RECORD, Volume 25, No. 1*

Atlanta Spring Meeting
May 24–25, 1999

Session 76CS
Model Risk

Track: Investment
Key Words: Investments

Moderator: RISHI KAPUR
Panelist: ELLIOT NOMA†
Recorder: RISHI KAPUR

Summary: Actuaries forecast future events and the implications of those events. We know that the results of our models are only as good as the assumptions and inputs. In financial modeling, one of the key "inputs" is the stochastic economic generator. Do the commonly used generators develop an appropriate distribution of possible outcomes? With value-at-risk analysis, in particular, the focus is on the left tail of the distribution. Do our economic generators produce good data for the tails or are they better used for generating expected values? The presenters demonstrate the potential risks of relying blindly on our models without sensitivity testing.

Mr. Rishi Kapur: Elliot Noma is from Deutche Bank in New York. He is not an actuary and has never been at an SOA meeting. This is his first talk at a SOA meeting as well, and he will be representing things more from the investment banking point of view than the insurance point of view that we're mostly familiar with.

I am going to start first and talk a little bit more about the overview of models. I'll discuss some of the major model risks and some of the problems that I have faced when modeling. Then Elliot is going to talk more specifically about a topic that is very important for a lot of us, which is interest rate models, especially over long periods of time.

---

†Mr. Noma, not a member of the sponsoring organizations, is Director, Risk Management Group at Deutsche Bank, New York, in New York, NY.

Note: The charts referred to in the text can be found at the end of the manuscript.

I am currently with a fairly new company. It is about a year old and it is called RGA Financial Products. We are based out of Toronto. Prior to joining RGA Financial Products, I was working at the Canadian Imperial Bank of Commerce (CIBC) in Toronto. At CIBC, as part of the Global Analytics group, we were responsible for validating all the models developed in the bank. I have looked at a lot of models across all asset classes (equities, interest rates, and commodities). When I joined RGA, one of the first jobs I had to do was validate their insurance models on pricing segregated funds and guaranteed minimum death benefits (GMDBs). I have seen a lot of models out there —the good, the bad, and the ugly—and I am going to just share a few thoughts about those models.

The title of my presentation is, "Model Risk: The Dirty Secrets of Every Model." Models do have a lot of dirty secrets. We use a lot of models in our everyday life. I just want to step back a little and talk about some of the problems that are faced by models.

I have basically divided the presentation in three parts. I am going to give a brief introduction about model risk and talk a lot about what can go wrong in models. I am not really going to focus on the positive side of models. I just want to focus on the negative side because I want to focus more on what can go wrong with models that we use.

The first question that we ask ourselves is, why do we need a model? There are three reasons: valuing a product, assessing the risk exposure and sensitivity to various factors we have in our model, and developing an appropriate hedging strategy or risk management strategy.

One thing that I want to talk a little bit about is what is driving all these models—product innovation, increasing model complexity, and greater computing power. As we have more and more innovative products being sold in the marketplace, a combination of insurance products and investment products, it generates a need for us to have more complex models to value these products and manage the risks associated with these products. To handle these increasingly complex products, we need greater computing power for simulation purposes and scenario testing. Therefore, product innovation is driving increasing model complexity, which is using greater computing power. But it also works the other way. Because of the greater computing power that is available to us now as opposed to five or ten years ago, our models are becoming more complex. Some of the things that we never would have attempted to do ten years ago we can easily do right now because we have very fast computers. So in some respects greater computing power is driving us toward more complex models. I'm going to talk a little bit about the risks associated with that.

To some extent, these increasingly sophisticated models are driving product innovation. Some of the products that we would normally not sell are being sold today because we have some idea on how to model them. We might not have sold them five or ten years ago because we did not really understand them or know how to handle them. Of course, there are a lot of products sold in the market that we do not know how to model.

I divided these models into two groups. There are traditional actuarial models and there are capital market models. I'm going to talk a little bit about what distinguishes these models from each other, at least as I see it. The actuarial models that a lot of us use are based on a historical analysis of data regarding mortality, morbidity, policyholder behavior, lapses, stock returns, interest-rate changes, etc. These are things that we are familiar with, and we have been using these models for a while. The prices that we charge are based on expected claim costs, plus some sort of padding or risk charge that we apply for profit and provisions for adverse deviation.

Capital markets models are generally based on current market information about stock prices, interest rates, and volatilities that we observe in the financial markets all the time. The prices in capital markets are not based on expected claim costs but on expected hedging cost (cost of production). It does not matter what the payoff from a particular product is. The price is driven by the cost of replication.

The market price of risk is implicitly included in the pricing methodology in the capital market. However, on the actuarial side, we actually have specific risk charges or provision for adverse deviation. Actuaries try to use an explicit approach in pricing while in the capital markets pricing methodology the market price of risk is implicit.

The most famous capital markets model is, of course, the Black-Scholes model, which has been around for 25 years. The authors have won a Nobel Prize for the formula. It provides a simple and robust technique for option pricing. The newer, more complex models incorporate multiple factors like stochastic volatility, but these models can be, in general, difficult to use. They are not as easy to use as the Black-Scholes model.

Interest-rate risk management has moved from simple duration and convexity analysis to multifactor stochastic interest rate models. A few companies have started using those. A lot of them are still using the duration and convexity approach for asset/liability matching. However, the wave of the future is multifactor stochastic models, not a simple duration/convexity approach.

Credit risk management, which insurers are quite familiar with in terms of having huge bond portfolios, has moved from simple limit management to more complex stochastic models incorporating default probabilities and recovery rates. We have had a few presentations on that topic at this meeting. This is just sort of a new wave for credit risk management. We are moving away from the simple limit management that has traditionally been the case.

There are two broad classifications of models out there. One type is the equilibrium no-arbitrage model that is quite common in the financial markets. Examples of such models would be the Black-Scholes model for options and the Heath-Jarrow-Morton model for interest rates.

The second class of models is the statistical econometric models. Examples of such models are generalized auto regressive conditional heteroskedasticity models for asset price volatility. Econometric models for asset prices, interest rates, health care, and pension costs trends are based on economic variables. This second classification is the one that actuaries are familiar with in terms of trying to project future scenarios of interest rates, inflation rates, and equity markets.

We can combine both classes of models. For example, the price of an option may be generated using the Black-Scholes model, but the volatility in some of the other inputs may actually come from the econometric or statistical model.

What can go wrong when using models? We are going to go through a few of those scenarios. First, there are errors in the pricing formula. We can come up with a pricing formula, but there could be an error in the formula. This is especially important when pricing new, complex past-dependent products that depend on several risk factors. Complex programming may have bugs that are difficult to detect. Numerical approximations may not be appropriate. As actuaries, we use numerical approximations all the time for calculating integrals and all sorts of other things. Another problem is the slow convergence of results, especially if the pricing formula is based on simulation. How many scenarios should you use: 300, 3,000, or 10,000?

Benchmarks to compare results may not be publicly available, especially for new products. You come up with a price, you really do not know whether it is the right price, and you do not really have benchmarks to compare your numbers to. You may be totally off and have no idea.

These errors are often difficult to discover. You may have errors for a long period of time. You may not be able to detect bugs for a long period of time until something really bizarre happens five years down the road. Then you'll have to go back and realize that there was a bug in the code.

Incorrect assumptions for the behavior of risk factors can create errors. All our models depend on risk factors. We presume that they behave in a certain manner. We can always make an incorrect assumption.

The capital markets generally use a constant volatility assumption for stock prices, but that is inconsistent with volatility smiles and skews observed in the market. Basically the market may assume constant volatility, but it tries to offset that assumption by charging different volatilities for different options.

