1996 VALUATION ACTUARY SYMPOSIUM PROCEEDINGS

SESSION 4

Interest Rate Model Risk

Michael E. Mateja, Moderator

David N. Becker

Peter Fitton

Douglas A. George

INTEREST RATE MODEL RISK

MR. MICHAEL E. MATEJA: What is interest rate model risk? Technically, you could consider it one of those dreaded other risks, a C-4 risk. Classically, it's the risk that judgments that you make about reserve adequacy based on a given interest rate model are in error. For this purpose, I'm going to define the seven scenario test that is embedded in valuation law as an interest rate model.

Is there really such a risk? Is it possible that cash-flow-testing results over the last 15 years or so based on the seven scenario tests have not really produced a sufficient indication of reserve adequacy? I think the answer to this question that will come out of the substance of the presentations to follow is a qualified yes. Certainly, for those who passed all of the seven scenarios, the statutory definition of adequacy was met. But, statutory adequacy may not be the same thing as an adequate reserve, or at least that's the premise of this entire session. Before probing the difference between the statutory definition of reserve adequacy and the theoretical definition of reserve adequacy, it might be useful to review briefly the history of the seven scenario tests that are embedded now in statutory valuation law.

Cash-flow testing is about 15 years old, and I was heavily involved in the process that culminated in the adoption of the seven scenario test embedded in valuation law. At the time the seven scenario test was adopted, we fully realized that a seven scenario test was simply a stress test. It did not address the issue of level of adequacy in a probabilistic sense. To assess reserve adequacy in probabilistic terms, you need to start thinking of stochastic tests. We fully realized at the time that the seven scenario test was a stepping stone into a more robust testing process. Practically, we needed to wait until the actuarial profession caught up. So, here we are 15 years later, and we're still stuck on step one.

What I hope will happen in this session is that our understanding of what it means to have an adequate reserve based on cash-flow testing will be enhanced. We're going to try to bring you up to date on the state of the art in this area. I hope we'll establish that the results of deterministic cash-flow tests may not support a conclusion of reserve adequacy. We're going to take you into the world of interest rate models and interest rate model generators, which represent the next major step in terms of the work of a valuation actuary.

I also want to address the practical side of this question because I know there are many companies where the primary goal is to meet the literal requirements of the law. If the law says pass seven scenarios, then these companies are only interested in passing seven scenarios. I take the position that our profession is grounded in the tradition of good science. I think we have an obligation to understand and apply the best science that we can in the discharge of our responsibility to both management and regulators. Remember, a statutory opinion is rendered on behalf of the board of directors of the company, so a report from a valuation actuary represents a report to the board as well as a report on behalf of the board. When you think of an actuarial opinion in that light, you need to ask yourself this question, do I feel comfortable rendering the opinion to my board? When I was preparing the actuarial opinion a few years back, I started to get very apprehensive about just what seven scenarios meant. I have talked to a lot of people about interpretation of results of the seven scenario tests in recent years. I believe that concern in this regard is growing.

What is an adequate reserve from a technical standpoint? I think one simple definition is that the probability that you can mature your obligations with the reserves (assets) you're holding is somewhere above 50%. I don't want to debate just what level of probability that might be, but it could be has high as 70, 75, or even 80%.

There's a recent article in the *Financial Reporter*, by Alastair Longley-Cook, in which he attempted to associate a probability with the various statutory tests. What he ultimately did was use an interest rate generator to produce many tests and then cleverly allocate the tests to the prescribed seven tests. The results yielded a crude probability associated with each test. If you simply generate the stochastic tests and develop a probability distribution, you could get the required probabilities directly.

I submit that we should be worried about the probability that reserves are adequate more than we are worried about passing the seven deterministic tests of statutory adequacy. Our opinions today are relied on by a very wide audience: our management, regulators, and the public at large. The consequences of not getting opinions right are probably far more severe than any of us can really imagine. The fact that we may be rendering an opinion based on what I would call faulty science leaves me very concerned. I hope after this session that you'll have an appreciation of just why I have that concern.

Here's our plan. Very briefly, Doug George is going to provide a practitioner's perspective. Doug is with Avon Consulting Group. Peter Fitton, who is an associate at SS&C/Chalke, is going to address the theoretician's perspective. Peter has a very sound understanding of the underlying theory. David Becker, who bats cleanup, will provide what I call the management perspective.

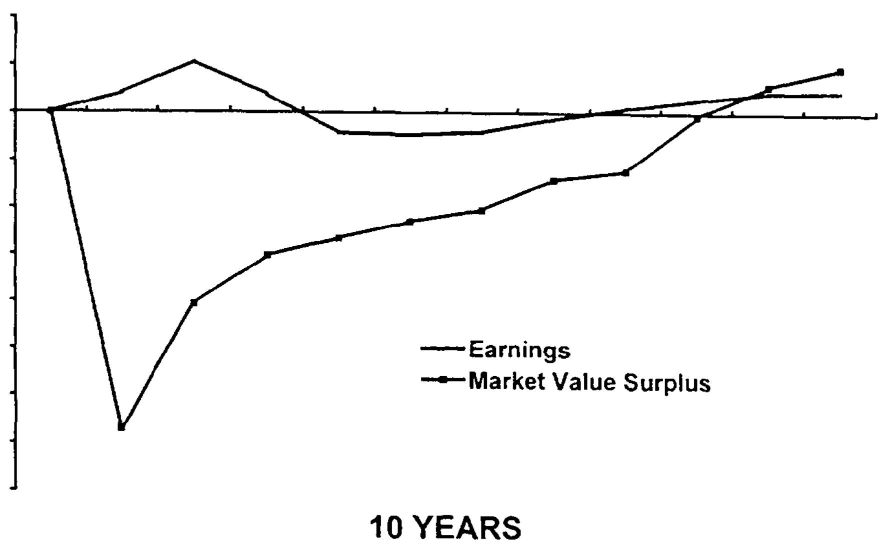
MR. DOUGLAS A. GEORGE: As Mike pointed out, over the last few years most of us have based our reserve adequacy opinions on the New York seven scenarios. Now, the New York seven scenarios are a pure stress test. They represent what might happen to our blocks under deliberate interest rate movement and under shock interest rate scenarios. They're really not realistic, and they're not meant to be. Maybe five or ten years ago, it was adequate to perform our cash-flow-testing analysis based on the New York seven. Five or ten years ago we didn't have the level of understanding of interest rate risk and asset/liability interaction that we do now. We also didn't have our asset/liability models to the point where they are today in terms of credibility. So, while I think the New York seven were appropriate five years ago, I think it's now time to move forward and enhance the analysis.

Now let's take a look at why I think more enhanced analysis is necessary. Chart 1 shows a typical single premium deferred annuity (SPDA) block and supporting asset portfolio, and how it performs under the pop-up scenario. This, of course, is our "killer" scenario for SPDAs. Let me describe the pattern that we might follow in terms of our earnings and our market value surplus under this scenario.

Immediately we get a bump up in earnings due to a shock lapse and the fact that we have some surrender charges on our block that we can make money on. We get out to the middle years of the scenario, and our surrender charges run off, and we tend to lose money because our portfolio yields aren't where they need to be to cover our credited rate. Finally, we get out to ten years or so, and we start to make money again because our portfolio has rolled over and we're able to earn the yields that we need to make our spread.

CHART 1

SPDA Pop-Up Scenario



The market value of surplus produces a pattern that is related to the earnings. We get an immediate pop-down, which relates to the pop-up in the interest rates. The market value of surplus then slowly creeps up as we make money, tends to level off somewhat in the middle years where earnings are flat, and finally, we hope, by the end of the tenth year, creeps back up into the black as we begin to make money again. What we have designed here is somewhat of a game. Can we get our market value surplus back in the black by the end of the tenth year?

If we were to go out more than ten years, I think we'd have a lot better chance of winning this game. If we went out 15 years with this projection, it would be a lot easier to win. It would be a lot easier to pass this scenario.

