good account of the Milken story; see particularly page 28–29 on historical return studies, and pages 266–270 on skepticism about their continued relevance.

8.2.8.3 Changes in the Academic Literature or in Market Practices Periodic reviews offer a good opportunity to consider any new approaches to modeling a particular type of transaction that have appeared in the academic literature, have been discussed at conferences, or have begun to be used by other market participants. This is where the emphasis in Rebonato (2003) on “market intelligence and contacts with the trader community at other institutions” and the reverse engineering of observed prices from other firms can be of particular value.

8.2.8.4 Changes in Technology Increased computational capacity may change the conditions on which previous decisions about approximation techniques have been made. Increased computational capacity could be due to newly purchased or upgraded hardware or to advances in computational theory. New conditions should lead to a reassessment of prior decisions—replacing existing approximation techniques either with more accurate ones or with full computations.

8.3 LIQUID INSTRUMENTS

Models for liquid instruments are robust and easy to test, since they can constantly be checked against actual liquid market quotes. This is why they lend themselves so readily to the interpolation approach to model validation outlined in Section 8.2.6.1. Risk reports only need to look at exposures, such as delta and vega, measured against current market prices. If changes in price levels lead to new exposure levels that concern senior managers, the liquidity of the instrument will allow for reduction in positions at the time the exposure exceeds desired levels.

We illustrate with an example of a model review for a very liquid instrument. Consider a portfolio of U.S. dollar interest rate instruments (e.g., interest rate futures, forward rate agreements, interest rate swaps, government bonds) with no option component (see Chapter 10 for a detailed discussion of how the risk on such a portfolio is managed). There will be liquid market quotes available throughout the day for trades on a large subsection of these positions. But many instruments will need some form of modeling for valuation, since even a very liquid ("on-the-run") instrument at the time of original transaction may soon become less liquid ("off-the-run") through the passage of time (e.g., a five-year swap, for which direct market quotes are readily available, soon becomes a four-year 11-month swap, whose liquidation price needs to be inferred from market quotes for on-the-run instruments). The models used for this off-the-run valuation will also be needed for computing the change of the portfolio's value in VaR and stress test simulations.

The models needed for these computations are quite standard throughout the financial industry by now, but there are still choices in interpolation methodology that need to be made that constitute forecasts of relative movements between instruments (Section 10.2.1 provides details). These modeling choices are best made by the front-office personnel who have the product expertise and superior data access. In addition, it is the front office to whom the profits from correct forecasting decisions (and losses from poor forecasting decisions) properly belong.

Model validation by outside reviewers only requires periodic checking of model valuations against actual market prices. Close agreement shows model adequacy; significant differences indicate the need to establish limits and/or valuation reserves and may serve as clues for model revision. The most robust price checks come when there is an actual transaction in an off-the-run instrument, but price checks can also be performed by polling brokers, dealers, and other independent sources of pricing information (issues involved in obtaining such quotes are addressed in Section 6.1.3). While actual conduct of the price checks may be performed by support staff, model reviewers and other senior control personnel should be involved in the design of the price check procedures, with regard to frequency and standards for confirmation of model adequacy.

The type of price check just discussed should be complemented by the daily P&L explanation exercises discussed in Section 8.2.7.1. As observed there, the P&L explanation process often identifies model deficiencies when there are unexplained P&L changes, particularly around transaction dates and dates for scheduled payments and resets. But even a thorough daily P&L explanation process should not be regarded as a full substitute for price checking; it may be that a model performs very well in handling on-the-run transactions from inception through maturity and is rarely tested on off-the-run transactions because the trading desk almost always transacts on on-the-run dates. But what happens if the desk goes through a stop-loss limit or if the firm's appetite for risk decreases? A reduction in positions may need to take place by reversing previously booked transactions that are now off-the-run. Risk managers will want to have prepared for this eventuality by testing model pricing of off-the-run positions.

If the disagreement between an observed market price and a model value represents a clear difference between where a risk can be sold at the current time and a theory as to the value of the asset over a longer period of time, then no matter how sound the reasoning behind the theory, I would recommend holding to the mark-to-market principle. If a firm deviates from this principle and values based on longer-term values that it believes can be realized, rather than on prices at which risks can currently be exited, it is turning short-term risks into much harder-to-evaluate long-term risks. Morini (2011, Section 1.2.1) supports this view colorfully: "on intuitive grounds, anyone who claims that arbitrage opportunities are abundant in the market should always be asked if he is fabulously rich. If he is not, it is not clear why he knows of so many free lunches and yet, rather than exploiting them, goes around passing preposterous judgments about market functioning."