Arbitrary adjustments made to match volatility skews and smiles may have unintended effects on risk management. Capital markets traders are very good at tweaking their inputs to affect market prices or to get some sort of an estimate for price. That tweaking makes quantitative analysts quite uncomfortable because the process is ad hoc. Although traders are very comfortable with these adjustments, we do not really know how this would affect risk management. That is one of the big problems that we have.

One of the things that can be done is that theoretical stochastic behavior can be replaced by empirical distributions to take into account things like fat tails for stock prices. Stock prices do not follow the lognormal distribution. The chances of large downward movements in stock prices are very big compared to the standard assumptions we use. We can try to use an empirical distribution to take into account the fat tails. Some people try to use distributions such as chi-squared and inverse gamma.

Now these so-called theoretical distributions are great for risk management—things like value-at-risk or credit risk management—but they are not so good for options pricing. Certain approaches that you might use to solve problems in risk management fail when you are pricing products and vice versa. I will talk a little bit about that later on.

Another thing that can go wrong is that we all rely on risk factors. We may be missing some risk factors.

Another risk is that multifactor stochastic models may be required for pricing, but we use single-factor models. Material risk factors may be judged to be immaterial partly to keep the model simple. We do not really understand how a particular risk factor should be modeled; hence we ignore it. It is partially motivated by the fact that we do not really know how to handle those risk factors. And these so-called nonmaterial factors can suddenly become material because of changing economic scenarios. Things that were thought to be totally irrelevant suddenly become relevant. An example of that is guaranteed interest rates for policy loans, which led to losses in the early 1980s when interest rates rose.

Minimum credited interest rates on universal life products. Again, guarantees that were felt to be worthless at the time they were given led to huge losses.

Another thing that can go wrong is nonperfect capital markets. We all know that capital markets, like insurance markets, are not perfect. Things can go wrong in these markets. Option pricing models rely on the existence and execution of a hedging strategy in perfect capital markets.

Basis risk and hedging strategy can be high if the underlying assets cannot be used for hedging purposes. This is especially important for things like GMDB and segregated funds in Canada because the underlying assets are mutual funds. The underlying assets are not the Standard and Poor's (S&P) 500 but are mutual funds like the Templeton Balanced Fund. These funds are not traded in the capital markets. You may sell guarantees on the Templeton International Technology fund that invests in high-tech companies around the world. There is no index that you can buy and sell to manage the risk associated with that. The fund does not necessarily behave like the S&P 500.

Transaction costs for hedging strategies can be very high, especially in less developed markets. In the U.S. you may be able to execute hedging strategies. In countries like Malaysia and Indonesia, the liquidity just isn't there. For you to be buying and selling all the time, you may pay a huge bid offer spread, causing your transaction costs to be very high. You may not have taken that into account when pricing your products.

Execution may fail if liquidity of underlying assets fall. Distressed sales can increase losses. The most famous example of that is October 1987. You could not buy and sell. The brokers and the investment banks were not buying or selling. Now your whole assumption in terms of risk management relies on the fact that when the markets move, you can buy and sell. What if you cannot do that?

Another thing that can go wrong is model calibration errors. We always have these wonderful models. The only problem is we have to start estimating parameters when using them. What are the errors that can happen with that? One is inappropriate statistical techniques to estimate parameters; another is an inappropriate time horizon used, which is especially key for volatility estimation. How far back do you go? Three months, six months, or one year? You can get radically different results. What is the right time period? We do not know. There is estimation error in the statistical technique. Even though you're using the right technique, every statistical technique has errors associated with it.

Estimators are sensitive to the treatment of outliers. How do you treat outliers? Ignore them? Are they really outliers? Depending on how you treat them, you get different results. I will talk a little bit about outliers later on in the context of the Asian economic crisis in 1997.

Low frequency of parameter estimation can lead to significant losses. We may estimate parameters today, but how often should we update them? Do we ever think about updating them? Do we only update them once a year? This may not be as much of a problem on the insurance side, but it matters in the capital markets where markets change every second. If you estimate your parameters today and then forget about estimating them for the next month, your parameters may be totally unrealistic and that can lead to significant losses. It has led to significant losses among some of the banks in the U.K., like NatWest. It has calibrated its models based on certain information. It forgot to calibrate its model for a few weeks and found that they were underpricing some of their products as a result. Traders know a good bargain when they can find one. Suddenly, the sales of these products at NatWest rose because they were underpricing the product. Everyone knew NatWest was underpricing when they bought from them. Once the error was discovered, the traders were let go.

High bid/ask spreads and/or infrequent trading of financial securities do not provide a good basis for parameter estimation. If you have prices in the market but the bid price is one dollar and the ask price is two dollars, what price do you use for calibration? Do you use the midpoint price? In some of the more illiquid markets, the bid/ask spread has nothing to do with what the price is traded at because the bid/ask spread is not firm. You might quote a price of one dollar, but you are not obliged to charge the buyer one dollar. You could charge them three dollars. In a lot of cases you cannot find prices in the financial markets. They're not publicly available. Which one do you use? Certain bonds, like corporate bonds or some of the junk bonds, might not trade for weeks. What price do you use in the meantime if you're calibrating? Should you use the price at which it was last traded, which may be a month or two months ago? It might be that no one is willing to actually quote you a price on that, especially if you're not their client.

One of the key things that actuaries are getting more and more concerned about is the so-called volatility in the market. There are two basic approaches used for volatility calculations. People try to say, "I'm going to estimate volatility historically." That is one approach that actuaries are quite familiar with. The other approach is saying, "I'm going to use the volatility that the market is assuming—it is the so-called implied volatility on the option prices." I might think volatility is 20%, but the market, in general, feels it is about 25%. Which one should I use? Should I substitute my judgment for the market? Chart 1 looks at historical estimation of volatility versus the so-called implied volatility that you observe in the market. The idea is to try to see which is the better estimate. Is historical volatility a better estimate of future volatility or is this so-called implied volatility that you observe in the market a better estimate? I have plotted the difference between the two. The dark line is the actual volatility minus the implied volatility. It gives you the error between the forecasted volatility from the market versus the actual volatility that was observed in the market. You can see that sometimes the implied volatility is higher; however, sometimes it is lower than the actual volatility. With implied volatility, you may end up overestimating or underestimating. It depends on the time.

The lighter line is the actual volatility minus the historical volatility. Again, in that case, sometimes you underestimate and sometimes you overestimate. Now which of these is a better estimate? The line closer to zero on the x-axis is a better estimate. You might argue that during some time periods historical volatility is generally better. During other time periods, you might argue that implied volatility is generally better. Which is the right answer? Nobody knows. You don't know what the future is going to be like. As actuaries, we have these wonderful models, and we want to estimate volatility. Our first step is historical or implied. Which is better? There's no clear answer.

This reflects a three-month period from July 1994 to about February 1999. The implied volatility that I have used actually comes from an index traded at the Chicago Board of Exchange. The volatility index is quoted just like the S&P 500 or the Dow Jones, and it is a measure of volatility in the marketplace.

Chart 2 shows the overnight interest rate in Hong Kong between January 1997 and the end of September 1998. Even though the rate is pretty stable, it's a little bit volatile here. This is October 23, 1997. The overnight interest rate in Hong Kong jumped from 9% to 129% in one day. The next day it fell to 16%. Now imagine if you're at October 24 or 25. Is this an outlier? Suddenly there was a huge shock in the market. Is it going to repeat itself? The market still remains very volatile. You can see interest rates jumping all over the place. August 1998 is again very volatile because of Russia defaulting on its debt and questions about long-term capital management.