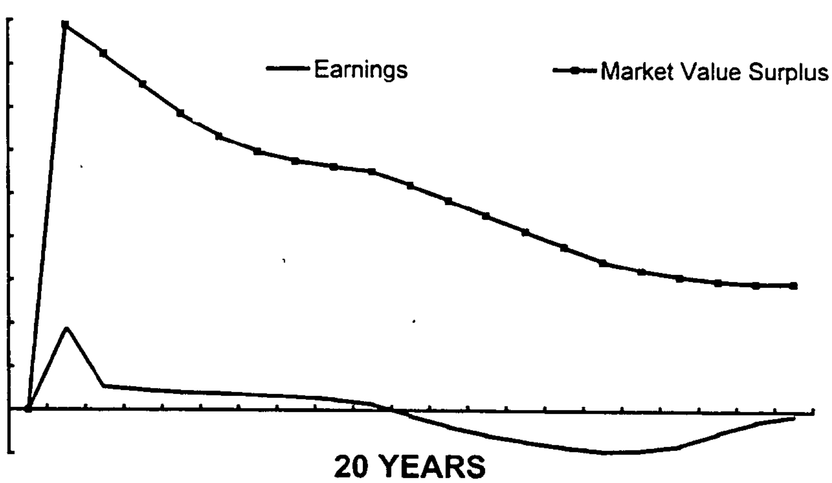
Some people might think that the reason we only use ten years is because most of our block runs off by the end of the ten-year period, but the real reason that we only use ten years is because Regulation 126, as first adopted by New York, only allows you to use ten years. You must get written approval from the New York Superintendent to use a period longer than ten years for SPDAs. So, the regulators really designed this to be a test to get your market value surplus back in the black by the end of the tenth year.

I think this is a worthy test, but my point here is that I think the information that it provides on our real reserve adequacy is somewhat limited.

Chart 2 shows a traditional life block under a pop-down scenario. This would be the most detrimental scenario for traditional life business. We get an earnings pattern that starts off high, but slowly goes down and drops below zero as time goes on. Our assets are rolling over, and the new yields that we're picking up are not able to cover the fixed rates that we're crediting on our traditional life business. Finally, we get out 20 years or so, and our earnings tend to come back up because our liabilities are running off and we're earning money on surplus at this point. So while earnings can still be negative, they tend to move back up closer to zero.

CHART 2

Traditional Life
Pop-Down Scenario



Our market value of surplus takes an immediate pop-up in the first year in accordance with the pop-down in interest rates. The market value slowly decreases over time, and we hope stays above water by the end of the 20th year. This again is somewhat of a game. In order to win this game, we need our traditional block to run off fast enough before we lose money in those years toward the end of the projection that might bring down our market value of surplus below zero. This game is a little easier to win than our SPDA game, but it's still a test to see what happens to a block under deliberate, nonrealistic interest rate conditions.

Like the SPDA under the pop-up scenario, I think this is a worthy test, but I think the information that it provides on long-term reserve adequacy is still somewhat limited.

Chart 3 shows what happened to interest rates going from the end of 1993 to the end of 1994. Now, I've had the privilege of working with a number of companies on cash-flow testing over the years, and I work with some big annuity writers. In 1993, my annuity writers didn't have much problem with the pop-up scenario. Interest rates were obviously very low. But when we got to 1994, many of my annuity clients were having a lot more problems with that pop-up scenario -- some people failing, some just borderline passing.

Now, my traditional life writers had somewhat of a problem with that pop-down scenario and also the falling scenario in 1993, because interest rates were getting so low. But in 1994, my traditional life writers didn't have a problem with those scenarios anymore.

These results aren't much of a surprise, but my point is, I'm not sure I believe that the long-term reserve adequacy of these companies was affected that much because interest rates changed 2% in one year.

I do think that the New York seven provides insight, but I think it's time to start doing some more enhanced analysis beyond relying on the New York seven to determine your reserve adequacy. Most of this enhanced analysis can take one of two forms. It can take the form of pricing, determining fair value of liabilities, measures like duration and convexity, and key rate duration. It can also take the form of additional scenario testing.

To do this enhanced analysis, the first thing we need is an interest rate generator. An interest rate generator can be one of the most critical pieces of our asset/liability analysis. A lot of us spend a great deal of time creating a very sophisticated asset/liability model, but very little time evaluating our interest rate generators when we're doing asset/liability work. The interest rate generator is the driver of the asset/liability analysis, a faulty generator can produce misleading results.

My own definition of an ideal interest rate generator would: (1) include all possible scenarios; (2) exclude impossible scenarios; and (3) assume correct probabilities. It seems like it shouldn't be too hard to make a generator that meets my criteria. However, I've never found a generator in practice that does. In practice, I've found that generators are designed for different purposes, and each one serves a purpose well, but also has shortcomings in other areas.

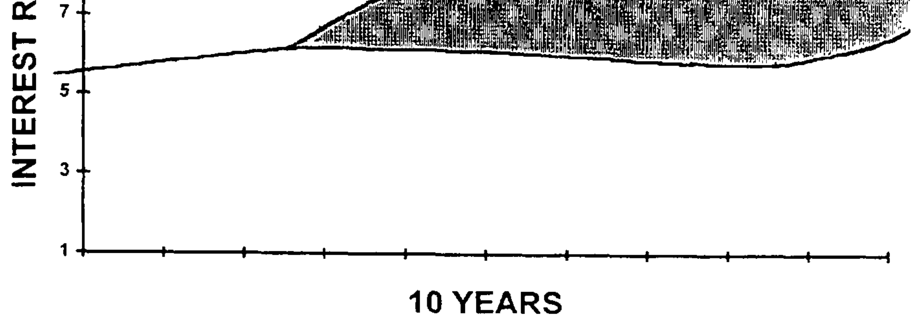
Primarily, I've found that generators are designed for two purposes. The first purpose is for pricing, the other for scenario testing. The pricing generators are characterized mainly by being arbitrage-free. They are based on the approach that you can take the current yield curve, break it down into forward rates, and use them as the "basic" prediction of where interest rates are going. They're designed for option adjusted spread (OAS) analysis, and that's why they need to be arbitrage-free, so you can't arbitrage the yield curve. They are used to calculate measures like effective duration and convexity.

The other general type of generator that I've seen is designed for scenario testing. These are not arbitrage-free, but are based more on historical statistics, based on interest rate movement in the past. I just want to give a brief introduction into these concepts. Peter Fitton is going to elaborate a little bit more on the different types of interest rate generators and how they're characterized.

First, let me show you my definition of an ineffective generator. When I started a few years ago doing fair value of liability analysis, I picked up a generator that had a pretty good reputation, one that came from Wall Street, and started running my liabilities through and started with about 250 paths. I came up with a liability duration of 1.72. Then I said, maybe I don't need 250 paths, what if I use 100 paths? Maybe that can still give me an adequate liability duration. I came out with about the same duration number. I went down to 50 paths and had still about the same. I went all the way down to one path and still had about the same duration. I figured, well, either I have the best generator in the world or I have the worst, because my generator can price all my liability options with only one path. Unfortunately, that's impossible. The real answer that I found out under further investigation was that most all the paths were following the same interest rate movement. Chart 3 shows a graph of the path coverage.

CHART 3

Ineffective Generator
100 Paths -- Rate Coverage



For the first three or so years, all the paths followed the same pattern of interest rates. Going out even five or six years, there was still very limited interest rate coverage. This wasn't a problem with the algorithm behind the generator. It was more of a problem with the way the algorithm was implemented. The generator was implemented to price mortgage-backed securities, where the options were out five to ten years. But when I used it for an SPDA, which has much shorter options, especially an in-force block where the options are maybe one, two, or three years out, this generator was not effective.

This I think is just one example of the need to look underneath your analysis. When you're doing this type of analysis, you can't just blindly use the numbers that are coming out. You really have to roll up your sleeves and investigate why the numbers are coming out the way they are, because they can be very misleading.