However, sometimes the difference between an observed market price and a model value represents two different ways in which a risk can be sold at the current time. Although this would seem to violate several important axioms of finance theory—the *no-arbitrage principle* and the *law of one price*—these are just models and cannot expect any absolute deference in the face of empirical exceptions. However, there needs to be careful evaluation of what lies behind an observed difference between a market price and a model price before an intelligent decision can be made as to which is the best of two different ways to represent the risk.

Let's focus on a concrete illustration. You have observable market prices for a European call option, a European put option, and a forward to the same expiry date, with the same underlying and, in the case of the put and the call, the same strike price. The combined prices, however, do not agree with put-call parity. This would imply, for example, that a position in the put that you have sold can be offset in two different ways—you could buy a put, or you could synthetically create a put by buying a call and entering into a forward. It also implies that the call-forward combination will offset the position at a cheaper price than the direct purchase of the put.

What should a risk manager recommend in such circumstances? Since the main argument behind a no-arbitrage principle such as a put-call parity is that the lack of parity will be quickly eliminated by profit-seekers taking advantage of a riskless opportunity to make money, any persistence of parity violation is suggestive of some liquidity difficulties preventing the opportunity from being exploited. We'll consider some possibilities:

- This is an arbitrage of which very few market participants can take advantage, but your firm is one that can. This could be because the market for the put is in some way restricted to only a few firms. It could be an arbitrage that is difficult to identify computationally and your firm has a computational advantage. It could be a diversified basket of assets that is difficult to accumulate and your firm has an advantage in its market access (see the discussion in Section 12.4.1.1). In such cases, it is right to base valuation on the model-derived price (in this instance, the call-forward combination), since this represents a liquid external price at which risk can actually be extinguished in the short term.
- One of the prices is less liquid than the others. For example, the amount of trading for that strike and date could be much more active in calls than in puts. This would be a strong indication of the desirability of using a model (put-call parity) to supply a price based on more liquid quotations rather than utilizing a less liquid price. The same reasoning would apply if the call and put markets are significantly more active than the forward market, in which case I would recommend replacing an illiquid forward price with a put-call parity–derived price based on liquid put and call prices.
- A timing difference exists in price quotations. Perhaps the options market posts closing prices at an earlier time of day than the forwards market. It is certainly legitimate to use a model to update both call and put quotes to adjust for changes in the forward since the time the options market closed.
- Some contract features make the model not completely applicable. Sometimes, on closer examination, contract provisions call into question the applicability of a model. In this case, it might be an allowance for early option exercise in certain circumstances, whereas put-call parity applies only to options without early exercise provisions.

This last type of case has led to a considerable number of disputes between risk managers and trading desks. One example that has arisen at several firms is traders' desire to unlock stock option values contained in convertible bonds. Option models applied to convertible bond prices frequently indicate implied volatilities that are quite low compared with the implied volatilities that can be derived from plain equity options on the same stock, leading traders to conclude that buying the convertible is a good value trade. Trading desks hungry to book immediate profits have pressed for overriding reasonably liquid convertible price quotes with a model-driven quote based on the implied volatility from the equity options market. But a convertible bond contains the option to exchange a bond obligation for a stock obligation rather than to exchange cash for a stock obligation, so it cannot be completely reduced to the value of an equity option (see the discussion in Section 12.4.4). When turned down on their first attempt, some trading desks have shown good enterprise in marketing total return swaps on the bond portion of the convertible in an attempt to isolate the equity option portion. So long as the swap has been properly engineered to cover all contingencies, such as canceling the swap without penalty in the event that the bond is converted for equity, a complete decomposition can be achieved and it is legitimate to value the resulting position as an equity option. Risk managers have, however, been very careful to check that no uncovered contingencies are present before allowing this valuation change.

8.4 ILLIQUID INSTRUMENTS

8.4.1 Choice of Model Validation Approach

Model use for illiquid instruments is much more critical than it is for liquid instruments and, unfortunately, model validation is also much more challenging. There may be a complete absence of actual market prices at which positions can be unwound, so modeling assumptions and inputs for unverifiable model parameters now become a key driver of model valuation.

Both Derman (2001) and Rebonato (2003) have strong statements as to the difficulty these risks can entail.