At the end of September 1998, you might feel that the 129% and the number from August 1998 were probably exceptions to the rule. You would not have that information, so if you're trying to calculate your model, what do you do?

Say you are at the end of October 1998 or September 1998 and you are trying to measure volatility historically. Let us say you want to use one-year data. If you use all the data, historical volatility is 466%, so it's absolutely huge. The current volatility for the S&P 500 is about 25–30%. Even the Internet stocks do not have that much volatility, and that is because of that one outlier.

If I just remove that one outlier, the volatility drops to 277%—which is still a great deal of volatility. Your numbers are radically different depending on whether you include that one point or not. It is very sensitive to the treatment of outliers, right? I ran another test saying, how many more points would I have to remove to get this volatility down? It might remove one or two points. If I remove the next 5 most volatile days in 1 year—there are 250 trading days in a year—I would be saying that about 2% of the data are outliers, which is huge. The volatility drops to around 230%, which is not a very big drop. If 2% of your data is outliers, that's a lot of data. Perhaps it's not really an outlier.

Handling situations like this is very difficult. I was at CIBC around this time. The reason I took this example is because I've worked with this. When we started estimating volatilities after these big jumps, the volatility numbers rose a lot. The maximum loss on the portfolio could be $1 million, but according to our models and our assumptions, we could lose more than the value of the portfolio. That was simply because of a function of our models. Our models were not designed to handle stuff like this.

What Else Can Go Wrong? Input Data Errors
As actuaries we use a lot of data, and we cannot always be sure of the integrity of that data. What are the input data errors? You have incorrect data or the interpretation of policyholder information. Incorrect interpretation can lead to other negative consequences. On the financial market side, incorrect data feed from internal or external sources of financial market data. The data that you rely on from data vendors that provide data to you on a daily basis might have errors in them.

Incorrect Use of Financial Data
Should you use historical volatility or implied volatility? In some cases historical should be used, and in other cases implied should be used. You can always switch the two if you are not totally familiar with the subtleties of what the differences are.

Interest Rate Curves
Should you use the zero curve from the Treasury? Should you use the yield curve or the swap curve? I can probably present arguments to use all three under various situations.

Model Application Errors

The model may be unsuitable for current market conditions, or the model may be unsuitable for the current product. Let us look at both of these in a little more detail.

The model may be unsuitable for current market conditions. The lognormal assumptions, which are fairly commonly used in the capital market, might not be appropriate in periods of high volatility; for example, the Asian crisis of 1997. The chart that I showed you about interest rates moving up to 129% should never happen under the lognormal assumption. Your assumptions generally fail in that instance. The example in Canada shows that at the time of the referendum in 1995, the financial markets became very volatile when it seemed as if Quebec was going to vote to separate and all the models were useless. How do you manage risk when the markets are moving up and down 5% and there is so much fear in the marketplace? The lognormal assumptions are great most of the time, but they can fail.

The second example is that loans to Microsoft have much higher capital requirements versus loans to South Korea. South Korea is an Organization for Economic Cooperation and Development (OECD) country. According to the Bank of International Settlements, if banks lend money to an OECD country, they have to put aside zero reserves against that because they are considered default-free. OECD countries include the U.S., Canada, the U.K., and South Korea. But if banks lend money to a corporation like Microsoft or G.E., they had to put aside 8% of the loan as some sort of a risk-based capital charge.

That might have been good a long time ago. The Asian crisis hit in 1997. Many of the Asian countries' currencies are devaluing; the South Korean stock market is down 80%, right? Companies are defaulting, and banks might default. How much in reserves do you have to put aside if you're lending money to South Korea? Zero. You have to put aside 8% of the money if you're lending to Microsoft. If you're a bank at the time of the Asian crisis, who would you rather lend money to—Microsoft or South Korea? From a risk management point of view, Microsoft, but from a balance sheet point of view, South Korea. Your model that you used to allocate capital is no longer suitable.

The other example is the model might have been developed for other products and may not be suitable for this product. An example from capital markets is that a one-factor interest rate model may be appropriate for a callable bond; insurance companies have a lot of callable bonds on their books. But it's not suitable for a callable convertible bond where the return is not just based on interest rates, but is also based on the equity market. So a model that might work for callable bonds will not work for callable convertible bonds. You can have great problems if you start using a one-factor model to price callable convertible bonds. It just doesn't work.

Let's discuss the Monte Carlo simulation. It is something that actuaries are quite familiar with and use all the time for generating scenarios and managing risk. What are the problems associated with the Monte Carlo simulation? I alluded to the first one a little bit earlier. How many runs should you have? How many scenarios should you generate? You may have developed benchmarks for the number of simulations based on one product, but it may not be appropriate for all products. One size does not fit all. A thousand scenarios might be great for some products, but for more complicated products even 10,000 may not be enough. It is very important to keep that in mind.

Effectiveness of variance reduction techniques varies from product to product and also within each product. Variance reduction techniques are used to try to improve the convergence of Monte Carlo and somehow reduce the number of paths that is required for convergence. There are several techniques out there. Which one should you use? It depends on the product, but it's a little bit more complicated than that. Even within the product, it really depends on the input. For example, for option pricing the technique you use may depend on whether the option is in the money, at the money, or out of the money. The method that you use not only depends on the kind of option you're pricing, but the "in-the-moneyness" of the option. Again, that is not something that is very widely known. There are not too many papers that will actually tell you the best technique for each product and for the kind of input. Which one should we use and when? When you get into the implementation and the nitty-gritty, that's where the questions arise and the errors can be made. You may use a variance reduction technique that works for one product, but it can be shown that there are other products in which they actually were soft by using variance reduction. You might think you're helping yourself, but you're actually hurting yourself and you don't even know it.

A quasi-Monte Carlo simulation, which is something new in terms of trying to improve the speed of Monte Carlo, comes from number theory. It uses some sort of a structure in coming up with the scenarios. They work for low dimension but may give unpredictable results for large dimensions, depending on the dimension of the problem. To give you an idea of dimensions of a problem, a lot of insurance companies have mortgage-backed securities on their books. You may have a 30-year mortgage-backed security for which payments are made once a month. That's 360 months. That's a 360-dimensional problem that has to be solved, right? You may know good things about quasi-Monte Carlo, but how good is it for those high dimensions? The answer is that it is not very good. It's very important to keep that in mind if any of the new methods come along. It may have been shown to work very well in low dimensions, but blindly applying it to high dimensional problems like mortgage-backed securities can give you unpredictable results.

Chart 3 is an example of a quasi-Monte Carlo sequence. We used 100 dimensions, which is not a very high number for the needs of the insurance companies, and we plotted the first dimension versus the second dimension. We can see that there is a fairly uniform spread of points. We like to see that in Monte Carlo. We like to see randomness, and we like to see a good spread. We want to cover all the bases. These are just 1,000 points. It is better than the results you get from standard Monte Carlo. Quasi-Monte Carlo is used because it is better in low dimensions.

Let's look at the higher dimensions. Dimension 99 versus dimension 100 is not quite random. There's a very well-defined structure to these points. Would we use these? Probably not. Would we know to look for errors like this? Now if you see a 100-dimensional vector, and you see 10,000 of those, you cannot really see the pattern between the 99th and the 100th number. It could be a problem. So what I did was I generated these 100 dimensional numbers and looked at the relationship between the 99th and the 100th dimension (Chart 4). It's not at all random. This is actually a little bit better than some of the other sequences. You might just have one line that sort of extends all the way from a point, and then all the numbers will lie along this line. You can get that. Quasi-Monte Carlo is great for low dimensions, but it is not very good for high dimensions. Again, this is a good model that's applied incorrectly.

