Theory and Practice of Model Risk Management

Riccardo Rebonato
Quantitative Research Centre (QUARC) of the Royal Bank of Scotland
Oxford Financial Research Centre – Oxford University

1 – Why Model Risk Matters

Model risk is a topic of great, and growing, interest in the risk management arena. Financial institutions are obviously concerned about the possibility of direct losses arising from mis-marked complex instruments. They are becoming even more concerned, however, about the implications that evidence of model risk mismanagement can have on their reputation, and their perceived ability to control their business.

Model risk inhabits, by definition, the opaque area where the value of instruments cannot be fully and directly ascertained in the market. The valuation of these products must be arrived at by means of marking-to-model, and therefore always contain a subjective component. In these days of heightened sensitivity to aggressive or opaque accounting, the ability to rely on sound practices to control model risk can have share-price implications well beyond the monetary value of the mis-marked positions. It is also for this reason that financial institutions place increasing importance on the effective management of model risk. Unfortunately, there is a widespread lack of clarity as to what model risk management should achieve, and about which tools should be used for the purpose. In this chapter I intend to answer these questions, to provide theoretical justification for my views and to suggest some practical ways to tackle the problem.

2 - Theory

2.1 – What is Model Risk?

There are several distinct possible meanings for the expression 'model risk'. The most common one refers to the risk that, after observing a set of prices for the underlying and hedging instruments, different but identically calibrated models might produce different prices for the same exotic product. Since, presumably, at most one model can be 'true', this would expose the trader to the risk of using a mis-specified model. This is the angle explored, for instance, by Sidenius (2000) in the interest-rate area: in his study of model risk, significantly different prices were obtained for exotic instruments after the underlying bonds and (a subset of) the underlying plain-vanilla options were correctly priced. The question has intrinsic interest, but it is not clear what use can be made of this information, unless additional assumptions are made about the behaviour of the trader after the transaction. Shall we assume that he will hedge his position on the basis of the model recommendations? Should we assume that the trader will hedge at all? What constraints and limits (VaR, vega, etc) is the trader subject to? Will the trader recalibrate her model using historical or market data? Under what accountancy regime will the value of the position be recorded?[i]

Another common interpretation of the term 'model risk' is the following. Let's assume that the trader, after making a price using a given model in a derivative product (complex or plain-vanilla), follows the dynamic hedging strategy prescribed by the model. How well will the trader fare by the expiry of the derivative? Green and Figlewski (1999), answer this question assuming that the trader recalibrates her model based on historical data. Hull and Suo (2001) and Longstaff, Santa-Clara and Schwartz (2001) address the same topic assuming daily re-calibration of the model to market prices. What all these approaches have in common is the implicit or explicit identification of model risk with the risk that a complex product might be bought or sold for the 'wrong' price, or that losses will be incurred because of an 'incorrect' hedging strategy.

These are certainly interesting questions, and they are probably the most relevant ones from the trader's perspective. For a single exotic desk, in fact, selling optionality too cheaply is likely to cause an uneven but steady haemorrhaging of money out of the book, and can ultimately cause the demise of the trader. From the perspective of the financial institution, however, losses incurred in this fashion are unlikely to be of such magnitude to have a major bottom-line or reputational impact. In this respect the exotic trader is, from the bank's perspective, little different than a proprietary trader in, say, spot FX. *As long as the exotic trader's positions can be accurately marked to market*, the risk for a financial institution arising from trading in complex options using a mis-specified model can be controlled with the same non-VaR risk-management measures applied to other traders (eg, stop losses). 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 products. 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[ii].

In this chapter I will therefore argue that, from the perspective *not of the trader,* but of a financial institution the most relevant question is the following: if the price of a product cannot be frequently and reliably observed in the market, how can we ascribe a price to it in between observation times in such a way as to minimize the risk that its book-and-records value might be proven to be wrong? This question directly leads to the following definition (adapted from Rebonato 2001):

> *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.*

This is the meaning of 'model risk' that I intend to cover in this chapter.

There are several points worth noticing about the definition. First of all, from the definition it follows that, if reliable prices for all instruments were observable at all times, model risk would not exist. The market prices might be unreasonable, counterintuitive, perhaps even arbitrageable, but they constitute the only basis for a robust mark-to-market.

Another important observation is that the instrument in question may, but need not, be complex or held off-balance sheet: model risk has often been associated with complex derivatives products, but a deeply out-of-the money call (the plain-vanilla option *par excellence*) and an illiquid corporate bond (a prototypical on-balance-sheet instrument) can both present substantial model risk. What both these instruments have in common is that the value at which they would trade in the market cannot be readily ascertained via screen quotes, intelligence of market transactions or broker quotes. A model must therefore be used in order to associate on a daily basis a value to these instruments for books-and-records purposes.

The last proviso is essential, and brings me to the last important observation about the definition above: in my discussion, model risk arises not because of a discrepancy between the model value and the 'true' value of an instrument (whatever that might mean), but because of a discrepancy between the model value and the value that must be recorded for accounting purposes. From the perspective I have chosen, model risk is therefore intimately linked to trading-book products (that must be marked to market on a daily basis). I shall argue below that this need of marking to market daily, together with institutional and regulatory constraints, have profound implications for the price of these opaque instruments. This does not mean that the value of instruments held on the banking book cannot be arrived at using (sometimes very complex) models. The dynamics between 'fundamental value', price adjustment and modelling practice is however so different, that, in the present context, it is counterproductive to analyze trading-book and banking-book products from the same perspective.