- Derman: “Because of their illiquidity, many of these positions [in long-term or exotic over-the-counter derivative securities that have been designed to satisfy the risk preferences of their customers] will be held for years. Despite their long-term nature, their daily values affect the short-term profit and loss of the banks that trade them.”
- Rebonato: “What differentiates trading in opaque instruments from other trading activities is the possibility that the bank might accumulate a large volume of aggressively-marked opaque instruments. When, eventually, the true market prices are discovered, the book-value re-adjustment is sudden, and can be very large. Stop-loss limits are ineffective to deal with this situation, since the gates can only be shut once the horse has well and truly bolted.”

Let's consider which of the model validation methodologies of Section 8.2.6—cost of hedging or prevailing market model—is more appropriate for illiquid instruments. My own view is that the cost of hedging approach is the more relevant for independent model reviewers for two reasons:

1. Even if a given model prevails in the market place, so long as the trading desk can't actually extinguish positions at the prices implied by the model, owing to illiquidity, it is actually the hedging costs that will determine the firm's P&L on the product. The model might continue to prevail in the marketplace for many years, and all the while the firm loses money on its hedging strategy. An advocate of the prevailing market model approach might respond that if, in fact, the model leads to hedging losses, then firms will eventually replace the model, so this is just a case of anticipating the direction in which the prevailing market model may evolve, in line with Rebonato's proposed criteria. But then I would still want to utilize the cost of hedging approach as a key tool in anticipating prevailing market model evolution.
2. I am wary of the ability of risk managers to anticipate prevailing market model evolution using any other tool besides cost of hedging simulation. Some of the tools Rebonato recommends—market intelligence and contacts with the trader community at other institutions—seem much easier for traders to utilize than independent reviewers.

If independent reviewers do rely on the cost of hedging approach, it would still be valuable for the front-office reviewers to utilize the prevailing market model approach as a supplement. This is particularly true when mark-to-market policies require marking to the prevailing market model, so that even if an instrument is being priced and hedged in a way that will virtually guarantee long-term profit, accounting losses may need to be booked in the shorter term. Rebonato (2003) makes this clear in Section 2.1, saying that "model risk arises . . . because of a discrepancy between the model value and the value that must be recorded for accounting purposes." This would not be the case for the mark-to-market policy I advocate in Section 8.4.4.

In a thorough review of cost of hedging, Monte Carlo simulation allows systematic consideration of many possible future paths of relevant liquid market prices and other economic variables. The soundness of the model can be judged only over longer time periods, when longer-term unobservable prices transform into shorter-term observable prices, when there is enough time to observe the impact of required rehedges, or when trades reach maturity and require contractual payments. Over a short time period, almost any model chosen will appear to perform well by a type of circular reasoning: The instrument with unobservable prices will be valued using the model and the observable price inputs. Therefore, the movement of the unobservable prices relative to the observable prices will seem stable since the same model is being used for valuation throughout the time period (see the example of this discussed in Section 10.2.1).

Proper design of a model review of an illiquid instrument utilizing Monte Carlo simulation has two parts: (1) choice of the liquid proxy, which will be analyzed in Section 8.4.2, and (2) design of the simulation, discussed in Section 8.4.3. I will then look at issues this approach raises for mark-to-market policies in 8.4.4 and for risk measurement in 8.4.5.

Another application of this approach to creation of liquid proxies and simulations of hedge slippage involves investments in hedge funds for which you lack data on current holdings of the hedge funds. Trying to draw conclusions from the historical pattern of the returns for the hedge fund and historical correlation of these returns with other positions is dubious, given the possibility that current holdings of the hedge fund may not resemble historical holdings. The case for treating these investments as illiquid rests both on limitations on hedge fund withdrawals and on lack of information on true exposure. A liquid proxy can be built by making reasonable inferences about the current style of the hedge fund, based on whatever public and private information you have available from the fund. Ineichen (2003) is an excellent starting point for explanations of the differing hedge fund styles (Chapter 5) and detailed examination of each style in relation to indexes of liquid investments (Chapters 6 through 8). A statistical approach to creating liquid proxies for each hedge fund style can be gleaned from Hasanhodzic and Lo (2007). The liquid proxy can be used for representing hedge fund investments in VaR and stress test calculations. Statistical analysis can then be performed on deviations between the liquid proxy and historical returns on hedge funds.