I think I focused a lot on what can go wrong with models, but models have their good side too, even though I'm not focusing on them, and I think we all need them. It's just important to keep in mind these are just models.

Just a quick summary. It's important to have an inventory of models, not just one model fits all. Make sure to have lots of different models. Use the ones you feel are most appropriate. It is very important to stress-test every model to determine strengths and weaknesses. There are things like the quasi-Monte Carlo. It's very important to test the 99th and 100th dimension. If you don't test it, you may never figure it out. Do not replace common sense with mathematical complexity. There's a tendency among mathematicians to do that a lot because they're comfortable with the model. A variance/covariance matrix does not fully explain human or financial market behavior. You cannot put everything down to a variance/covariance matrix and look at all risk factors and say, "One's related to this, this way." A lot of times there is no real structure; it's totally random.

This is my view on things. Not everyone may believe in it. Given a choice between an experienced trader with no model and an inexperienced trader with a model, I'll choose the former any day. That's why I always say the traders are paid more than the mathematicians that sit on the trading desks because the mathematicians are good on the models. They're not as good on common sense. There have been so many cases, especially at the bank I used to work at, where you get mathematicians making these trading decisions and losing a lot of money because the markets did not behave the way they thought the models implied. It is very important to keep an idea of common sense and not be overwhelmed by the math.

The other thing that has been my experience in all the models that I have seen is simple, robust models work better than complex, inflexible ones. The standard Black-Scholes model will generally work a lot better than these various multifactor, multidimensional models with large numbers of parameters that need to be estimated. They're inflexible. You lose the intuition, and you can lose common sense. I would rather choose simple, robust models with good inputs than complex and flexible ones where the inputs are difficult to estimate.

Mr. Elliot Noma: When I was first asked to speak, I had never been to an SOA meeting, and I really did not know what the level of complexity was and what assumptions to make. Therefore, I want to be as simple as possible, in terms of my thoughts about models. It turns out that the way we think about this is a matter of the way our brains process. Our brains are building a model as far as what is actually going on out there, so we remove the lines and see that they're actually parallel. Our brains have developed a model, and the model is somewhat incorrect. I think the first requirement of any model I look at is that it has to satisfy some intuition about the markets. The model can be perfect mathematically, but if it doesn't match the way I think the bond might be traded you have a major problem, so I think that's probably our first test. The second test is historically alternative models are also important. I will concentrate on the cost of adopting new models and the risks that are involved. I'll look at one particular example in a fair amount of detail. I know there's at least one of my colleagues in the audience who will say these models are really obsolete. Ten years ago they gave you great models. Now that they're obsolete, they're not here to show you the state-of-the-art. The basic problem is to look at specific bond prices and bond prices in terms of the overall framework of the bond prices. These will be used to price swaps and caps.

In order to value a book, you need a single framework model. You know certain things about the model. It can't be the same as a Black-Scholes model. There is a variety of different reasons for that. One of the things about bonds is that they have to converge to par. There was a series of early attempts made to patch up Black-Scholes initially so that the bond moves with random walks and will eventually return to par. Nevertheless, it tells that there's an intrinsic difference between bonds and stocks. We know we have to develop a model, so this is our motivation.

Another major difference is the volatility of bonds versus stocks, but we have a limited historical series of what the bond prices were over time. In some sense, there's a series out there that's being looked at. If I wanted to look at a two-year bond, I have a maximum of two years worth of history on that bond. I can't go any further because the bond matured. The only reason we can do this is to patch together a series of consecutive bonds and make a model as far as how the bond would move, assuming the bond would continue to roll over. One of the ways to do it is we convert our prices into the yield curve, so we're doing the modeling on the yield curve. We have a couple of different ways of modeling.

Let me quickly review the yield curve. I know you're all very familiar with the picture in Chart 5. This is the yield curve as of April 20, 1999. You probably already know it seems to be a little bit on the low side because of things that happened recently. This shows the different characteristics of upward sloping—3 month T-bills on up to 30-year bonds. Then we look at it in terms of the historical context. They find the difference, whether they are investing in ten-year bonds or in six-month T-bills. We really should think in terms of more than just parallel shifts to the yield curve. Obviously we have to start somewhere, so we start with parallel shifts. There is deficiency, and we want to see how far we can go with this and what the limitations are. The next jump after that will be what happens when you go to the more complex models such as multifactor models. It allows you to reshape your yield for certain prices. I'll talk about what prices bring more realism to your model.

In contrast, Chart 6 depicts a very flat yield curve. It is not only a flat yield curve, but also the rates are in a different range than the rates on the left side.

The basic idea of the one-factor model is that you have short rates that dominate. In other words, the three-month T-bill rate moves up and down and the entire yield curve shifts. One of the reasons why we stick to this is because it is easy to model. It provides closed-form solutions. It eliminates a lot of the problems that Rishi talked about at the very end of his presentation, such as how do we decide what variance reduction techniques to use? Do we do the quasi-Monte Carlo or the standard Monte Carlo? In a closed-form solution, at least within the assumptions of the closed form, we know we have a correct solution. But we do start off with the question about how to model the short rate.

Many of the answers depend on what you're doing. I've arbitrarily taken the three-month T-bill. That's one of the standard assumptions. The model can be modified to take the London Interbank Offered Rates from the swap curve.

The three-month T-bill over a time period has a fair amount of volatility. There are probably certain things that we know about the very, very short rates. One of the things that we know about is that we should be very cautious in terms of what we mean about a random process of the very short rate. We already have a difference in terms of the way we want to think about it, in particular the way we would think about it in a Black-Scholes world. I think as far as starting up with a 10% short rate, which is obviously very high, one of the things to look at is the extrapolations to the lognormal world with a 70% volatility. What would the short rates be? These are the various density curves. There is a trivial probability that within 12 months, the 1 year rate could go from 10% down to 6% or up to 18%. Rest assured. I don't think the Federal Reserve Bank would smile upon this as what they would allow in terms of interest-rate changes. Over a ten-year period it becomes even more strained.

Looking at the short rates, we have a certain amount to pay. First of all, as I mentioned, the acts of the Federal Reserve strongly affect the three-month T-bill rate. We have to consider that. I think the Federal Reserve has a history in terms of moving rates. Sometimes it can be very aggressive with 1% moves. It is very realistic. The traders know what the Federal Reserve is doing. It is looking at the unemployment figures. We're looking at what the trade deficit is. We move it up and down. We try to micromanage a lot of ways, so we have that concern. We have a series of other considerations in terms of yields. I don't expect yields to rise too rapidly, even for Hong Kong. The rates may be 129% in Hong Kong, but that's the overnight rate. The economy cannot be run at that level of rates. At some point, Hong Kong would be paying out more than its principal every night, and I don't consider that a reasonable possibility.

Yields probably don't go below zero with the possible exception of the Japanese market a few years ago. The yields are mean-reverting. High yields probably indicate that. That's from an economic sense, and it will be followed by lowering yields. The same situation exists with very, very low rates. One of the things that has come out of this is that short rates display higher volatility than long rates.
I want to talk about two stochastic models that were developed. They're equilibrium no-arbitrage models. There were two early attempts in 1990.

The current approach is to say, "Let's take our rates and when rates go up, we force them down." It's the mean-reverting process. Basically, we're going to make sure the short-term rates stay within a band. In another way of doing that, there's a term structure of volatility. If we start with regular volatility over time, the volatility decreases, and it should be given the pure random walk.