2.2- Isn't There Just One Value? Efficient Markets Revisited

In the previous section I have repeatedly referred to 'true' and 'fundamental' value on the one hand, and to mark-to-market value on the other as two potentially different concepts. I have also hinted that the value of an instrument might have to be looked at differently depending on whether it resides on the trading or the banking book. Doesn't this fly in the face of financial theory (that equates price and value), and of the efficient-market hypothesis in particular? The whiff of sulphur associated with these claims is so strong, that it can't be ignored. I maintain that understanding why 'value' can have different meanings in different contexts is crucial to model risk, and well worth the effort. Let's rehearse the standard argument first (Shleifer 2000).

The efficient market hypothesis (EMH) can be formulated in forms of wider and wider applicability. The most radical form requires that all economic agents are fully informed and perfectly rational. If they can all observe the same price history they will all arrive at prices by discounting the expected future cashflows from a security at a discount rate dependent on the uncertainty of the security and their risk aversion. In this sense the value of the security is said to embed all the information available in the market, its value to be linked to 'fundamentals', and markets to be informationally efficient. All securities are fairly priced, excess returns simply reflect excess risk, and five-dollar banknotes cannot be found lying on the pavement.

A weaker form of market efficiency (but one which arrives at the same conclusions) does not require all economic agents to be rational, but allows for a set of investors who price securities on sentiment or with imperfect information. Will this affect the market price? Actually, one can show that as long as the actions of the uninformed, irrational investors are random ('uncorrelated'), their actions will cancel out and the market will clear at the same prices that would obtain if all the agents were perfectly rational.

Surely, however, the zero-correlation assumption is far too strong to be swallowed by anybody who has witnessed the recent dot.com mania. The very essence of bubbles, after all, is that the actions of uninformed, or sentiment-driven, investors are just the opposite of uncorrelated. If this is the case, supply and demand, rather than fundamentals, will determine the price of a security. Is this the end of the efficient market hypothesis? Not quite. Let there be irrational *and* co-ordinated investors. As long as there also exist rational, well-informed agents who can value securities on the basis of fundamentals, price anomalies will not persist. These pseudo-arbitrageurs[iii] will in fact buy the irrationally cheap securities and sell the sentimentally expensive ones, and by so doing will drive the price back to the fundamentals. Whether it is due to irrationality and sentiment or to any other cause, in this framework excess demand automatically creates extra supply, and vice versa. Therefore, as long as these pseudo-arbitrageurs can freely take their positions, supply and demand will not affect equilibrium prices, these will again be based on the suitably discounted expectation of their future cashflows (the 'fundamentals') and the efficient-market hypothesis still rules[iv].

It is important to stress that the EMH is not only intellectually pleasing, but has also been extensively tested and has emerged, by and large, vindicated. The 'by-and-large' qualifier, however, is crucial for my argument. In the multi-trillion market of all the traded securities, a theory that accounts for the prices of 99.9% of observed instruments can at the same time be splendidly successful, and yet leave up for grabs on the pavement enough five-dollar notes to make a meaningful difference to the end-year accounts and the share prices of many a financial institution. This is more likely to be so if the instruments in questions are particularly opaque and if the reputational amplifying effect mentioned in the introductory section is at play. The possibility that the pseudo-arbitrageurs might not be able always to bring prices in line with fundamentals should therefore be given serious consideration.

2.3 – Pseudo-Arbitrageurs in Trouble

What can prevent pseudo-arbitrageurs from carrying out their task of bringing prices in line with fundamentals? To begin with, these pseudo-arbitrageurs (hedge funds, relative-value traders, etc) often take positions not with their own money, but as agents of investors or shareholders. If the product is complex, and so is the model necessary to arrive at its price, the ultimate owners of the funds at risk might lack the knowledge, expertise or inclination to asses the fair value, and will have to rely on their agent's judgement. This trust, however, will not be extended for too long a period of time, and certainly not for many years[v]. Therefore, the time span over which securities are to revert to their fundamental value must be relatively short (and almost certainly will not extend beyond the next bonus date). If the supply-and-demand dynamics were such that the mis-priced instrument might move even more violently out of line with fundamentals, the position of the pseudo-arbitrageur will swing into the red, and the 'trust-me-I-am-a-pseudo-arbitrageur' line might rapidly loose its appeal with the investors and shareholders.

Another source of danger for relative-value traders is the existence of institutional and regulatory constraints that might force the liquidation of positions before they can be shown to be 'right': the EMH does not know about the existence of stop-loss limits, VaR limits, concentration limits etc.

Poor liquidity, often compounded with the ability of the market to guess the position of a large relative-value player, also contributes to the difficulties of pseudo-arbitrageurs. Consider for instance the case of a pseudo-arbitrageur who, on the basis of a perfectly sound model, concluded that traded equity implied volatilities are implausibly high, and entered large short-volatility trades to exploit this anomaly. If the market became aware of these positions, and if, perhaps because of the institutional constraints mentioned above, the pseudo-arbitrageur had to try to unwind these short positions before they had come in the money, the latter could experience a very painful short squeeze[vi].

Finally, very high information costs might act as a barrier to entry, or limit the number, of pseudo-arbitrageurs. Reliable models require teams of quants to devise them, scores of programmers to implement them, powerful computers to run them[vii] and expensive data sources to validate them. The perceived market inefficiency must therefore be sufficiently large not only to allow risk-adjusted exceptional profits after bid-offer spreads, but also to justify the initial investment.

In short, because of all of the above the life of the pseudo-arbitrageur can be, if not nasty, brutish and short, at least unpleasant, difficult, and fraught with danger. As a result, even in the presence of a severe imbalance of supply or demand, relative-value traders might be more reluctant to step in and bring prices in line with fundamentals than the EMH assumes.

2.4 – Why does it matter for model risk?