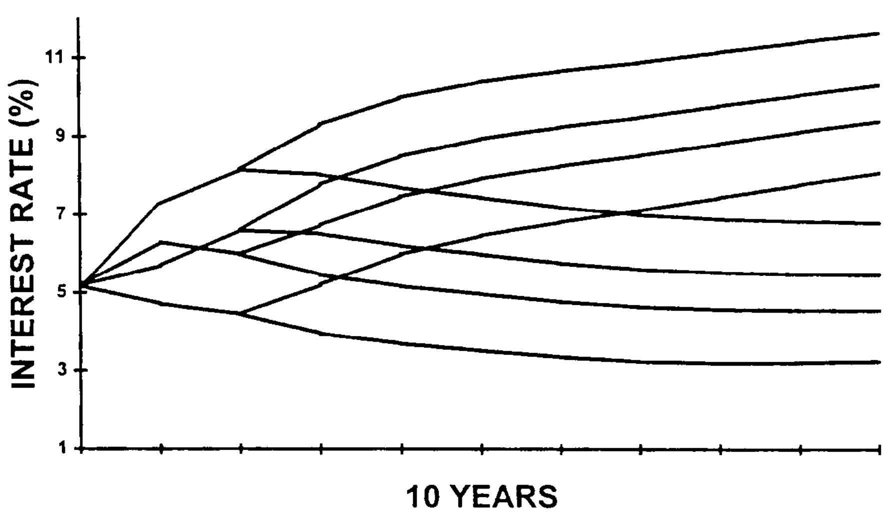
I picked up a generator from Capital Management Sciences (CMS). This is a one factor model, based on the Heath-Jarrow-Morton interest rate model. It's arbitrage-free; it's designed for pricing where the one factor is based on the short-term interest rate. A one factor model tends to give you good parallel movement or parallel shift movement to the yield curve. Different studies that I've seen have shown that parallel movement to the yield curve accounts for between 80 and 85% of all yield curve or interest rate movement, which is pretty high, but does not cover all interest rate movements. This model is also characterized by having a volatility input. It's got mean reversion. One of the differences in this model versus others is that all paths are assigned the same probability, whereas most arbitrage-free models will actually assign more probability to the central-most paths. This model designs the paths so each one has an equal weighting.

The model comes out of the BondEdge asset analytical software, asset pricing software. It holds credibility with my investment people because it's coming from their side of the shop, so if I'm using my analysis to change the way I manage my asset portfolio, I have some "buy-in" from my investment people because I'm using their analytics to produce the results. It also gives me consistency because if I use this model to do a duration calculation for liabilities and match it up to the durations on my assets that are coming through BondEdge, I have an apples-to-apples comparison of the results.

Chart 4 shows the first eight paths that come out of this generator. This generator provides pretty good coverage of future interest rate possibilities with those eight paths. It doesn't provide all the up-and-down movement because it's not really designed to be a scenario generator, but it does provide good coverage of future interest rate space. The idea obviously is to try to optimize how much coverage of interest rate space you can get with the fewest number of paths. Although the generator is designed for pricing, I have seen some people use this generator for scenario testing. Usually, the only time I might recommend that is if your models are so big that you only can really run on a small number of paths. Then it might be adequate to run something like this because you do get good coverage of future interest rate space with only a few paths. But what you don't really get is all the up-and-down movement that you'd like for more realistic scenarios.

CHART 4

CMS Generator
8 Paths



As part of this analysis, we created a hypothetical asset/liability portfolio for an SPDA block, and we ran a bunch of tests. I ran duration analysis through the CMS generator and found out with about 16 paths I could get pretty good convergence (Table 1). I get a pretty accurate number with about 16 paths -- it's very close to the value produced with 128 paths. I can be pretty comfortable using 16 to produce results. In my own analysis with real data, I've actually seen pretty good coverage with even only eight paths.

TABLE 1

CMS Generator

Paths	SPDA Duration	Change from 128 Paths
8	1.40	0.17
16	1.53	0.04
32	1.56	0.01
64	1.58	(0.01)
128	1.57	

One thing you need to watch out for with these generators is your volatility input (Table 2). A change in volatility input changes the duration of my SPDA block pretty significantly. For this reason, you need to make sure you're calibrating your model. You need to make sure you have a reasonable volatility. Unfortunately, the only way you can really calibrate that is by turning to the asset portfolio and pricing assets so that they match market value because you really don't have accurate liability pricing to calibrate to. But the volatility input can definitely affect your results.

TABLE 2

CMS Generator
Volatility Change

Volatility	Duration
12%	1.53
16%	1.84

What happens with the increased volatility is that more of the down scenarios go lower with the increased volatility. What happens is we bump into our interest rate guarantees more often. That tends to increase the value of the liability under the down scenarios, and that's what brings up the duration.

I've run some results using the CMS generator to actually do scenario testing analysis (Table 3). Again, we can get convergence with about 16 paths. We're passing about 92-94% of our scenarios. When I looked at the ones that I'm failing, I notice that it's basically the up scenarios that I'm failing. In our one factor model producing mostly parallel interest rate movements, when the scenarios get to those higher levels, those are the ones that we are tending to fail. By the way, the model passed all seven of the New York seven scenarios when we ran it through them.

TABLE 3

CMS Generator

Total	Pass	Ratio
8	7	88%
16	15	94%
32	30	94%
64	59	92%
128	118	92%

Consider another generator coming from Wall Street. This one is from Global Advanced Technologies (GAT). It has a one factor model in it. It also has a brand new two factor implementation as well. The probabilities for this model are based on the GAT's linear path space, where there's a binomial lattice that underlies a trinomial lattice that they use to project interest rates. The probabilities are based on the number of paths that underlie the trinomial lattice, i.e., the number of paths in the binomial lattice that underlie it.

One of the benefits here is that you also have a two factor model, so not only can you get parallel interest rate movement, but also you can get changes to the steepness of the yield curve. Going forward, yield curves can either get steeper or flatten out, and that's the second factor coming into play.

Different studies have shown that a steepness change to the yield curve can account for 10-15% of yield curve movement on top of the 80-85% that you're getting with parallel movement. So you're capturing pretty close to 100% of yield curve movement with a two factor model. Again, one of the benefits is I have some "buy-in" from my investment people because this is a "Wall Street" model, and if my investment people are using the GAT software, they're going to buy into the results that I'm getting.

Table 4 shows a comparison of the one factor model in GAT versus the one in CMS. The GAT model produces a lower duration of our SPDA, but it also produces a lower duration for assets so that, when you look at the duration mismatch between assets and liabilities, they're about the same for both GAT and CMS. My point here isn't so much to try to compare these two or to say that one is right and one is wrong, but it's more to point out that all interest rate models do have biases, and you need to take this into account. That's why it's especially important to run your liabilities through the same model that you're running your assets through. In this way, you can minimize the effects of those biases by looking at your duration mismatch. You can get more consistency in the results that are coming out -- the duration of assets and the duration of liabilities, so you can make an apples-to-apples comparison.

TABLE 4

Duration Comparison
One Factor Models: ~ 100 paths

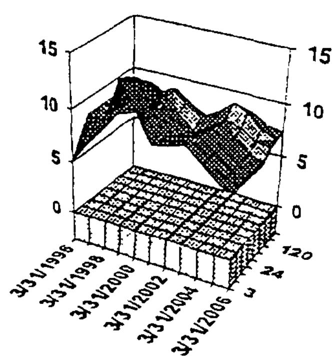
	GAT	CMS	Difference
SPDA	1.34	1.57	0.23
Assets	2.63	2.95	0.32
Mismatch	1.29	1.38	0.09

Chart 5 shows some of the interest rate scenarios from the GAT two factor model. You can see the effect of not only up-and-down interest rate movement in a parallel fashion, but also changes to the shape of the yield curve. For instance, the yield curve gets very much steeper going from 1996 to 1998. You get a big steepness increase in the yield curve. But then going out further to 2002 and from there forward, the yield curve tends to flatten out. That's one of the benefits of a multifactor model over a one factor model -- you get those yield curve changes.