8.4.2 Choice of Liquid Proxy

A choice of liquid proxy is equivalent to a choice of what liquid market prices are utilized in modeling the illiquid instrument. Every model choice implies a liquid proxy, and every liquid proxy choice implies a model.

In evaluating whether a liquid proxy choice is correct, it is necessary to ask whether the implied model makes adequate use of available liquid market prices. This is closely related to one of the key questions in Derman (2001): "Has the model been appropriately calibrated to the observed behavior, parameters and prices of the simpler, liquid constituents that comprise the derivative?" This point can be most clearly made using a concrete example, which is discussed more fully in Section 12.4.2.

Consider an option written on a basket consisting of two stocks. You could choose two different ways to model this: (1) have a complete model of the price evolution of each of the two stocks individually and assume a correlation between them, or (2) directly model the price evolution of the basket. We'll call these the correlation model and the direct model, respectively. Assume that there are liquid market prices for options on the individual stocks but no liquid market prices for options on the basket or on the correlation, which is a fairly standard situation.

It can be argued that either model is a reasonable choice. In either case you will need input for a variable that cannot be observed in the market. In both cases, you have included all the sources of risk in your model.

But, as will be shown in our more detailed discussion in Section 12.4.2, where correlation is not expected to be too negative, the first model offers definite advantages in terms of making better use of liquid market prices. Options on the individual stocks will serve as effective partial hedges for the basket option, so utilizing the first model, which can be calibrated to current market quotes for these options, offers the following advantages over the second model:

- The correlation model implies a liquid proxy that represents the basket trade in the exposure reports for options positions on the two individual stocks. This encourages the use of liquid hedges.
- The correlation model will require valuation changes in the basket when there are changes in the implied volatility of the two individual stock options. The direct model does not require such valuation changes and so can result in stale valuations not fully reflecting the cost of unwinding some of the risk in the trade.
- The correlation model exhibits significantly lower statistical uncertainty of results compared with the direct model. This should permit lower required reserve levels and larger limits than could be allowed if the direct model was used.

Note that these advantages of the correlation model over the direct model are based on empirical, not theoretical, findings. As can be seen in the fuller discussion, if correlation levels are expected to be very negative or if the product were structured differently (for example, an option on the difference between the stock prices rather than on the basket), the advantages of the first model over the second would diminish to the point of indifference between the models.

In some cases, the liquid proxy used could consist of an instrument that is not itself liquid, but for which modeling in terms of a liquid proxy and simulation of the remaining risk have already been incorporated into the firm's risk management system. This is reminiscent of the quote from Derman in Section 8.2.6.2: "Derivative models work best when they use as their constituents underlying securities that are one level simpler and one level more liquid than the derivative itself." For example, in Section 12.3.3, in examining a liquid proxy for barrier options based on Peter Carr's approach, I advocate the use of illiquid binary options as part of the liquid proxy, noting that "techniques we have already developed for managing pin risk on binaries" in Section 12.1.4 "can now easily be brought into play."

8.4.3 Design of Monte Carlo Simulation

Modeling the differences between the actual trade and its liquid proxy must go all the way to final payout or to when the trade becomes liquid. Modeling must reflect the possibility that the model used for pricing and trading the product may be wrong. Modeling should be by Monte Carlo simulation to reflect a full range of possible outcomes and to generate a statistical distribution that can be used in assessing issues such as capital adequacy. Let us take these points in more detail:

- Don't assume that an illiquid instrument will become liquid—it may happen but it shouldn't be assumed. Another way of saying this: It is important that statistical analysis of the distribution of parameters be based on actual market observations and not on derived values, since the derived values often themselves contain modeling assumptions subject to error. For example, if a given market is currently liquid only out to seven years, use only quotations out to seven years in your hedging simulations; 10-year quotations derived by extrapolation should not be used. This is analogous to the point made earlier in Section 8.4.1 about avoiding circular reasoning in model validation.
- Statistical assumptions used in determining distributions should not be constrained by any assumptions made within the valuation model. For example, the valuation model may assume a normal distribution of a factor because it is computationally simple and the increase in accuracy from using a different distribution can be shown not to be worth the added investment. This would not in any way justify assuming that the corresponding input variable is normally distributed in a model-testing simulation, since the computational trade-offs motivating the model-building decision do not apply to the model-testing calculation.
- Independent reviewers must be careful not to rely on statistical analysis prepared by traders. It is notoriously easy to employ data-mining techniques to find statistical proofs of nearly any relationship by selecting the right historical data set. Statistical controls, such as careful

discipline about segregating historical data into sample periods to fit parameters and out-of-sample periods to test results, are useful, but can still be defeated by sufficiently industrious data mining. It is better to have truly independent analysis, even at the risk of inaccuracy (on the side of conservatism) from lack of insider information.