Let us look more closely at the features of the trades that, on the basis of the above discussion, one might most plausibly expect to be priced out of line with their 'true' value. One feature was their complexity (with its double repercussions on the agency relationship and on the information costs); another was their poor liquidity and transparency; a third was the existence of limit structures that might cause the unwinding of trades before they can be shown to be right. All of these features perfectly apply to those products for which model risk is most acute. If this analysis is correct, however, it has profound implications as to how to control model risk, and, ultimately, as to what the source of model risk is.

If you belong to the EMH school, 'things can only get better', and do so quickly: markets will always become more efficient and liquid, models truer to life, and their truth will always be swiftly recognized[viii]. Therefore model risk is simply the risk that the model currently used to mark positions might not be sophisticated and realistic enough. For an EMH believer, finding a 'better' model is the surest way to ensure that the marks of a financial institution will be in line with the (efficient) market prices. An important corollary of this view is that there should be no distinction between the model a front office trader would want to use to price a given instrument, and the model to be used for recognizing its value for books-and-record purposes.

Much of the existing literature on model risk tends to implicitly endorse the view of the world underpinned by the EMH hypothesis, placing the emphasis as it does on the possibility that the model used might not be 'accurate' or 'sophisticated' enough. In this strand of works, creating a 'better' model is seen as the best defence against model risk. The model validation department of an institution that subscribes to such a view typically has enshrined in its 'mission statement' the requirement that the hypothesis underlying the model used for mark-to-model should be tested for 'soundness and validity'. If asked to choose between models to be used for books-and-records valuation, the same model validation unit will generally go for the theoretically more appealing and sophisticated one. If one keeps in mind that the ultimate goal of marking to model is to guess the price where the same product would actually trade in the market, the corollary of this policy is that the 'better model' and the 'market-chosen model' (ie, value and price) must coincide.

If, on the other hand, one believes that the EMH, while generally valid, might not apply to many of the products for which model risk is relevant, the situation is very different. Over the time scales relevant for model risk, the market might stray even further away from the righteous path. The model that will prevail tomorrow (or, rather, the model that will reproduce tomorrow's prices) need not necessarily be more realistic than today's model. Anecdotal evidence suggests that, for instance, after the liquidity crunch associated with the Russia crisis, traders reverted to pricing Bermudan swaptions with cruder and more conservative models, such as the Black-Derman-and-Toy, rather than with the indubitably more sophisticated and realistic LIBOR market models).

This state of affairs creates a fundamental tension between front-office traders and risk mangers insofar as model risk is concerned.

2.5 – Traders and risk managers part ways

Within the EMH framework the goals of traders and risk managers are aligned: a superior model will bring the trader a competitive advantage, and will be recognized as such by the market with very little time lag. From this point of view, an accurate mark-to-market is purely a reflection of the best information available, and 'true' (fundamental) value, market price and model price all coincide. It therefore makes perfect sense to have a single research centre, devoted to the study and implementation of the best model, which will serve the needs of the front-office trader and of the risk manager just as well. Looked at from this angle, model risk is simply the risk that our current model might not be good enough, and can be reduced by creating better and better models that will track the monotonic, if not linear, improvement of the markets' informational efficiency.

If we believe, however, that pseudo-arbitrageurs might in practice be seriously hindered in bringing all prices in line with fundamentals model risk looks very different. Across both sides of the EMH divide there is little doubt that what should be recorded for books-and-records purposes should be the best guess for the price that a given product would fetch in the market. For the EMH sceptic, however, there is no guarantee that the 'best' available model (ie, the model that most closely prices the instrument in line with fundamentals) should produce this price. Furthermore, there is no guarantee that the market, instead of swiftly recognizing the error of its ways, might not stray even more seriously away from fundamentals.

Ultimately, whenever a trader enters a position, he must believe that, in some sense, the market is 'wrong'. For the risk manager concerned about model risk, on the other hand, the market must be right by definition. For the EMH believer the market can only be wrong for a short period of time (if at all), so there is no real disconnect between the risk-manager's price, and the trader's best price. For the EMH-sceptical risk managers, on the other hand, there is an irreconcilable tension between front-office pricing, and the risk management model price.

2.6 – The parable of the two volatility traders

To illustrate further the origin of this tension, let's analyze a stylised but instructive example. Two plain-vanilla option traders (one working for Efficient Bank, and the second for Sceptical Bank) have carefully analysed the volatility behaviour of a certain stock, and both concluded that its level should be centred around 20%. The stock is not particularly liquid, and the brokers' quotes, obtained with irregular frequency, do not deviate sufficiently from their estimate to warrant entering a trade. One day, without anything noticeable having happened in the market, the volatility quote appears of 10%. Two huge five-dollar notes are now lying on the floor, and both traders swiftly pick them up. Both traders intend to crystallize the value of the mis-priced option by engaging in gamma trading, ie they are both long gamma and will dynamically hold a delta-neutralizing amount of stock, as dictated by their model *calibrated to 20% volatility*.

Life is easy for the Efficient Bank trader. He will go to his risk manager, convince him that the model used to estimate volatility is correct and argue that the informationally efficient market will soon adjust the implied volatility quote back to 20%. This has important implications. First of all, the profit from the trade can be booked immediately: the front office and risk management share the same state-of-the-art model, and both concur that the price for the option in line with fundamentals should be obtained with a 20% volatility. Furthermore, should another five-dollar bill appear on the pavement the trader of Efficient Bank will have every reason and incentive to pick it up again.