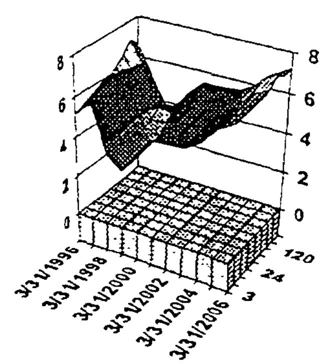
CHART 5

GAT Generator

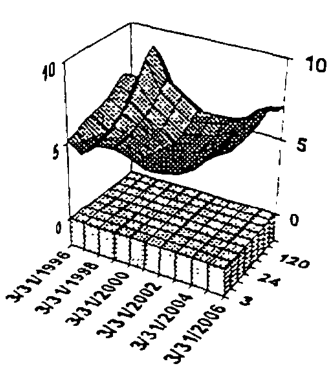
3/31/1996 2-Factor LPS Spot Rates
++++000,+-+-+-0



3/31/1996 2-Factor LPS Spot Rates
+0-+0+0,--0+0++



3/31/1996 2-Factor LPS Spot Rates
--++-++,0+--0++



3/31/1996 2-Factor LPS Spot Rates
-+-+-+-,+-+-+-+

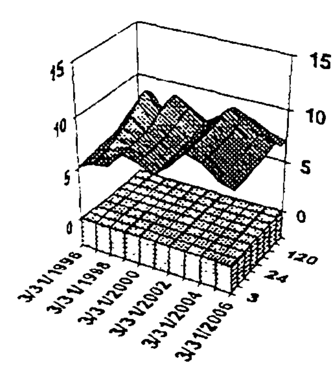
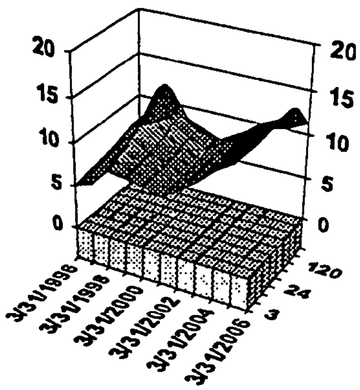


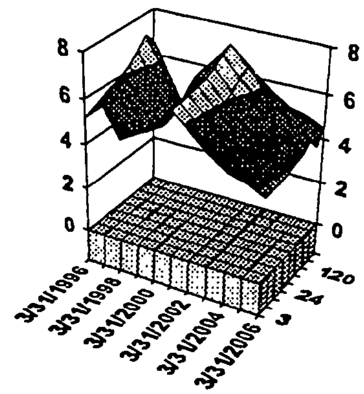
CHART 5 (continued)

GAT Generator

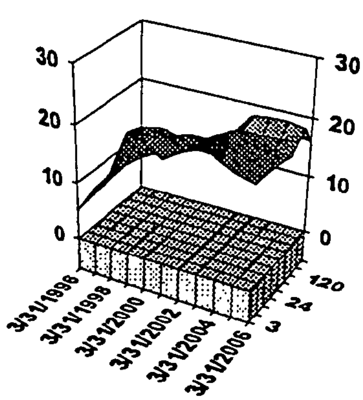
3/31/1996 2-Factor LPS Spot Rates
--++--0,++--++0



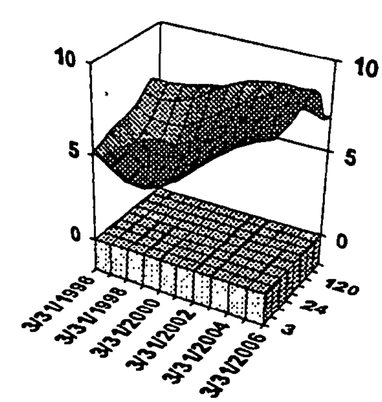
3/31/1996 2-Factor LPS Spot Rates
++--000,--++--0



3/31/1996 2-Factor LPS Spot Rates
++++000,+0+0+00



3/31/1996 2-Factor LPS Spot Rates
----000,0000000



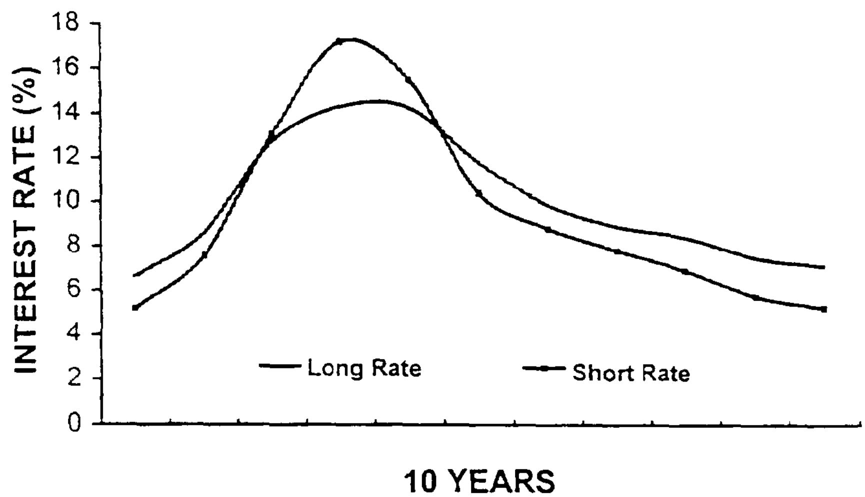
As I said, GAT is just finishing up refinement of this generator. One of the issues is that with the two factor model, you can produce over four million possible paths. No one has time to run four million paths, so GAT is going to have to find a way to limit the number of paths to use when testing.

For this experiment what I did was pick out some real extreme scenarios, some really volatile ones, to see what would happen to my model. I picked 35 paths, ran my model through and ended up passing 22, or 63%. If this were a standard set of 35 stochastic scenarios and I only passed 63%, I think I'd be pretty concerned, but I think the thing I've really proven here is that it is possible to create scenarios that you will fail. A lot of these scenarios had interest rates going up over 20%. A lot of them had inversions for a few years in a row. I'm not really worried about the fact that I only passed 22, but what I want to do in a situation like this is look at the ones that I failed and find out why.

Chart 6 shows that was a little bit of a concern. We failed this scenario. Obviously, it was volatile. What happened was interest rates went up, and my portfolio yield couldn't keep up with my credited rate, so we lost money on our spread. But the interesting part is when interest rates were on the way down. My model strategy didn't follow interest rates down very quickly. The strategy used a lag and moving average approach to following interest rates down, so I lost money when interest rates were coming down, as well, because I credited such a high rate.

CHART 6

GAT Scenario #20



So, if I were strategy testing here rather than cash-flow testing, and I want to try to find an optimal crediting strategy for my liabilities, I thought I'd test the scenario where I would bring down credited rates a lot quicker, which is probably what I actually would do in this type of situation. On the way down, I want to be a lot more aggressive with my interest rates. When I tested the more aggressive strategy, I actually passed the scenario.

Next, I ran the more aggressive strategy through my duration analysis. I was surprised to find out that the more aggressive strategy actually had a lower duration than my base line strategy (Table 5). This was somewhat of a surprise because you would think in general it's desirable to have higher durations on our liabilities. Usually, we want to implement strategies that produce higher durations on liabilities, so they match up better with our assets, especially under a rising interest rate scenario. So, why does this aggressive strategy give me a lower duration, but it actually looks like it performed better under my increasing scenario?

TABLE 5

Strategy Comparison

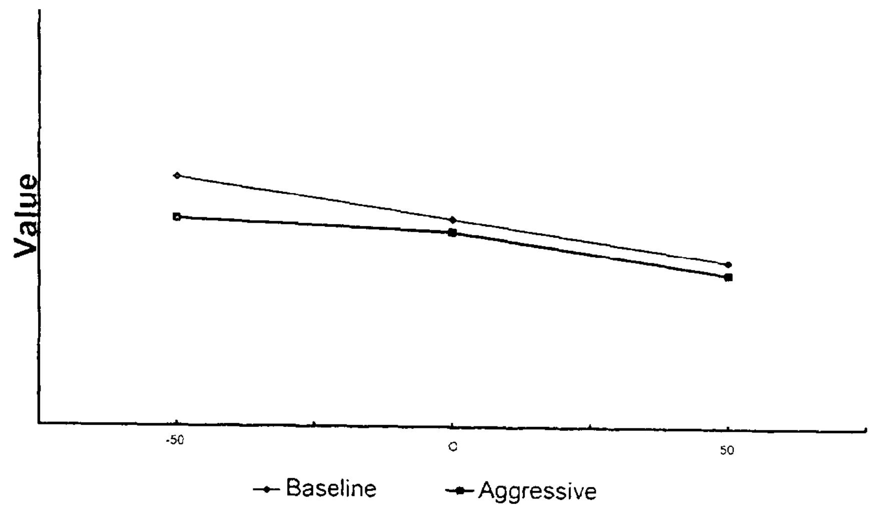
Strategy	Duration	Convexity
Base	1.36	0.03
Aggressive	1.13	-0.37

The answer is that the aggressive strategy is better. The aggressive strategy line on my price behavior curve is always lower than the baseline strategy (Chart 7). The aggressive strategy is actually less costly. The reason the duration goes down on my aggressive strategy is because of the kink in the curve under my aggressive strategy.

If I was doing duration analysis by itself, I might conclude that the baseline strategy is better because it has a higher duration, but in fact, the aggressive strategy is better.