First we'll look at the mean-reverting process.

MEAN REVERTING DRIFT (CIR & VASICEK)

r =short rate

θ =long-term reversion level

κ = speed of mean reversion

σ =volatility of random process dz

β =1/2 for the CIR model, β = 0 for the Vasicek model

$$dr = k(\theta - r)dt + \sigma r^{\beta} dz$$

We have a variety of parameters to estimate. First is the long-term mean reversion level. You have kappa in the equation, but it can't be equal to infinity. As soon as I deviate out a little bit, I can even snap back down a bit at the mean-reverting rate. When kappa is very, very close to zero, it takes a long time. It can wander around a lot.

I have volatility and a random process. We will consider two types of stochastic types of models. One is the Cox-Ingersoll Ross (CIR) model. As my rates start going towards zero, my volatility is multiplied by zero, which is when my volatility starts going down. This prevents me from dropping through to having a negative rate. In the Vasicek model there is no restriction on the rate, so I can have a negative rate. We already have some idea in terms of what the trade-offs are. One of the reasons the Vasicek model is preferred is because there are a lot of closed form solutions for various types of derivatives. Obviously, if you are using CIR, you're forced to do a more difficult process.

What does this mean in terms of our short rates? Let's look at a regular lognormal process as shown in Chart 7. We started down at 5%. Because it is simulated, it wanders all over the place and ends up with grossly high peaks very early on. I took the same set of random numbers and put them into the Vasicek model, and, as soon as the thing starts moving away from 5%, which I assume will be my long rate, the thing starts upwards. That's because the peaks aren't nearly as high as the lognormal process. The CIR model is different. As I plot these things out, it is very similar in terms of its feel.

What are the parameters? We have volatility and as long as we're dealing with stochastic processes, we'd better have a way of thinking about that. Again, as Rishi pointed out, how do you estimate that? How you get there is a question of different estimation techniques and different underlying models. How could I deal with that? We should be able to look up on our Bloomberg screen and see what short rates are. It is not a difficult process.

Now we have three other parameters here. Kappa is something you've already seen. That's the speed of mean reversion. Lamda is another parameter, the price of risk. There are certain people who are risk lovers and certain people who are risk haters. People, in general, are averse to risk, and that's going to affect the way our yield curve looks. If it goes up in slope, then people are risk-averse. The longer my money is with somebody else, the more I'd better get paid in terms of my risk premium. In other words, I have five parameters. So the first price I'm paying is an extra two parameters.

The second thing here is that there are certain limitations about the old world structure that we have. What I've done here is I've taken very short rates. I've started from the particular short rates and the particular set of parameters and asked, "What does my yield curve look like? Is it up and sloping or down and sloping?" It turns out that I'm very restricted in terms of the types of yield curves I can really see. If I have a yield curve that goes up sharply and flattens off, I can model that with CIR. If I have a hump, I can't model it unless I have a conversion curve called the impossible. I've set three parameters. I'd added some realism to it, but I'm still paying a price because I can't model every possible yield curve. That also means I can't model every move of the yield curve.

The next generation after that is a single-factor model and the Black-Derman-Troy (BDT) model. This model is a little bit different. Instead of having a mean-reverting process, let's assume that we have a standard Brownian motion. My short rate has changed. Then, over time, my volatility starts decreasing, as shown in the formula below:

SHORT-RATE TIME DECAY MODEL (BDT)

r_t = short rate

u_t = median of the short-rate distribution

σ_t = the short-rate volatility

z_t = standard Brownian motion

$$r_t = u_t e^{\sigma_t z_t}$$

ν = decay constant ($\nu > 0$)

$$\sigma_t = \sigma_0 e^{-\nu t}$$

I can think about this as a tree as shown in Chart 8. The tree starts at 5%. I can random-walk over a period of time either out to six and down to four and so on and so forth. Instead of the tree looking like it does in Chart 8, it would start going toward the extremes. I get a funneling-in effect.

The one advantage that I also have over the other previous models is that I can price swaps and caps off these models. This is a more general model. Now we talk about model-setting parameters.

One of the problems we always have is, what's the length of the history? As Rishi mentioned, this is very significant. You have a very long history, and a lot of your data are very relevant. If I took my interest-rate history and went back 100 years, I know it's going to look quite different. Up until well into the 1960s, interest rates didn't change that much. Then in the early 1970s interest rates were up. Now you have the current regime, which is sort of intermediate. How far do I go back? On the other hand, if I used the last ten days, experience is statistically unstable. If I look at it from a statistical point of view and the rate was known, it was fixed over that ten-day period. What we all know about statistics is that there is just an unreliable estimate, so we have a trade-off between a long history and a short history. Some parameters like my risk preference are totally unobserved. I can fit them. I have no way of validating this. This violates one of the early things I was saying about the psychology of the model. We have no idea whether we should tweak it up or tweak it down. It is just a number. We also have a question about different combinations of parameters. We may have a parameter that has a conflict between a mean-reverting parameter and the speed of mean reversion. There's a trade-off. If we can adjust one, we can adjust the other and end up with the same yield curves. We end up with very similar results. We had no way to look at that history to tell us which was the correct way. We had no psychological insight into a lot of cases to tell us which way to go. The parameter changes are often correlated. If I have a model that looks at long and short rates, like a multifactor model, the long and short rates will have a correlation. You have to include the correlation in terms of the way we were thinking about the model.

What are the limitations of CIR? The main one is that you can't get all types of yield curves. You have more flexibility than you would have in, let's say, a random walk of Black-Scholes, but you still have a restricted yield curve.

In the BDT model, you have another thing that is often unobservable—the term structure of volatility. You can assume that volatility is dropping off exponentially. I don't know whether an exponential level is a good level.

This pretty much answers what Rishi was talking about. There is probably a little bit more detail in some ways, but he mentioned trading frequency is very important. Obviously, October 1998 is very important. I spent a long time looking at commercial mortgage statistics. I look at the mortgage. Then there is the series of commercial mortgage statistic spreads that say you've got to be kidding. There is nothing traded there for two months. There's no way those spreads mean anything at all.

You have situations of worldwide trading, except the data are not synchronized. I have a series of closings of the German and Japanese exchanges. There's a lot of information that's taking place across time. How do I synchronize that? There is going to be 12 hours of information from the time that one closes to the time the other one closes. I have to adjust it. Rishi also mentioned bid offer spreads. What do I use? Bid, offer, or mid? Whenever we are doing a parameters fit, we have to get to some fitness of fit measure. There's a variety of fitness to fit measures. Which one do you choose? It matters sometimes. Some pick outliers and emphasize them. Some underemphasize outliers, and some are concerned with recent history. Some are concerned with long-term trends. It matters that they have to interact with the type of models that have good yields.

We have a variety of different techniques—maximum likelihood methods, method of moments, and Kalman filtering. In the best of all worlds, it wouldn't matter what type of estimation technique you'd use. In actuality, sometimes it can be very, very material. For a lot of these methods, like maximum life yield, you have conversion problems and conversion problems situate. You can't tell that you have conversion problems.

Whenever you fit a model you have two things. You have historical rates and then you have the market implied rates. Generally, you make a decision that says, "I'm going to estimate six of my parameters using historicals. I'm going to use two of my parameters to tweak the model to fit with the current market." Some of the traders sitting next to you might say, "I'm going to take four and three and they're going to be a different four and a different three." It's a judgment call. This is all too important when you're thinking about it. What's in the model risk? If it's not just a model, it is a mathematical entity. It's also about how is it used, how it is controlled, and where the prices come in. How is the estimation being done?