The coincidence of the front office and middle office models has yet another consequence. The trader works on a 'volatility arbitrage' desk, and his managers are happy for him to take a view on volatility, but not to take a substantial position in the underlying. They have therefore granted him very tight delta limits. This, however, creates no problem, because his strategy is to be delta-neutral at every point in time and to enjoy the fact that he has bought (at 10%) 'cheap convexity'. Crucially, in order to crystallize the model profits from trade, he will engage in a dynamic hedging strategy based on the superior model (calibrated with a 20% volatility), not on the temporarily erroneous market model. Since middle-office again shares the same model, the risk manager calculates the delta of the position exactly in the same way as the trader, and therefore sees the whole portfolio perfectly within the desk's delta limits (actually, fully delta neutral).

Life is much harder for the trader at Sceptical Bank. She also works on a volatility arbitrage desk with tight delta limits, and her middle office function also recognizes that the model she uses is sound and plausible and concurs that the market must be going through a phase of summer madness. The similarities, however, virtually end here. Her risk management function does not believe that a superior model must be endorsed by the market with effectively no delay, and therefore is not prepared to recognize the model value implied by the 10% trade as an immediate profit. A model provision will be set aside. Since the trader will not be able to book (all) the model profit upfront, she will have to rely on the profit dripping into the position over the life of the option as a result of trading the gamma. This process will be relatively slow, the more so the longer the maturity of the option. During this period the trader is exposed to the risk that another ‘rogue’ volatility quote, say at 5%, might even create a negative mark-to-market for her position. Her reaction to a second five-dollar bill will therefore be considerably different from that of her colleague at Efficient Bank. Furthermore, in order to carry out her gamma-trading programme she will have to buy and sell delta amounts of stock based on her best estimate of the ‘true’ volatility (20%). Middle office, however, who have observed the 10% trade, uses the model calibrated with the lower volatility to calculate the delta exposure of the trade, and therefore does not regard her position as delta-neutral at all. She utilizes more VaR than her colleague, might soon hit against her delta limit, and, if her trading performance is measured on the basis of VaR utilization, she will be deemed to be doing, on a risk-adjusted basis, more poorly than her colleague.

This parable could be expanded further, but the central message is clear: different approaches to model risk management can generate very different behaviours and incentives for otherwise identical traders. The natural question is: which approach better serves the interests of a bank? In the long run, which bank will be more successful, Efficient Bank or Sceptical Bank?

2.7 – The Role of Liquidity and of Risk Aversion

In order to answer this question at least two important elements, absent in the hyper-stylised example above, must be brought into the discussion: the role of liquidity and of risk aversion in the formation of prices. In the following discussion the two concepts are linked, but, for simplicity, it is simpler to analyze them in turn. Let us start with liquidity. The analysis underpinning EMH, which automatically equates the market price with the informationally most efficient price, pays little attention to the role of liquidity. Yet model risk is, by definition, most acute for illiquid products (that trade, and whose prices are therefore observed, relatively infrequently). In situations of market distress investors are happy to pay a premium to hold securities which can be sold easily at a readily ascertainable price. The trading dynamics in the months that followed the Russia default and the LTCM crisis of 1998 have often been described as a flight to quality, but a careful observation of the instruments that were sought after or disposed of would make the name ‘flight to liquidity’ more appropriate: US Treasuries, of course, were very popular, but, since they enjoy both safety and liquidity, it is difficult to tell which features the investors were after. The unprecedented yield discrepancies that occurred at the time between on-the-runs and off-the runs Treasuries (both backed by the full faith of the US Treasury) gives a first indication that it was liquidity that loomed large in the investors’ mind. Another popular product of the days were Danish mortgage-backed securities. These are bonds of the highest credit quality, (no default has been recorded in over 300 years) but with an outstanding stock that cannot compare with their US cousins (and therefore with more limited liquidity). In some of most the troubled days of the late summer of 1998 they traded for as little as 60c in the $ on an option-adjusted basis. These, and similar, observations suggest that investors not only pay a premium for liquidity, but that this liquidity premium can be strongly state- and time-dependent.

This poses some crucial questions for model risk management. Looking closely again at Danish mortgage-backed-securities, the typical trade of the day, that many hedge funds and several banks were engaged in, was simple: they would buy a mortgage-backed security, purchase a Bermudan swaption on the amortizing principal to protect themselves against pre-payment risk, fund themselves at LIBOR and still lock in a small but almost risk-free profit. Sound as the strategy might have been, it took into no account that, in order to put it in place, the pseudo-arbitrageurs had to take a long position in an instrument (the Danish mortgage bond) with adequate liquidity in normal market conditions, but prone to trade at a liquidity discount in stress situations. Arguably, the ‘fundamentals’ (the suitably discounted expected cashflows from Danish mortgagors) had not changed because of the Russia/LTCM turmoil; yet, because of the high value placed on liquidity, the market price of the securities did. Those market players who were subject to daily marking to market of their trades were quickly hitting against sop loss or other limits, or were facing margin calls, and were forced to liquidate the self-same positions, further contributing to the negative price dynamics. The traders at Sceptical Bank had to cut their positions first, those at Efficient Bank held on to their ‘sound’ positions longer, but ultimately found themselves no better able than King Canut to stem the tide, and liquidated at larger losses.