CHART 7

Price Curve



Now, my numbers haven't really failed me. What I forgot to look at was convexity. The bend in the curve actually produces a negative convexity for my aggressive strategy, and that is a very desirable effect for your liabilities (Table 5).

The way you can think of this is that the benefit of the negative convexity is actually more important and tends to outweigh the detriment of the lower duration for the aggressive strategy. This again is just one more example of the need to dig deeper into the numbers. You can't just do duration analysis and stop with that. You really need to dig and see what's underneath, so you can really understand what's going on with your model before you make decisions.

My final model is what I'd call my brute force approach. This is a typical stochastic generator. It's a two factor model based on a lognormal distribution, with mean reversion and volatility inputs. We have a couple of different ones in our shop, but a typical stochastic generator is one I like to use when I do asset/liability analysis, because it takes away most of the constraints or a constraints that I get with my pricing models. You tend to get a lot more varied interest rate movement, a lot more random interest rate movement, and a wider range of coverage of future interest rate space.

Another one of the benefits is, I don't get the drift in my paths that the OAS models tend to produce. Because the OAS models are based on the current yield curve and the current yield curve is usually always upward sloping, OAS models and pricing models tend to produce paths that go up over time. Even the down scenarios can tend to increase over time because they need to follow the current yield curve or the forward rates that underlie the current yield curve. With more of a brute force approach, we don't get that drift.

Table 6 shows some results with my stochastic generator. I'm getting some convergence maybe even after 100 or so scenarios. I tend to think I got a little bit lucky. I think that it takes more than 100 scenarios with a Monte Carlo type generator to get good convergence, but in this particular example, maybe 100 or 150 produces pretty meaningful results. Many people don't have time to do 100, 200, or 300 different stochastic scenarios. I've seen a lot of people produce results with only 20 or 40 scenarios. I think this can be meaningful to see how you perform under those specific scenarios, but I don't think you really can rely on the percentage that you pass as being a credible indicator with only 20 or 40 scenarios. What you can see is what happens to your block under those scenarios or to your company under those specific scenarios that do come out, but I'm not sure you can really rely on the pass ratio and say you have something that you can be confident with.

In any case what I can do is look at the ones that I failed. I want to see why I'm failing. When I looked at those 300 scenarios and the 28 that I failed, once again, we failed when interest rates were increasing -- some with rates up to 20% or more. We failed with inversions in the yield curve, sometimes over a longer time period. Another area where I failed was when yield curves flattened out and stayed relatively flat for a few years in a row. That was a little bit of a concern to me. When I looked a little bit deeper I saw that my investment strategy was investing a good portion of my cash in longer-term bonds -- five-year, seven-year bonds. I'm not sure I'd really follow that under a level yield curve. I think I would change my investment strategy.

TABLE 6

Stochastic Generator

Total Scenarios	Cumulative Passing	Ratio
50	42	84%
100	92	92%
150	133	89%
200	178	89%
250	224	90%
300	272	91%

Our current models and our current state-of-the-art in commercial software tends to lock us into a given set of management strategies. I'm not sure we would really follow those strategies under all the different economic conditions that we want to test. I think we could act very differently as yield curves change in terms of our investment, or disinvestment, even our crediting. We always need to remember that we do have model limitations. I've always found that when I want to do real asset/ liability analysis, in practice, I've always found that I modify the commercial software systems to give me more dynamic management action to go along with dynamic interest rate movement.

Nevertheless, given that I've passed 91% and what I've found about my model and the scenarios that I'm failing, I'm pretty comfortable with my pass ratio in terms of what it says about my asset adequacy.

That leaves us with the question of, where do we go from here? I think there's a lot more research to do in interest rate generators. I think there is a movement towards multifactor models from one factor models as people realize you need to capture more interest rate risk. I think there's a lot more work to do on the equilibrium side of the house rather than the OAS models, so try to put more effort into using historical data and statistics to drive interest rate generators. I also think there's a need to incorporate term premium into our interest rate generators, which most of them don't do right now.

The term premium that is represented is because the yield curve is upward sloping because people demand a higher return for a longer investment, rather than because everybody thinks interest rates are always going to go up. That's what a lot of our pricing models are really saying. When they based interest rate movement on the current yield curve and the forward rates that underlie it, what they're implying is that there's a general belief that interest rates will rise and the current yield curve has been proven to not be a very good predictor of future interest rate movements.

In the meantime that leaves us with the question, what is a good interest rate model? I conclude that it really depends on your intended use and what you're trying to get out of it. In my work, I've never found a perfect interest rate model. I've found ones that are good to use for a given purpose, but almost always have shortcomings in some other area.

In the end I think our interest rate models are tools for making decisions and selecting the most appropriate tool is more important than selecting or making the most sophisticated model.

MR. PETER FITTON: I am here to provide the theoretician's perspective on interest rate modeling. My background is in making interest rate modeling software; creating a superior interest rate model is a task to which I have devoted the last two years of my life.

What I am here to tell you, in a nutshell, is that a little theoretical understanding is a key prerequisite to success. In order to use stochastic interest rate models to accomplish "stochastic cash-flow testing" (as a supplement to cash-flow testing using the seven deterministic scenarios), there are a few concepts that you must understand.

The most basic concept is, of course, what is an interest rate model. I will introduce the framework of interest rate modeling, and explain how the interest rate model changes according to your intended use. The two main intended uses are generation of realistic scenarios for cash-flow testing -- something similar to testing the seven scenarios, where you're trying to generate a probability distribution of what might happen in the future -- and generation of risk-neutral scenarios for pricing or duration calculations. The same model used for pricing is also appropriate for duration because calculating duration simply involves determining two prices with different initial yield curves. The form of the interest rate model is going to be different for each of the two intended uses.

I am also going to talk abut the evolution of interest rate models -- where things have been, where things are going. Then I will talk a little bit about how cost considerations come into play, because the decision to use a good interest rate model is really a decision about building or purchasing software. I will describe the current state of affairs in interest rate modeling software and what that means to a valuation actuary.

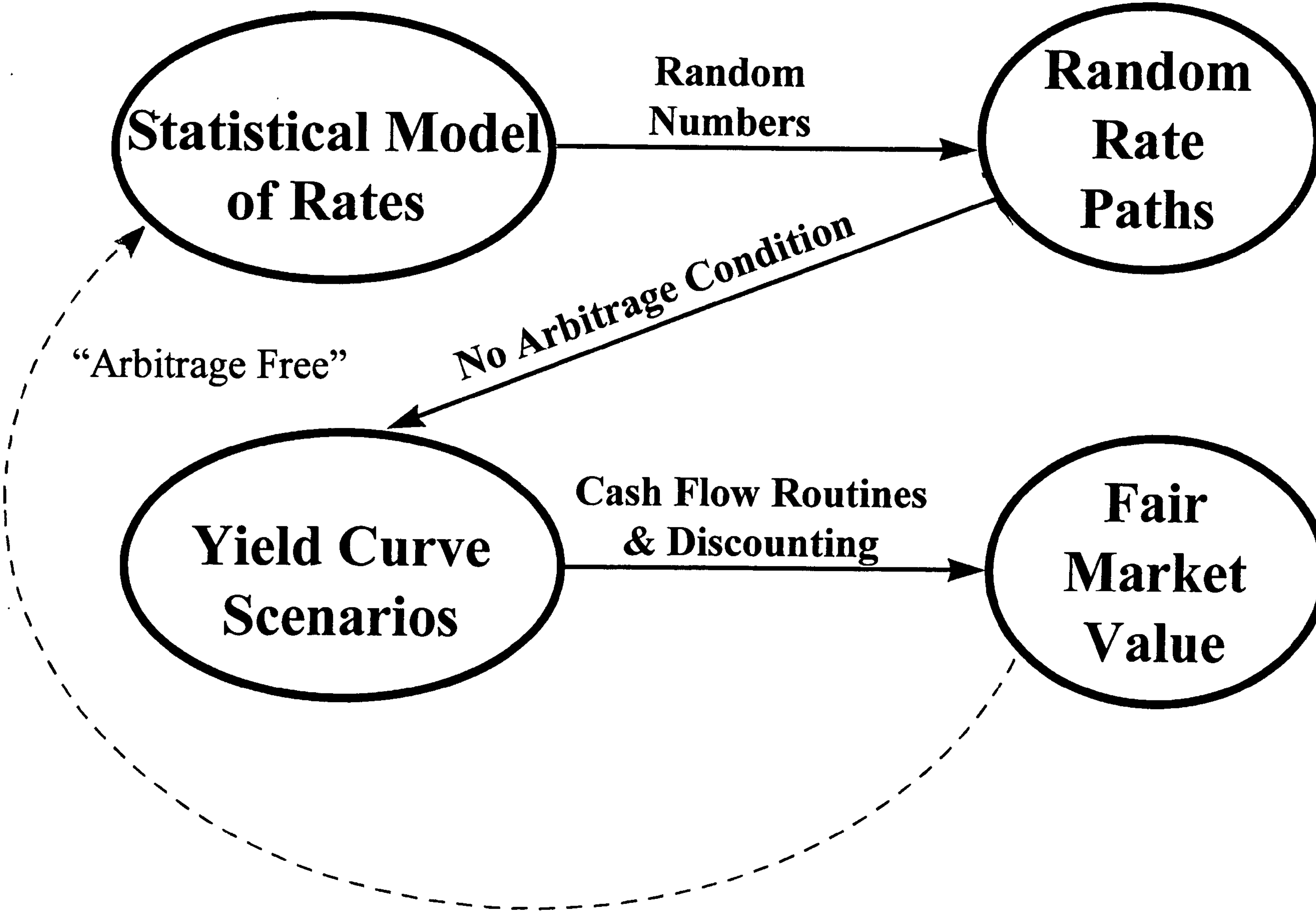
An interest rate model is really a statistical model of interest rates (Chart 8). That's the most important thing to understand. Whether it's realistic or whether it's risk-neutral, whether it's arbitrage-free or not, it's still intended to be a statistical model of future interest rate behavior. Suppose we think of an interest rate model as a tool for generating random yield curve scenarios. We have a statistical model, implying some distribution for future interest rates. We can sample from this distribution using random numbers, resulting in random rate paths for a small number of maturities. If we are only generating the short-term interest rate randomly, and every other rate is a deterministic function of the short-term rate, then we have a one factor model. If we are generating more points on the yield curve randomly (or variables that can be identified with longer-term rates), then we have a multifactor model.