- Use of Monte Carlo simulation allows for generation of a full statistical distribution of results, which can be very useful for issues such as determining capital adequacy on illiquid positions. This is a necessity if the capital adequacy proposal of Section 8.4.4 is to be followed.
- It must be emphasized that any statistical distribution involving tail risks requires subjective probability judgments (as discussed in Section 1.3). Still, the basic approach of insisting that simulation be of hedge trades involving liquid instruments, and that simulation go all the way to the point at which the original position becomes liquid, means that there will be a lot of historical liquid pricing data that can be utilized in forming these probability judgments. In essence, while illiquid instruments cannot be fully evaluated based on current liquid prices, they can be evaluated based on the future evolution of liquid prices. For illustrative examples, see Sections 10.2.2, 12.1.4, 12.3.3, 12.4.2, 12.5.2, and 13.4.3.
- Use of Monte Carlo simulation avoids the overstatement of risk that can result from more formulaic risk calculations. For example, if the desire is to reserve to a 90th percentile degree of certainty, using 90th percentile values of the distribution of two or more input parameters will likely result in a far greater than 90th percentile degree of certainty in the reserve. In a Monte Carlo simulation, many reruns of the valuation model are made based on sample points chosen randomly from the assumed distribution of each nonliquid variable, and with explicitly assumed correlations between variables. The 90th percentile of model outputs can then be estimated.

Derman recommends a full simulation that includes both underlying behavior and trader hedging strategy. Section 11.3 contains an example that comes close to Derman's proposed full simulation: a Monte Carlo simulation of dynamic hedging of a less liquid option (less liquid because of a nonstandard strike). Sampling over the simulation paths yields a statistical distribution of the differences between the payout on the option and the costs of the hedge.

Derman's recommendation of a full simulation including trader hedging strategy represents an ideal that may sometimes be difficult to achieve in practice. In the simulation in Section 11.3, a full simulation is possible because the assumed trader strategy is very simple, just varying the delta hedge of the underlying forward. Trader strategies that involve changes in options positions are much more difficult to simulate, because a full specification of the volatility surface is required at each node of the simulation. An illustration of this point can be found in the discussion of barrier options in Sections 12.3.2 and 12.3.3.

When a full simulation is not practical, then I still believe that a simulation should be done, but computation can be simplified by restricting hedging strategies. Easier implementation comes at a cost of greater conservatism, since the full range of possible trader hedging strategies will not be captured. The simulations that I refer to in the next-to-last bullet point of the preceding list can serve as helpful paradigms.

8.4.4 Implications for Marking to Market

Choosing a good liquid proxy, following the guidelines of Section 8.4.2, should assure that illiquid positions are marked to market to reflect changes in liquid market prices. To illustrate with the example used in that section, when there is a change in the implied volatility of one of the two stocks in the basket, it will be immediately reflected in the marking to market (MTM) of the liquid proxy and hence in the MTM of the basket option, which consists of the sum of the MTM of the liquid proxy and the reserve for the difference between the basket option and the liquid proxy.

But should there also be an adjustment to the MTM of the illiquid position based on new information about parameters that cannot be sourced from a liquid market? Continuing with the same example, the question would be whether to change the MTM of the basket option based on new information about correlation between the two stocks. My answer would be that this should be done only very rarely. We have classified the correlation parameter as one that has no liquid market pricing source, so where would frequent updates be coming from? There are two possible sources:

1. Analysis of historical price data has led to a change in estimates of the correlation to be used. But this will only occur infrequently—if the correlation has been estimated from a long data history, then it will usually take months of new data before conclusions will change significantly.
2. There is evidence that the price being charged customers for correlation has changed. But since this is not a liquid market at which risk can be exited, the argument for making immediate use of such new data is not nearly as strong as it is for liquid instruments.

In both cases, new information on correlation might ultimately impact the reserve for the difference between the basket option and the liquid proxy, and thereby impact the total MTM of the basket option. But in both cases, you would expect to see this impact take place infrequently. In fact, I would argue for designing reserve calculations in a way that would make such changes extremely infrequent. For example, in this case, calculate the reserve based on an extremely unlikely *level* of correlation as opposed to an extremely unlikely *change* in correlation from the long-term average. That makes it less likely that new information about a shift in the long-term average will require a change in reserve level.