The second, and related, element missing form the previous discussion is the stochastic nature of risk aversion. It is well known that, when a pricing model is calibrated to market prices, the resulting 'implied values' (reversion levels, jump frequencies, reversion speeds, etc) are not the econometrically observed ones, but the ones that would prevail under the pricing measure (ie they are risk-adjusted). This is readily handled within the EMH framework, because the expected cashflows are discounted at a rate that takes risk aversion into account. If this risk aversion, however, can vary stochastically over time, the knowledge that 'some' future degree of risk aversion (unknown today) will prevail tomorrow, such that the future prices will be accurately accounted for, is of little comfort. Typically, for the sake of tractability, financial models tend to make the market price of risk either a constant (Vasicek), or a deterministic function of time (as implicitly happens in the Hull-and-White model), or a simple function of the state variable (CIR). The very notation typically used to denote the market price of risk, $\lambda(t)$, tends to suggest that it should be a deterministic function of time. Unfortunately, there is ample evidence that risk aversion changes in an unpredictable way (stochastically) over time: equity smiles, which can be plausibly linked to the risk-adjusted jump frequency, suddenly made their appearance after the 1987 stock market crash: sellers of equity put options became more 'afraid' of jumps after October 1987 than they were before. Similarly, non-monotonic smiles in the caplet and swaption volatility surfaces appeared after the turmoil that followed the Russia events. During the same period swap spreads widened beyond what any risk-neutral estimation of bank default risk might suggest plausible.

The consequences of this for model risk are far-reaching, and we do not need to invoke any shortcomings of the EMH to understand them. Let us assume that a trader has a perfectly correct model (ie a model that, *on a risk-adjusted basis*, prices perfectly in line with fundamentals), and that supply and demand does not appreciably influence prices. Still, if the compensation the market requires for the various sources of risk (jump risk, volatility risk, swap spread risk, etc) varies stochastically over time, the same trader cannot be confident that today's calibration to the market prices (which embeds *today's* risk aversion) will be valid tomorrow. As we have said, if the relevant market prices were always readily observable, model risk would not exist. But, in the absence of frequent reliable quotes, a model's ability to recover *today*'s prices would be no guarantee that next months' prices will also be correctly recovered. Of course, leaving aside again the possible shortcomings of the EMH, if we had a 'correct model' and we included in the state variables available for its calibration also the future (stochastic) risk aversion, we could hope to obtain at least a conditional distribution for the future market prices. Unfortunately, daily time series of risk aversion to jumps, volatility, spread risk, etc, from which a future joint probability distributions of these variables and of the 'traditional' observable risk factors could be obtained, do not belong to the world of the here-and-now.

2.8 – The Technology of Option Pricing

There is yet another reason why the way the model price of a given product is calculated might change in the future. The idea that the price of a financial instrument might be arrived at using a complex mathematical formula is relatively new, and can be traced back to the Black-and-Scholes (1973) formula. Of course, formulae were used before that for pricing purposes, for instance in order to convert the price of a bond into its gross redemption yield. Even when no closed-form solutions existed, as is indeed the case for the gross redemption yield, these pre-Black-and-Scholes early formulae generally provided totally transparent transformations from one set of variables to another, and did not carry along a heavy baggage of model assumptions.

Bona fide model-based pricing, on the other hand, can be characterized by the fact that the reliability of a pricing formula hinges on the applicability and realism of a complex set of assumptions about the market (eg, its completeness, the ability to enter short positions, the absence of taxes or transaction costs, etc.) and about the process for the underlying state variables (eg, continuous or discontinous paths, etc). Given the relatively short history of model-based approach to pricing, it is therefore not surprising that the models used by the trader community have evolved rapidly and radically. (See Rebonato 2002 for a survey of the evolution of interest-rate modelling). Very often, however, the adoption of new models has been driven not by their superior explanatory power, but by concomitant, and difficult-to-predict, technological and numerical advances. Trading houses did not, for instance, suddenly ‘see the light’ and begin adopting in the mid 1990s the Heath-Jarrow-and-Morton/LIBOR-market-model paradigm simply because they realized it was ‘a better model’. I have argued elsewhere (see again Rebonato 2002) that its generalized adoption (at least for path-dependent securities) would not have been possible had important and independent advances not simultaneously been made in the Monte Carlo area. It is also interesting that for a few years compound options were deemed to be too difficult to handle with the same approach, and trading houses were therefore using different models for different products on their interest-rate books. The provocatively-titled article ‘How to throw away a billion dollars’, which has created great controversy both in the academic and in the trading communities, indeed argued that those institutions using single-factor models of the old school were, by so doing, leaving on the table a significant portion of the theoretically available profits (the billion dollars in the title). It is easy to see how an article of this type might change the pricing consensus for a product over a very short period of time.

More generally, there is a spectrum of continuously evolving opinions regarding the criteria on the basis of which the quality of a model should be assessed. The ability to recover the prices of the underlying plain-vanilla options is an obvious desideratum. It has soon been realized, however, that it is not too difficult to achieve this goal in isolation if no control is imposed on other financial features, such as the time-homogeneity of the evolution of the smile surface or of the term structure of volatilities. The consensus as to which of these features are important also evolves over time, and does so at a strongly non-constant pace: non-monotonic interest-rate smiles ‘suddenly’ appeared after the turmoil triggered by the Russia events, and stochastic-volatility or jump-diffusion extensions of the LIBOR market model were just as suddenly deemed to be desirable.

2.9 Theoretical Conclusions

Recall that we defined model risk as the risk that our mark-to-model might not reflect the price an instrument would fetch in the market. If the EMH held exactly true, the price of all products would always coincide with their value. Guessing the level of a future transaction (its future market price) would therefore be tantamount to estimating its fundamental value. In an informationally efficient market, the better the model, the greater the likelihood that this identification of future price and value will be realized. Simply put, model risk would simply be the risk that our model is not good enough.