We proceed from the generation stage to fill in the rest of the yield curve around the randomly generated rates by using no-arbitrage conditions. What I mean is that our interest rate model, by providing the rate paths and a probability distribution of future rates at each point along these paths, along with a term premium, implies what the yield curve shape should be to preclude arbitrage.

Then, when we have the yield curve scenarios, we can go on and feed them into asset/liability models, produce cash flows and discount them back to today to get an estimate of fair market value. The quality of that estimate can be judged in comparison to market value, or what market value would be for something tradable. Of course, assets mostly are, and liabilities are not.

CHART 8

Interest Rate Model



The quality of the valuation is going to depend on the quality, or realism, of the interest rate model as well as the success of the asset/liability model in reflecting the true behavior of the asset or liability. It depends on the approximations made in the asset/liability model.

The concept of arbitrage-free comes in when we take a small set of assets and compare the fair market value implied by the model to the set of asset values observed in the market. Usually, we do this for a small set of bonds and adjust the original statistical model until the two sets of prices match exactly. That does not indicate, of course, that for any other asset you will necessarily have an exact match between the fair market value generated by the model and the market price. It means that for this small set of assets, you have a match.

If your purpose in using an interest rate model is valuation or calculating something like duration, then your requirement of the model is to produce the right expected discounted value of future cash flows. In that case, the model form is less important. An important thing to know about risk neutral valuation is that the only aspect of the distribution that has any relevance is the mean; that is, the mean of the distribution of discounted cash flows. The risk-neutral framework, and I'll get into this in more detail in a moment, makes two assumptions that counterbalance each other, the effect of which is that the mean of the distribution of discounted cash flows can be identified as the price, but the volatility and higher moments of that distribution have no economic relevance.

On the other hand, when it comes to dynamic financial analysis -- stress testing and cash-flow testing -- what is required is a realistic representation of what the future might hold. What is required is a proper statistical distribution of the future that reflects all the higher moments, particularly the second moment, since it is usually identified with risk.

I want to talk abut the difference between the risk-neutral and realistic forms of an interest rate model (Chart 9). They are separated by the term premium. In the realistic form of the model, which comes from what we call the risk-unadjusted process, we have a statistical model of interest rates that implies how interest rates are going to move in the future. We hope this reflects some reasonable expectations of what the future might hold. If we generate random paths, they are centered around that realistic expectation. On the other hand, in the risk-neutral model the paths are centered around the current forward rate curve (implied from the current yield curve). Forward rates, as Doug mentioned, are not good predictors for future interest rates, because they are separated from interest rate expectations by an expected return premium for interest rate risk.

If you invest in any security for more than a day, you bear some interest rate risk. Since investors require an additional return for bearing risk, a term premium is built into the initial yield curve. If you use that initial yield curve as a predictor of rates, you are going to go wrong.

This is why I said that the expected value of discounted cash flows is the only part of the discounted cash-flow distribution that's relevant for valuation. What we do in the risk-neutral method is to make two incorrect assumptions at the same time. We can assume on the one hand that forward rates are unbiased predictors of future spot interest rates, knowing that that is wrong. We can also assume that all investors are risk neutral; that is, they do not demand a term premium for interest rate risk. If we make those two assumptions at the same time, then we manage to get a good value out of the model. We get the same fair value that we would have if we assumed a realistic rate distribution and then discounted properly, taking into account not only the short-term interest rate, but also the premium for interest rate risk. However, when we use risk-neutral analysis we change the shape of the distribution of rates. We will not get any information about risk from the distribution of risk-neutral prices.

Chart 10 shows a history of Treasury rates -- the one-year and ten-year rates over a period of something like 30 years. We have only one path. This one realization is the only thing we have to make a judgment about the realism or lack thereof of an interest rate model. What that means is that the judgments we can make about realism are pretty limited. There is a certain degree of complexity that we can put into the model that may or may not be justified by the market. The observed path, being only one realization out of some large probability distribution, does not give us much information about anything except the most simple models. On the other hand, whatever information is embedded in that one history, we want to use it. Thus, we want to choose a simple model, but a simple model that captures some basic characteristics of realism, for stochastic cash-flow testing.