I think the main one I want to talk about is the appropriateness of a model. Another way of looking at this goes back to what Rishi said. For one thing, I have the same bias as he does. I have an inexperienced trader with a great model versus an experienced trader with no model. I'd take the experienced trader any day. I've seen it happen too many times on the trading desk. That's where the psychology comes in. The experienced trader understands the psychology, knows how that market is going to move, knows how they used to hedge, and knows when the market breaks down.

Another thing that's important is that if you have two models that explain equally well, take the simpler one. There are fewer parameters. It is probably easier in terms of the psychology. The other thing is, it's also easier to isolate problems with lots and lots of experience.

Mr. Bruce E. Booker: Your presentations add a lot and make us think. These are just models; they're not real life. Don't trust them. What if I change one of my parameters today because I don't trust what the model's saying versus what I would have to do in real life in either buying or selling or pricing something?

Mr. Noma: This is one of the reasons why you do stress tests. Stress tests give you some idea of the outer limits of your models. Those should be done before I get into the crisis situation. I should be comfortable with them or be able to say, "This model totally fails on the stress test." Therefore, as soon as I see the red flag coming up, I know it's time to yank the model. It's time to pull the trader into the office and have a major conference in which we say, "We have to change things in terms of how we think about this. We have to have senior management really focus on this problem."

From the Floor: Rishi, you mentioned sometimes it's appropriate to use historical volatility and sometimes it is appropriate to use implied. Could you say when it's appropriate for each?

Mr. Kapur: The chart that I showed basically gave the idea that you cannot really tell when it is appropriate to use what. It just shows what my experience is. I don't know if you can take this as a back-of-the-envelope rule. In cases of high volatility environments, implied volatility tends to be a little better than historical volatility. The market will adapt rapidly, so your implied volatility would reflect the fact that the market expects the world to be more volatile, but the historical volatility may take some time to catch up.

Mr. Noma: There are also particular instances in which we know just a priori that historical volatility is not the correct way to look at things. For instance, in January, when Brazil decided to stop defending its currency, it was after you already had the first break of the currency. Obviously we know, at that point, that historical volatilities are not appropriate at all. We have no road map, but we know that one of the road maps is totally wrong because the psychology of the market has totally changed. This is one of those situations in which the traders come in and say, "I know the market's really different. It's time to do something and think about it differently."

From the Floor: Rishi, you spoke about how quickly we need to change the parameters on these models. That made me wonder about how long a model is valid. Elliot, you just spoke about the red flags that help you decide when it's time to change. My question is, what is the time frame? I am hearing you say less than a month, so is a model a couple of weeks?

Mr. Kapur: There are really two sides. One is the model and one is the inputs to the model. My recommendation is you should always monitor the inputs to your model and estimate them; that is if it doesn't take very long to estimate them. If you have a really slow method of estimating parameters, it might take you a couple of hours to run. You can do it on an overnight basis and maybe monitor them every day but not change them until they differ sufficiently from your existing parameters. The models themselves can last for a long time until the market psychology changes (for example, the Brazilian currency crisis). With inputs, it's important to monitor what your input should be. In some cases, for example, for a lot of the banks, the valuation is done at the end of the day so every day they will use the latest stock prices and interest rate curves to value it. During the day when they make their trading decisions, the inputs that they are using are derived from calibrating to the market. Every time you need to run, you have to calibrate and the numbers may be different.

From the Floor: If that's the case, that makes me question whether the model is good for that day if you have to change it the next day based on new information. Was it accurate that day?
Mr. Kapur: You might have an interest-rate model. One of the things you might want to do is to say, "I want my interest-rate model to be able to at least match market prices of bonds." That is a reasonable expectation, and as the market prices of bonds change you may have to change your inputs. The fact that you have to change your inputs every day or every time the market moves doesn't necessarily mean your model is wrong. If you're having to change your model or if you're having large swings in your inputs all the time, then that's a red flag.

Mr. Noma: The life cycle of models and parameters are very, very short, so one of the things that should be done on a very frequent basis—frequent depending upon what market it is—is to do an out-of-sample-type test in which you use some of the market information you're receiving on a given day to fit the parameters. Then try to figure out whether you're predicting other indicators you see in the market or how well your model is performing. Let's say you have a yield curve model in which you're predicting what the yield curve is. You fit your parameters to the yield curve, and then you predict what the swap curve is doing. Assuming your model is correct, your swap model would be pretty accurate in terms of what the swap rates are. If you start seeing deviations, then you know it's time to really look at this and ask, "Is the market psychology really changing on this?"

Now the other way is to also look at your profits and losses (P&Ls) and your traders. If your traders are saying they are perfectly hedged and you see their P&Ls swinging around showing that there are irregular patterns, then it's time to go back and ask them how much confidence they have in their models. How are things really changing in the market? So, talking to the controllers and having them monitor, on a daily basis, where their P&Ls are versus where their value-at-risk is versus the historical patterns of the P&L, is a very important tool in terms of looking at the life cycle of the models.

From the Floor: This is also a general question about a model of interest rates. From a U.S. banking perspective, everyone agrees on how you should approach it. You should upgrade a model often and also try to calibrate your parameters once a day. Look at some of the actuarial literature, for example. There are some articles on how to generate interest-rate scenarios and, if you read very carefully, the model is a not necessarily arbitrage-free. So I think the purpose of using actuarial-based models would be different. I'm not sure it is for reserving purposes or for pricing purposes. Do you think we should use different approaches for different perspectives? To price a marketing instrument you have to use a model that is updated to the latest information. For some kind of model for long-term reserving purposes, you can use them all like an economic type of model. I would like to hear comments.

Mr. Kapur: Yes, there is quite a bit of difference between pricing and risk management. In pricing you like to be calibrated to the market because you want to get prices that are close to the market prices. For risk management, a lot of your models will be statistical and econometric models based on historical series. In that case, your frequency really depends on how often you use them. Now, if you are not pricing or if you are valuing liabilities, for example, you don't do that on a daily basis. You might do it every quarter at an insurance company, but at least every year you have to do them. Every time you use it, you may want to calibrate your parameters. It will not be necessary for you to be calibrating parameters every day because you're not valuing your insurance liabilities every day.