I have argued however that the way prices of complex and/or illiquid instruments are arrived at in the market can be influenced by a variety of factors that in the EMH analysis either should be irrelevant (ultimately, supply and demand), or are poorly accounted for (stochastically-varying risk aversion), or are generally altogether neglected (say, liquidity). Whether markets ‘ultimately’ do always move towards greater efficiency is a fundamentally important question, but one of a more metaphysical than empirical flavour. In the short run, I believe that the deviations from greater efficiency can be significant, and I have argued that this affects the concept of model risk, and the practice of its management. In my view, the task faced by the model risk manager can therefore be summarized as follows: guessing how future ways of arriving at prices might differ from the accepted wisdom of today, without assuming that the direction of model evolution is fully pre-ordained.

If we accept these conclusions we can derive some indications as to how the task can be carried out in practice. This is tackled in the next section.

3 – Practice

The theoretical treatment presented above can provide some useful guidelines as to how model risk management can be carried out in practice. The first casualty of the way to look at model risk I have advocated is the traditional concept of model validation.

3.1 – Inadequacy of Traditional Model Validation

Model validation is usually meant to be the review of the assumptions and of the implementation of the model used by the front office for pricing deals, and by finance to mark their value. Absence of computational mistakes is clearly a requirement for a valid valuation methodology. The idea, however, that a review of the model assumptions per se can give reassurance as to the correct mark-to-market is much more difficult to justify. Rejecting a model because ‘it allows for negative rates’, because ‘it does not allow for stochastic volatility’ or because ‘it neglects the stochastic nature of discounting' can be totally inappropriate, from the risk manager's perspective, if the market is happy to live, for particular products, with these blemishes. Similarly, requiring that a product should be marked to market using a more sophisticated model (ie, a model which makes more realistic assumptions) can be equally misguided if, for any of the reasons discussed in Section 2, the market has not embraced the 'superior' approach.

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 in order to arrive at the observed traded prices. In order to carry out this task the risk manager will require a variety of sources: in particular, market intelligence and contacts with members of the trader community at other institutions are invaluable. In addition, also very important is the ability to reverse-engineer observed prices using a variety of models in order to 'guess' which model is currently most likely to be used in order to arrive at the observed traded prices. In order to carry out this task the risk manager will need a variety of properly calibrated valuation models, and information about as many traded prices as possible.

The next important task of the risk manager is to surmise how today's accepted pricing methodology might change in the future. Notice that the expression 'pricing methodology' makes reference not just to the model per se, but also to the valuation of the underlying instruments, to its calibration, and possibly, to its numerical implementation. In the light of the discussion in Section 2, the risk manager should not assume that this dynamic process of change should necessarily take place in an evolutionary sense towards better and more realistic models and more liquid and efficient markets. An interesting question for a model risk manager, for instance, could be: 'How would the price of a complex instrument change if a particular hedging instrument (say, a very-long-dated FX option) were no longer available tomorrow?'. The next sections will describe in some detail how these tasks (gathering market intelligence, reverse engineering of prices, and guessing what the market-chosen future model might be) can be carried out in practice.

3.2 Gathering Market Intelligence

In the context of model risk, market intelligence is first and foremost the ability to 'see' the largest possible number of transactions, and the levels where they trade. When a trade has been lost in a competitive-tender situation, the salesforce can often provide information as to where the trade 'went in the market', how many other players submitted a more aggressive bid, etc. No matter how good or convincing a theoretical model might be, few state of affairs should worry a risk manger more than the trader who, using this model, consistently beats all competing banks in a competitive-tender situation. In this respect, contacts with brokers, traders or risk managers at other institutions can provide what is possibly the most effective 'early-alert system'.

Being aware of the latest market developments, and of academic papers can be very useful in guessing which direction the market might evolve tomorrow (the caveat of a non-linear evolution towards ever greater 'perfection' should, of course, always be kept in mind). Professional conferences are useful in this context in order to gauge the market's reception of new ideas, and the likelihood of their becoming the next market standard.

Another possible cause for concern is the sudden occurrence of large-notional trades for which it is difficult to establish a clear rationale on the basis customer-driven demand. The motivation behind these trades could be the very actions of pseudo-arbitrageurs who were discussed in Section 2. These players might have identified particular types of transaction for which common market assumptions or practices cannot be justified. Historically, examples of such trades have been, for instance, LIBOR-in-arrears swaps, forward-starting swaptions, CMS caps, etc. In these cases, the risk manager should not feel automatically reassured by discovering a *bona fide* 'customer' name on the trade tickets, since the trusted client could be in a back-to-back transaction with a market professional.

In summary, it should be stressed that, without at least some anchor points of solid knowledge about the levels and the nature of actual market transactions, the task of the model risk manager is utterly hopeless. 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.

3.3 Reverse Engineering- Plain-Vanilla Instruments

Let us assume that adequate market intelligence has been gathered. The next task is to reverse engineer these prices, ie to try and find the set of pricing methodologies that can bets account for them. The exercise should be carried out starting from the simplest building blocks. For instance, in the interest-rate arena a LIBOR yield curve generator should be tested first, with special attention devoted to slightly off-market deals: a small number of forward-starting or off-coupon swaps, swaps with 'stubs' at the beginning or the end, odd-maturity/reset/frequency transactions, etc, can often provide more useful information than thousands of totally standard trades.[ix]

The next stage in the reverse-engineering process would be the recovery of the prices of plain-vanilla options. Typically, market quotes are readily available for a relatively small range of maturities and for strikes which are not too far from the at-the-money level. The creation of a full smile surface that extends well into the wings is essential in model risk management both for the direct risk of mis-marked out-of-the-money options, and because large portions of the smile surface typically constitute the inputs to models used for more complex products.