CHART 9

Risk Neutral or Realistic Distribution

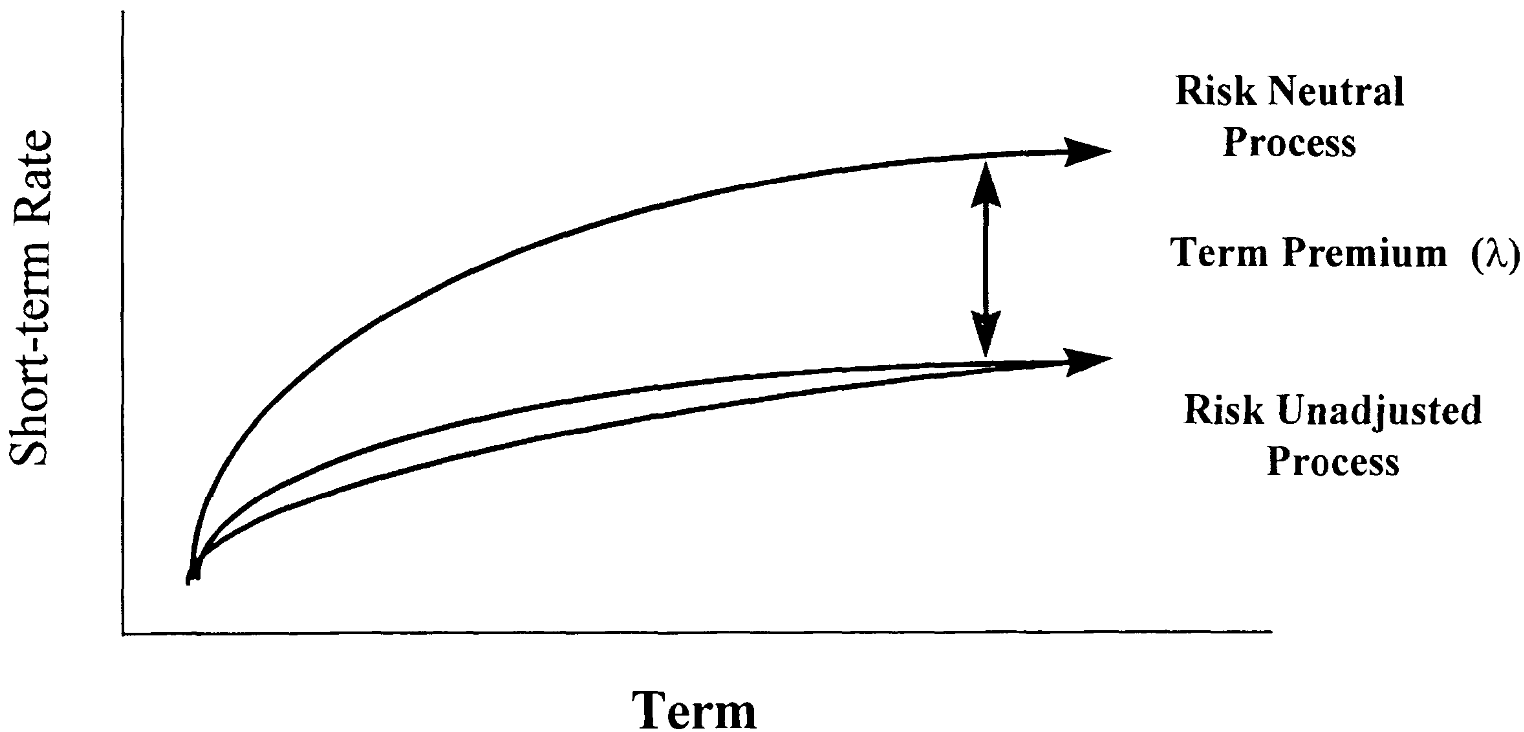
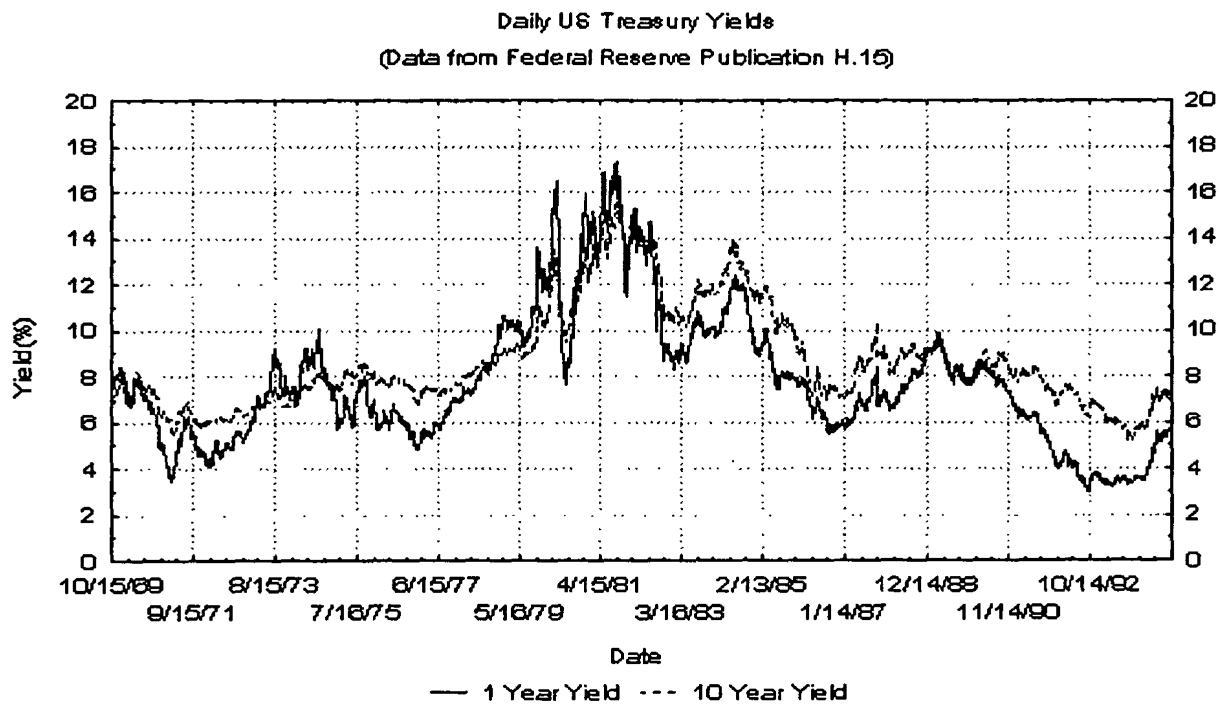


CHART 10

Evaluation Criteria

Realism



Our two design considerations for the model are:

1. How does a specific spot rate or yield (whatever equivalent you want to look at) move over time within the model (Chart 11)? We can look at history and collect some statistics about various rates: mean, volatility, correlation, and so on. We know that nominal rates have never been negative, and we have good economic reasons to believe that they never will be. We know that nominal rates have never been particularly high, the highest having been in the early 1980s at a level lower than 20%. We have good economic and political reasons to suspect that the Federal Reserve is going to do everything it can to keep that from happening again, and that it is likely (though not certain) to be successful in doing so. There are a few other things we can examine. We can look at historical distributions and try to do distribution fits. This is somewhat inconclusive if you try it, but it is another piece of information that forms an input to the model.

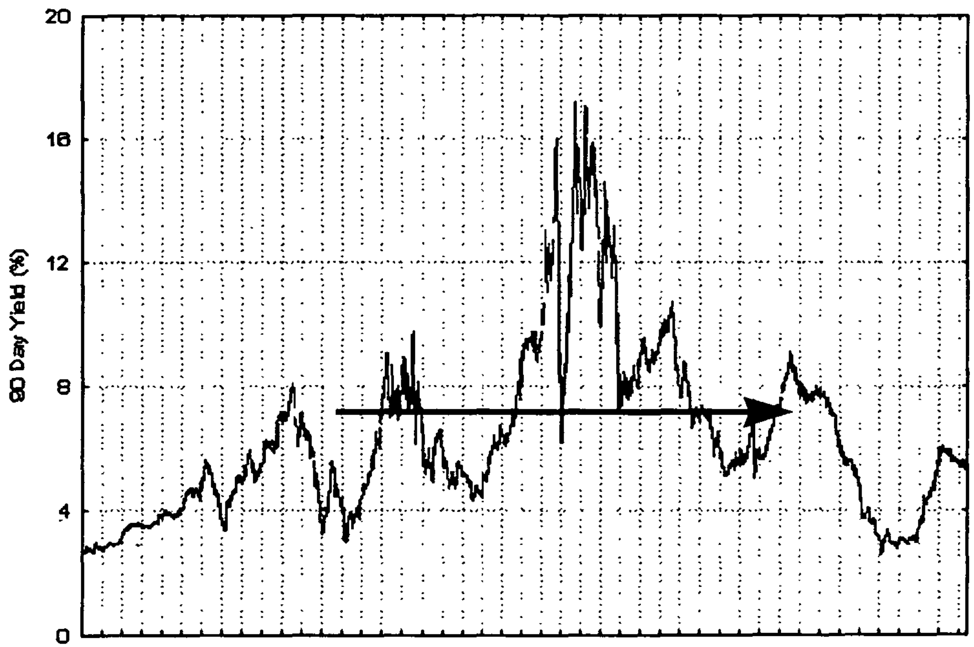
2. How, given random generation of one or more points on the yield curve, do we get the rest of the curve (Chart 12)? The real problem with using a one factor model for generation of rates is the model generates only one possible yield curve for a given short rate. At any particular point in time, the short-term rate that the model generates is a perfect predictor of the rest of the curve. If any rate rises, all rise, and vice versa. That sort of perfect correlation is simply not observed in the market. There are high positive correlations among rates at different points on the curve, but certainly not perfect.

Historically, if you look at the 30-year Treasury yield versus the three-month Treasury yield, you observe that the correlation over the last 40 years is about 85%. This is positive, but not perfect, reflecting the fact that we can have long rates and short rates move in different directions. In the past couple of years, we've seen long rates go down and short rates go up at the same time; this cannot happen in a one factor model. We have to be prepared to test the effect of such a yield curve movement on cash flows, so we have to include that sort of behavior in the model.

CHART 11

Alternative Models

Design Consideration 1: How does a specific spot rate move over time?