The reason I want to make reserve changes infrequently is that I don't think reserve level changes provide good incentives to traders and marketers. Changes in MTM of liquid instruments provide good incentives for exiting positions—either because stop-loss limits are being breached or because accumulating losses cause traders to rethink the desirability of positions (this includes changes in MTM of liquid proxies, which can trigger hedging actions in liquid markets). But changes in reserve levels won't provide much incentive to exit existing positions, since the illiquidity of the instrument makes such exits very difficult. It is true that raising reserve levels may send a signal to marketers to be more reluctant to book new trades, which might argue for raising reserve levels on new trades but not on existing ones.

Even if you are convinced this policy makes good risk management sense, you still might be reluctant to have it guide the MTM reporting of the firm. Financial controllers, independent accounting firms, and regulators all tend to be suspicious of policies that involve high reserve levels that shield reported earnings from fluctuation; it looks like an attempt to smooth reported earnings. Let me make the following points concerning this:

- I believe that the policies I am advocating here represent an accurate picture of what is known about earnings. The true earnings on illiquid positions are often not known until the trade matures. A highly conservative reserve level is therefore justified, and it is unreasonable to expect much new information to arrive from outside sources; the real information will come over time as the trade matures. There are exceptions—new information that would change your outlook for the whole distribution of an illiquid input. An example would be long-term default rates on home mortgages in 2007 when new information on deteriorating underwriting standards would have impacted reserve levels that were previously viewed as prudently conservative.
- Reserving policies can be designed to assure independence and shielding from manipulation that attempts to use reserve levels to smooth earnings. See Section 6.1.4.
- These policies could help to deal with some of the concerns being expressed about the harmful impact MTM policies are having on bank

management (see the reducing procyclicality discussion in Section 5.5.8.1). MTM losses for liquid instruments encourage banks to shed volatile assets, but MTM losses for illiquid instruments, since the banks can't shed the assets, result in a need to raise new capital, often in economic environments that are the most challenging for raising capital, leading to paralysis of the banking system. (This is discussed in greater detail, in the context of the 2007–2008 crisis, in Section 5.3.2.) My proposal causes large reserves to be taken up front, when the environment is still favorable for raising capital, and then releases the reserves, and hence frees capital for new investments, as the existing investments unwind.

- I would definitely advocate strong controls on the use of this accounting policy, only permitting it for positions the firm designates at the time of creation as illiquid.
- I have experience with a policy close to the one described working in practice over a several-year period, from 1996 to 2003, at Chase and JPMorgan Chase, with the full knowledge of risk managers, financial controllers, independent accountants, and regulators. Reserve levels established were sufficiently conservative that they almost always proved adequate at an individual product level, and always proved more than adequate at an aggregated firm level.
- In the current environment, following the debacle of 2007–2008, it may no longer be possible to get independent accountants and regulators to go along with a policy like this; it requires more trust of the motivations of firm risk management than may now be achievable. In that case, I think risk managers should argue for keeping an internal set of accounts that most accurately reflects the economics of a business, even where this diverges from external reporting.

8.4.5 Implications for Risk Reporting

In Section 8.3 we noted that for liquid instruments risk reports only need to look at exposures measured against current market prices, since future exposures due to changes in price levels can always be reduced utilizing the liquidity of the instrument. This approach will not work for illiquid instruments. To take an example, discussed at greater length in Section 12.1.4, a binary option might currently show very little gamma exposure but might have an unacceptably large gamma in the future if prices are close to the strike level when little time is left until option expiry. You can't just wait to see if this will happen, since if it does you can't count on being able to extinguish the risk by selling the digital option. You need to deal with this contingency at the time you are considering creating the option.

One way of handling this is to run risk reports at the time you are considering creating the position that look at a range of future possible price levels for future dates. Acceptability of possible future risk exposures are evaluated as part of the decision-making process for taking on the position.

Another way of handling this is to make sure that the liquid proxy and simulation methodology of Sections 8.4.2 and 8.4.3 adequately control possible future exposures. Continuing with the binary option example, you would make sure that the call spread liquid proxy chosen can only give rise to reasonable future gammas, by making sure that there is a sufficiently wide gap between the strikes of the call spread. As you will see in Section 12.1.4, widening the gap between the strikes will lead to more uncertainty in the simulation and hence higher reserve levels and tighter limits, but this should be viewed as a necessity for controlling future gamma exposure.