CHART 1
3 MONTH VOLATILITY FORECAST FOR S&P 500

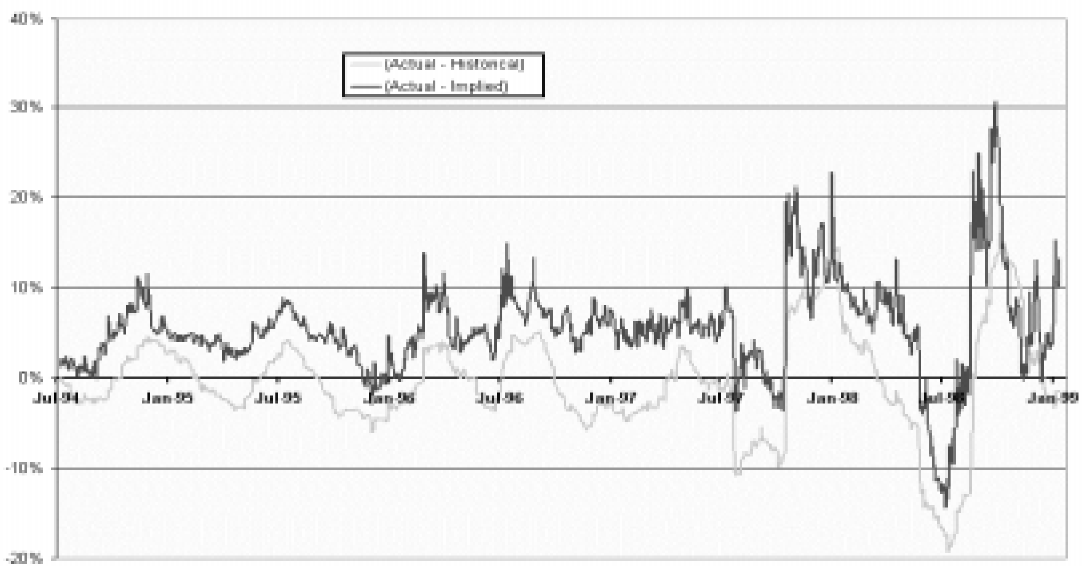

CHART 2
OVERNIGHT INTEREST RATES IN HONG KONG

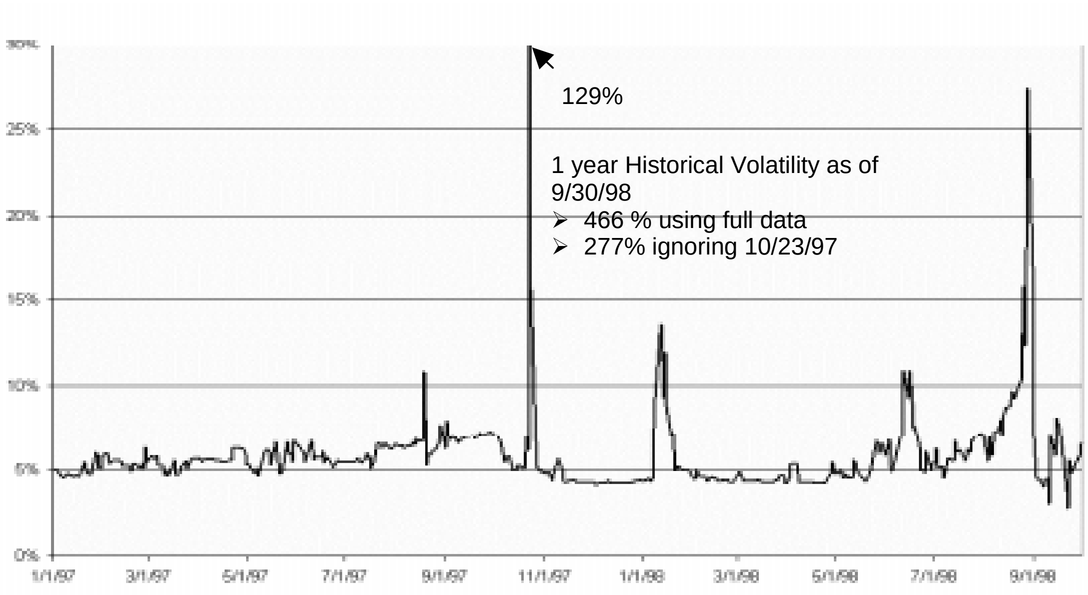

CHART 3
FAURE SEQUENCE—1000 PATHS
FIRST VS. SECOND DIMENSION

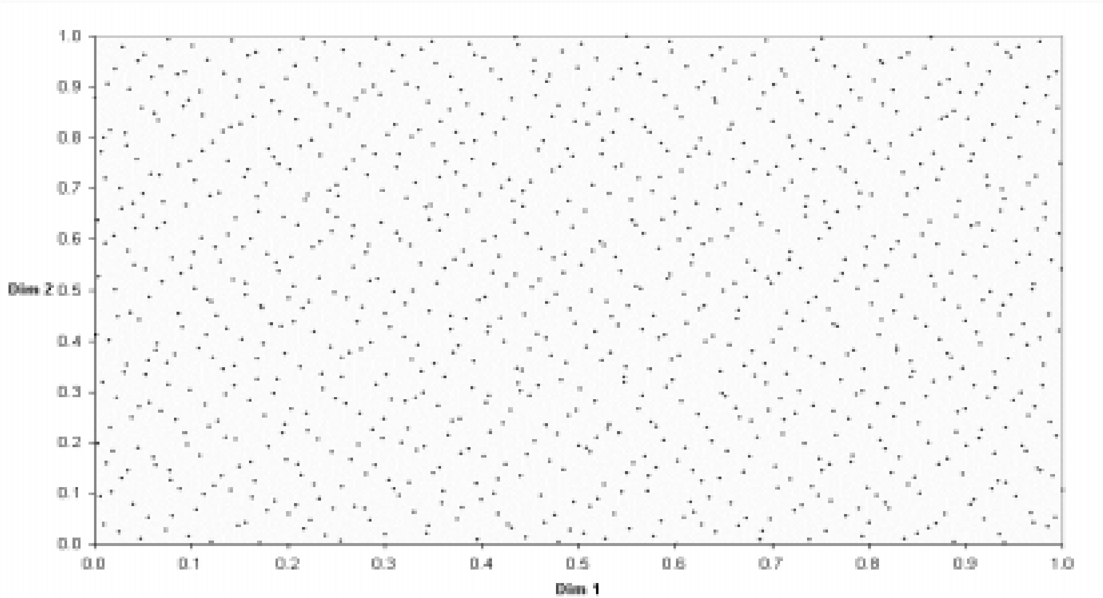

CHART 4
FAURE SEQUENCE—1000 PATHS
RELATIONSHIP BETWEEN THE 99TH AND 100TH DIMENSION