Much as it would be desirable, it is not always necessary, or indeed possible, for a 'model' in the traditional sense (ie, for a description of the process of the underlying) to account for the observed market prices of plain-vanilla options. A number of surface-fitting methodologies that do not assume a particular dynamics for the underlying state variables have been proposed for the task, and the criteria for choosing among them are not always clear-cut. Apart from the quality of fit, robustness of the estimation procedure and stability of the optimized parameters for small variations in the inputs are often desirable features. More fundamentally, producing a smooth and well-behaved *inter*polation between observed prices (implied volatlities[x]) is not too arduous a task. Where most of the model risk often resides, however, is for the opaque out-of-the-money strikes and/or long maturities where few reliable quotes are observed (the *extra*polation part of the exercise). It is therefore important to check that the chosen smile-surface-generator behaves well not so much in the centre of the smile surface, but also in the wings.

3.3 Reverse Engineering- Complex Products

At this point we can assume that the model risk manager has a satisfactory understanding of the market dynamics behind the largest trades (trading levels, competition with other players, nature of the demand for the product, etc); that she has confidence in the congruence between the methodology used in-house to price the fundamental building blocks (eg, the yield curves) and the corresponding market practice; and that she can trust the approach used to account for the prices of plain-vanilla options not just for close-to-at-the-money liquid calls and puts, but also for the far out corners of the smile surface. The next task that she must face is the valuation of complex products.

Moving from plain-vanilla to complex instruments, the number of opaque variables that can affect their prices sharply increases. For plain-vanilla options, the description of how the model-independent implied volatility changes with maturity and strike gives the full information about the associated prices[xi]. Therefore, a relatively simple, model-independent parametrization of the smile surface as a function of strike of maturity is not too arduous to achieve. Unfortunately, the price of complex products may depend on a much larger number of variables: for instance, on full yield curves and term structure of volatilities, on correlation matrices, on credit transition matrices, etc. As a consequence, finding a clever model-independent algorithm that can account for observed market prices over a large range of this multi-dimensional space is almost impossible, and an appropriately calibrated model (ie, a specification of the processes for the underlying stochastic variables) must almost always be used.

For practical applications and for illustration purposes the discussion can be profitably split between interest-rate products and equity- of FX-based instruments. For the former, it is often claimed that the LIBOR market model approach has 'won the day' and that a wide industry consensus has crystallized around this pricing methodology (see Rebonato (2002) for a history of the developments that led to the LIBOR market model). While this can be true, it is often forgotten that the LIBOR-market-model approach ultimately describes conditions of no-arbitrage among forward rates, and that it only becomes a 'model' once specific choices are made for the instantaneous volatility and correlation functions. Significant pricing differences can therefore arise between traders who use the same pricing approach, but who choose different functional forms for the input functions. The likelihood of encountering in the market different choices for volatility and correlations is increased by the fact that there exists an infinity of volatility functions that can perfectly reproduce the observed market prices of caplets or European swaptions. Furthermore, no consensus exists as to which market prices the model should be calibrated to (caplets or swaptions, or to both simultaneously). Finally, the number of factors that should be used for pricing certain products (notably, Bermudan swaptions) is still open for debate (the more-than-academic importance of this issue is demonstrated by the very title of the controversial Longstaff, Santa-Clara and Schwatrz (2000) paper that dealt with the pricing impact of the number of factors: 'How to Throw Away a Billion Dollars'). Therefore, even if a general consensus has emerged regarding the general pricing approach, substantial differences still exist, even in the standard version of the model[xii], as to its actual implementation (and hence to the prices it can produce). In my opinion, certain strategies for choosing the volatility functions are financially much more satisfactory than others, and should be employed by a trader. From the risk management perspective, however, the possibility should be entertained that the market might disagree with the 'optimal' choice, and that prices substantially at variance might be produced by different calibrations (to caplets only, to swaptions only, to both markets) and/or by different functional choices ('flat' volatilities, time-homogeneous volatilities, forward-rate and time-dependent volatilities, etc).

The state of affairs is significantly different for complex equity- or FX-based products. A similar broad consensus about the most desirable pricing approach has not been reached yet, and several major pricing philosophies uneasily coexist: jump-diffusions, stochastic-volatility, local-volatility and gamma variance, to name just a few. Matters are not simplified by the fact that combinations of these approaches (eg, of local-volatility with either jump-diffusion or with stochastic volatility, or of jump-diffusions with stochastic volatilities) have been proposed. The number of 'fitting parameters' is sometimes staggering (more than twenty for the combined stochastic-volatility/jump-diffusion approach), and the calibration methodologies, often based on chi-sqaure minimization techniques, tend produce solutions of similar numerical quality, but with completely different parameters.

The shift in market consensus as to the best 'model of the month' has been more rapid and fitful in equities and FX than in the interest-rate arena, perhaps indicating that each competing approach is perceived as a partial solution, whose blemishes and advantages must be weighed and reassessed on a continuous basis. The local-volatility approach, for instance, has been amply criticized for its poor hedging performance and for implausible evolution of the smile surface that it implies. Nonetheless, many traders and risk managers like to use it for the 'comfort factor' provided by its ability to recover by construction the prices of the underlying plain-vanilla options.

In the light of the discussion above, my recommendation for the prudent model risk manager is not difficult to guess: several models, with fits of similar (good) quality to the plain-vanilla market prices should be used to identify a range of possible future market prices. In the same spirit, different calibration methodologies, that produce similar prices for calls and puts but possibly very different prices for complex products, should also be used in parallel. And, finally, the risk manager should not unquestioningly assume that an intellectually more satisfactory model or calibration strategy will necessarily be chosen by the market.