8.5 TRADING MODELS

When a model is being used as part of a trading desk's decision-making process, it clearly requires internal model review by the model creators and users. For the model validation part of this process, it is particularly important to review how the model relates to the prevailing model being used in the market and to try to anticipate evolution of the prevailing market model, as argued in Section 8.2.6.3. The question I want to examine here is whether such models also require an external review by an independent group if the model is to be used only for trading decisions and not for the firm's official valuations and measurement of risk.

Major trading losses are frequently ascribed to the firm having the wrong model. What is often unclear in these claims is whether "having the wrong model" just means making incorrect forecasts about the future direction of market prices or if it means misleading the firm's traders and managers about the nature of positions being taken. A good illustration is the discussion in Section 4.2.1 of whether the reliance by Long-Term Capital Management (LTCM) on models should be viewed as a primary cause of the collapse of the fund.

Any firm engaged in making markets or investing funds must take positions whose profit or loss will depend on the correctness of forecasts of moves in market prices. Different strategies will be tied to different price relationships. Some depend on overall market direction, whereas others depend on the relative price of related assets; some depend on getting a long-term trend right, whereas others depend on correctly anticipating short-term moves. However, traders will always need to make judgments about an uncertain future, and firm managers in turn will always need to make judgments about how much of a risk of loss they will allow a trader to take in exchange for a possible gain. When making this assessment, management will be guided by evidence of prior accuracy of the trader's forecasts.

Nothing in the last paragraph will be altered by whether a trader uses a model as a computational aid in forecasting, unless perhaps management is lulled into a false sense of security by believing that the use of a model lessens the chance of errors in trading judgment. However, if a model results, either purposely or inadvertently, in misleading traders and managers about the relationship between positions being taken and the size of possible losses, then the accusation that model error resulted in the loss is far more plausible.

For example, a spot foreign exchange (FX) trader could be using a very complex model when deciding which positions to take. This could even extend as far as program trading, in which a computer actually issues the buy and sell instructions based on model output. However, spot FX positions can easily be valued based on external quotes, and position size is extremely easy to understand without the aid of models (see the discussion in Sections 9.1 and 9.2). So it is easy for management to see what the profit and loss (P&L) is every day and to cut the risk if P&L performance has been poor. Thus, the modeling does not have any of the dangers of hidden risk, such as Ponzi schemes (see Section 2.2). No FX trader would dream of asking to report more profits this year because he can "prove" that his model (or trading style) will work better next year than this year.

When I was in the position of managing the independent model reviews for a firm, I argued strongly against my group reviewing the validity of models that were being used only for trading decisions. Partly, this was an attempt to conserve resources for what I viewed as the more important task of validating models used for valuation and risk measurement. But even more, I was concerned that traders would use model validation by my independent reviewers as a stamp of approval that would discourage critical review of trading strategies by senior management. I argued that since we weren't being asked to review trading strategies that didn't involve models, the use of a model did not transform us into experts on trading strategy. In particular, how could we obtain the insider knowledge that could allow us to anticipate evolution of the prevailing market model? This is the position I advocated in the first edition of this book, but on reflection, I would reconsider my previous stance.

When position limits are being set and when actions following a stop-loss limit overage are being reviewed, there is no question that traders will utilize results from their trading models to make their case to senior managers. Since senior managers will not have the time or, usually, the skill set to form their own judgment of these models, it is only by having independent reviewers look at the models that an effective challenge to trader claims can be prepared. Independent reviewers must make clear the limited scope of their review, but can certainly raise issues concerning possible cherry-picking of historical data or reasons why shifts in the economic environment might bring conclusions based on historical data into question. These challenges may prove of value to traders as well as senior managers. And certainly, independent review of model mechanics—the model verification of Sections 8.2.3, 8.2.4, and 8.2.5—can add value.

FRB (2011) seems quite clearly to endorse independent review of trading models. Its Section III, which examines the criteria for which models need to be subject to the review standards of the document, states, "Models meeting this definition might be used for analyzing business strategies, [and] informing business decisions" and "The definition of *model* also covers quantitative approaches whose inputs are partially or wholly qualitative or based on expert judgment, provided the output is quantitative in nature."