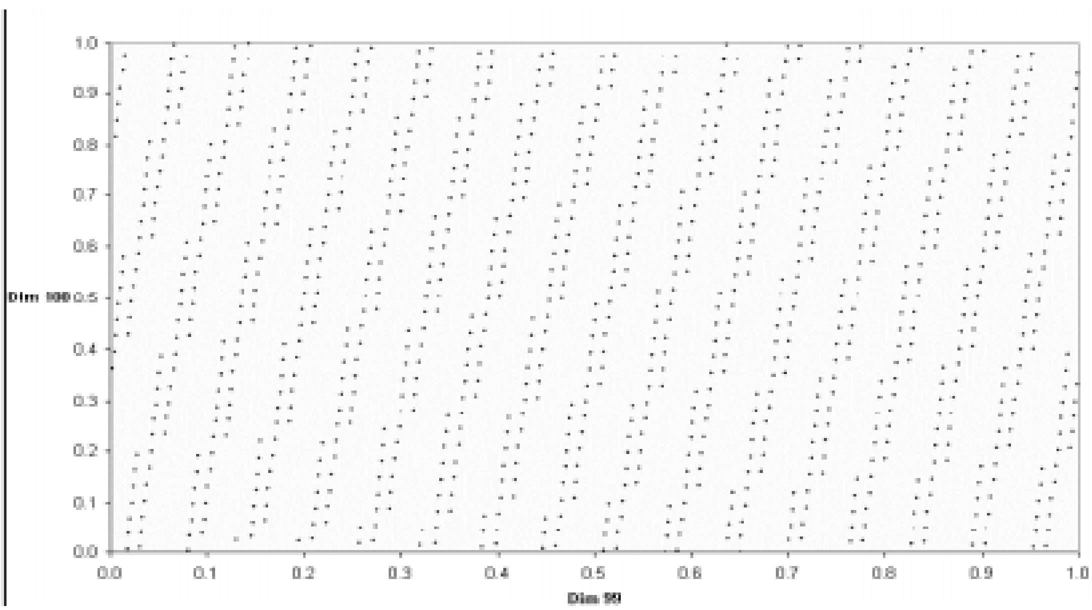

CHART 5
U.S. TREASURY YIELD CURVE AS OF 4/20/99

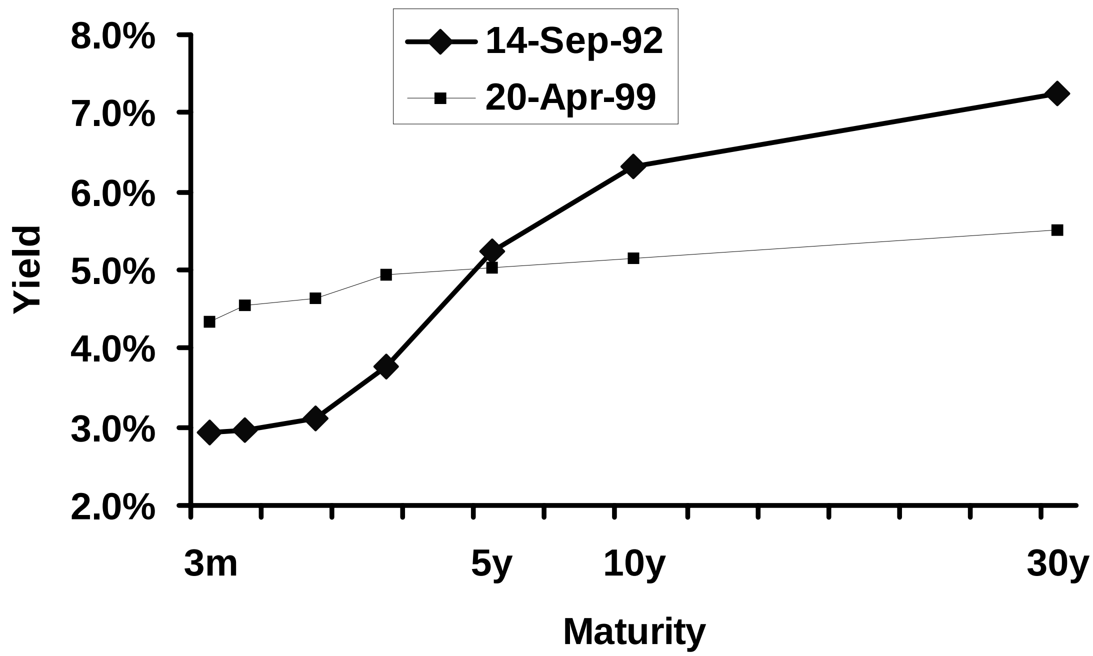

CHART 6
U.S. TREASURY YIELD CURVE

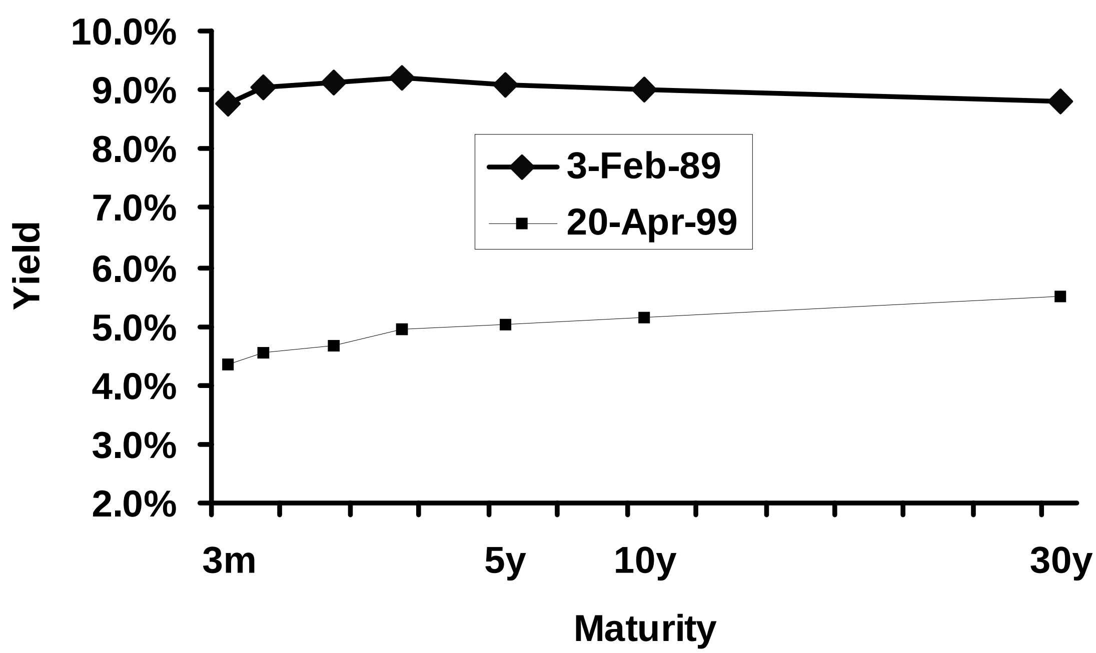

CHART 7
EXAMPLE OF MODEL-GENERATED
FUTURE SHORT RATES

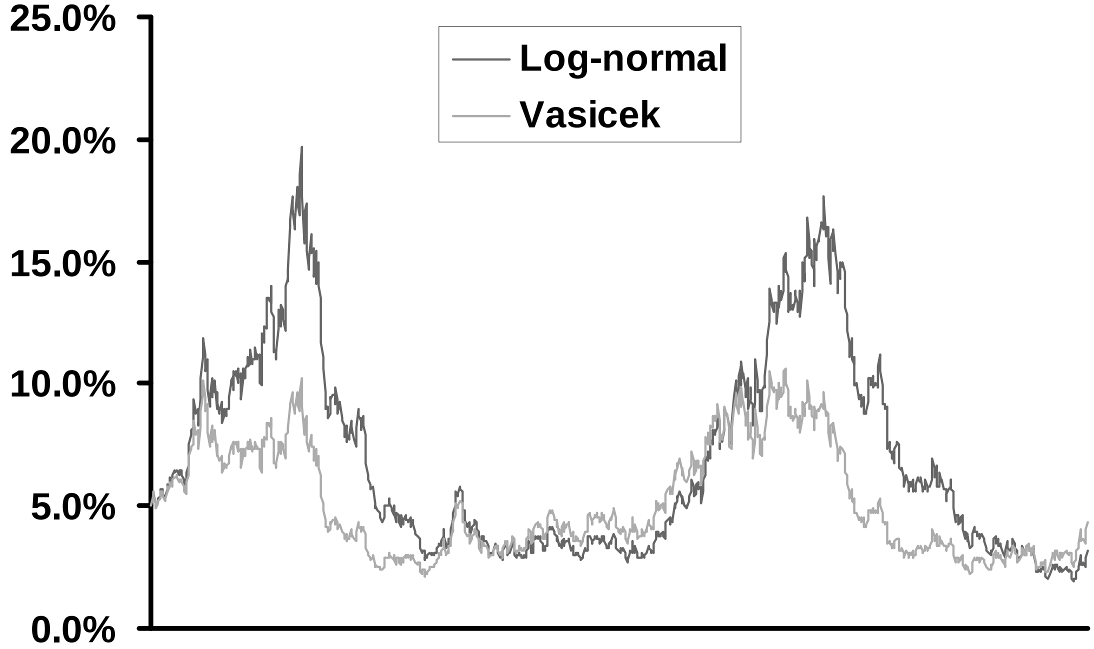

CHART 8
PROPERTIES OF TIME-DECAY MODEL