References

Andersen L and Andreasen J (1997) '*Volatility Skews and Extension of the LIBOR Market Model*', Gen Re Financial Products working paper, to appear in *Risk*

Dunbar N, '*Inventing Money*', (1999) John Wiley, Chichester

Glasserman P, Kou, S.G (2000) '*The Term Structure of Simple Forward Rates with Jump Risk*' - Working paper - Columbia University

Green TC, Figlewski S (1999) '*Market Risk and Model Risk for a Financial Institution Writing Options*'

Hull J, Suo W, (2001) '*A Methodology for Assessing Model Risk and its Application to the Implied Volatility Function Model*', working paper, Joseph L. Rotman School of Management, University of Toronto

Longstaff FA, Santa-Clara, Schwartz E, (2002) 'Throwing Away a Billion Dollars: The Case of Sub-Optimal Exercise Strategies in the Swaptions Market', forthcoming in *Journal of Financial Economics*

Rebonato R (1999) '*Volatility and Correlation*', John Wiley, Chichester

Rebonato R (2001a) '*The Stochastic Volatility LIBOR Market Model*', *Risk*, October 2001, 105-110

Rebonato R (2001b) '*Managing Model Risk*' in *Handbook of Risk Management*, Carol Alexander editor, FT-Prentice Hall

Rebonato R (2002) '*Modern Pricing of Interest-Rate Derivatives: the LIBOR Market Model and Beyond*', Princeton University Press, Princeton, New Jersey

Rebonato R, Joshi M (2002) '*A Joint Empirical/Theoretical Investigation of the Modes of Deformation of Swaption Matrices: Implications for the Stochastic-Volatility LIBOR Market Model*', accepted for publication in *International Journal of Theoretical and Applied Finance*, and Working Paper, QUARC (Quantitative Research Centre), available at www.rebonato.com

Sharpe W, Alexander G, (1998) '*Investments*', 6th Edition, Prentice Hall, Englewood, New Jersey

Shleifer A (2000) '*Inefficient Markets: an Introduction to Behavioural Finance*', Oxford University Press, Oxford

[i] The accountancy regime can make a substantial difference to the timing of the recorded profits. Assume that, perhaps because of poor information, an agent sells short-dated equity options of a wide range of strikes and maturities a bit too cheaply (ie a bit below where professional traders would manage to sell the same options). On a mark-to-market basis the trades will immediately show a loss. Options are priced in the market, however, as discounted expectations in the risk-neutral measure, not in the objective (econometric) measure. Therefore it is possible that the liabilities generated by the same options, left unhedged, might be more than covered, on an actuarial basis, by the premia received. The strategy of continuously selling options at a loss on a mark-to-market basis, but at actuarially advantageous prices might therefore be profitable on an accrual-accounting basis.

[ii] Unfortunately, the neat distinction between the 'haemorrhaging' of the trader's profitability that can be controlled by exercising the stop-loss discipline and the sudden realization of a loss due to mis-marking of position can sometimes become rather blurred: this can occur when the frequent, but relatively small, losses due to wrong model choice induce the unscrupulous trader to influence the marking of his book in a flattering fashion. The trader's ability, and temptation, to do so will be greater, the more opaque the products in which he trades.

[iii] 'Pseudo-arbitrage' is defined in this context as the 'simultaneous purchase and sale of the same, or essentially similar, security [...] for advantageously different prices' (Sharpe and Alexander 1998)

[iv] Criticisms of the EMH simply based on the observation that not all investors are rational are therefore irrelevant: the EMH can still hold even if there are irrational investors, and if their actions are co-ordinated. As long as unfettered pseudo-arbitrageurs exist we can simply *pretend* that all agents are rational, and we would get exactly the same results. The situation can be compared with game theory, which has also been criticized on the basis or requiring hyper-rational players. Yet, as long as suitable reward and punishment mechanisms are in place, we can successfully analyze evolutionary problems (Maynard Smith) *as if* we were dealing with purely rational agents (even if, in the evolutionary context, the 'players' might actually be amoebas). Showing that amoebas are not fully rational is neither very difficult, nor very productive in a discussion of the usefulness of game theory to tackle evolutionary problems.

[v] Indeed, in order to limit the risk of investors suddenly pulling their money out, many hedge funds impose restrictions on their investors' ability to withdraw funds at will.

[vi] As far as can be ascertained from public information (see, eg, Dunbar 2000) this is indeed part of what happened to the hedge fund LTCM in the autumn of 1998.

[vii] When, in the late 1990s, I used to trade complex interest-rate derivatives the farm of computers needed to price and hedge the trades on our books made my group rank immediately after Los Alamos National Laboratories in terms of computing power.

[viii] not *too* swiftly, though, if the pseudo-arbitrageurs want to be able to pick up their five-dollar bills. The EMH needs the pseudo-arbitrageurs, but, after all, their efforts can only be worth the trouble if markets adjust to fundamentals with a finite speed.

[ix] The idea behind LIBOR-in-arrears swaps, for instance, that caught quite a few naïve banks off guard in 1994, is based on a simple variation on the standard-swap theme (the pay-off time in the arrears swap coincide with the reset time, instead of occurring an accrual period later).

[x] Quoting an implied volatility does not imply an endorsement of the Black (or of any other) model. It is simply the 'wrong number to plug into the wrong formula to get the right price' (Rebonato (1999)).

[xi] In the case of swaptions the situation is somewhat more complex, because the implied volatility is a function of maturity, expiry of the option and length of the underlying.

[xii] The 'standard' version of the LIBOR market model does not allow for smiles, which have become more and more pronounced, and complex in shape, after the market turmoil of summer/autumn 1998. Several modification that can be 'grafted' on top of the LIBOR-market-model approach have been proposed (see, eg, Rebonato (2001), Rebonato and Joshi (2002), Andersen and Andreasen (2002), Glasserman and Kou (1998)).