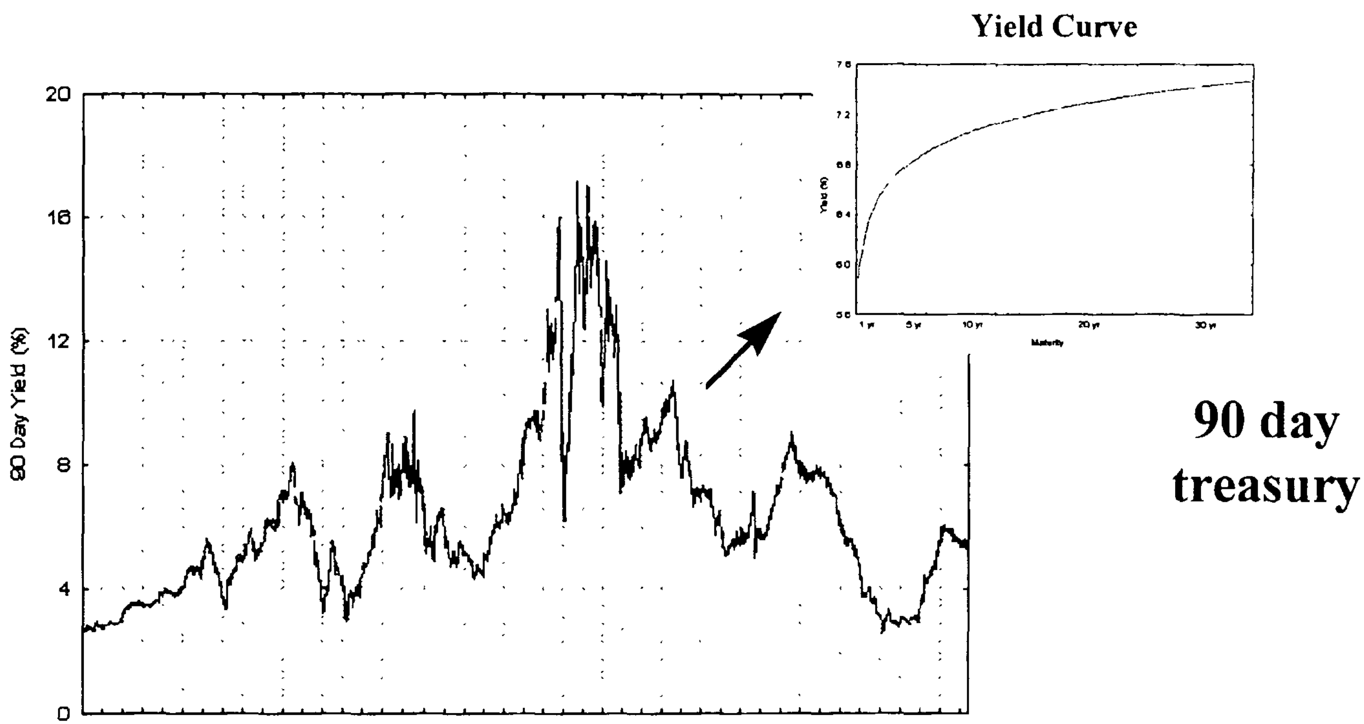


90 day treasury

CHART 12

Alternative Models

Design Consideration 2: Given a specific path, what is the yield curve at any point in time (to remain arbitrage free)?



This is why, when we talk abut the evolution of interest rate models, the starting point is simple one factor models. One random factor drives the yield curve process. The entire yield curve depends on one random variable. For these models, having mean reversion seems to be better than not having it; that is, the behavior matches historical distributions much better if our one factor model contains some tendency for rates to come back to a long-term average.

It is very hard for a one factor model to meet even some basic realism standards derived from our one historical path of rates, because of the perfect correlations, and because it is very difficult to control the relative volatilities of different yields on the yield curve with a one factor model. We end up understating interest rate risk with a one factor model. When we generate scenarios with a one factor model, we end up with a distribution of values that is not going to reflect any risks due to yield curve shape changes of the types that don't appear in the model. We get unrealistically high correlations between points on the curve, and unrealistically low volatilities of longer-term rates.

One factor models have typically been used on Wall Street because they're not bad for pricing the simpler interest rate derivatives. They're not bad for pricing because of what I said before: it's only the mean that really matters. Having some other elements of the distribution of rates that do not match historical behavior is not so much of a problem for pricing. Thus, one factor models are typically used only in the risk-neutral form. As Doug pointed out, that is also a problem if we want to examine what the future might hold realistically.

The next step in the evolution of interest rate models has been ad hoc multifactor models. By that, I mean they're intended to be statistical models of realistic interest rate behavior, but the designers haven't paid attention to the second design consideration, where we talked about how the path of rates relates to the yield curve appear at a particular point along that path. Typically, these ad hoc models have not derived their yield curves on the basis of no-arbitrage conditions, but instead have produced yield curves by some sort of interpolation, drawing a line between the randomly generated points on the curve.

If you're generating a three-month rate and a ten-year rate, then you just draw a line straight through them, and that's a yield curve, or you have some sort of prespecified curve fitting process that doesn't reflect no-arbitrage conditions. The result is that these models typically can't be made arbitrage-free, and if they were, they would lose their realism. However, they do give a more realistic picture of interest rate risk than a one factor model. Even though they're not strictly realistic in terms of maintaining no-arbitrage conditions throughout the future, they at least give you some randomness in yield curve shape changes and can match some elements of the distribution of interest rates better.

The next step in interest rate model evolution, and really the state of the art right now, is a rigorous two factor model -- one that's structured so that it can be made arbitrage-free. It can match a certain subset of market prices or yields at the beginning of scenario generation, but it has been econometrically designed so that it matches realism standards at the same time. This will give you the most realistic picture of interest rate risk that's available to date. To my knowledge, nobody has successfully implemented a three factor model with these characteristics for these purposes. Certain Wall Street firms have experimented with three and more factor models of interest rates, to my understanding entirely unsuccessfully in terms of accomplishing the dual goals of meeting realism standards and maintaining arbitrage-free capability.

Whenever we consider moving from what's done now to a newer technology, there is the cost consideration. What kind of cost burden does moving from seven scenario testing to stochastic testing put on the shareholders of the company? Of course, historically, stochastic testing has taken longer: you have to run more scenarios, and it costs money to build or buy stochastic interest rate models. However, at the moment, software prices for stochastic interest rate models have come down substantially. At the same time, processors and mathematical methods have improved to the point that it is quite possible to run stochastic scenarios through a complex model of assets and liabilities on a network of personal computers, with current asset/liability modeling software, and get meaningful results. It's now possible to afford an interest rate model that has stochastic characteristics, meets some realism standards, and requires a small enough number of scenarios that current processing power can handle it. At this point, we've reached the stage where, for the majority of large insurance companies, it's obvious that they're capable of stochastic cash-flow testing. In my opinion, it has become a matter of professional responsibility to at least examine whether stochastic testing is a viable option, and whether you're going to learn something from it about the risks facing your company.

In summary, right now the state of the art would be a rigorous two factor model that meets realism standards in comparison to history. It can be made arbitrage-free. I'd also like to point out that both realistic and risk-neutral forms of the model are important, and it's even more important to understand the difference between the two. When we're trying to project possible futures, we cannot use a risk-neutral model, since that model has deliberately been made unrealistic by embedding the term premium into expectations of future rates.

Finally, better models are becoming much more feasible and affordable for the valuation actuary, and so it's much more possible for many companies to adopt stochastic cash-flow testing this year and next.

MR. DAVID N. BECKER: Before the formal part of my presentation I will begin with a discussion of what our mission as appointed actuaries and valuation actuaries should be.

What is it that we as actuaries do? What's the perspective from which we operate? In all fairness, you shouldn't think of yourself as the "vice president of formula reserves and regulatory reports." More important, you don't want your company thinking of you this way. If you do have that as your mission statement, then I don't think you're going to be very happy over the long haul.

What our job should be is risk management and determination of the value of our firm and our firm's lines of business. We do these by identifying and quantifying risk and by evaluating strategies to optimize the risk/reward posture of the company. To do this you must have the proper science. And that is what we bring to the management of the business: science! It is our responsibility to extract information from data and to gain knowledge from information so that we can make decisions based on knowledge.

The applications of our science fall into two areas: risk management and valuation. I use the term *valuation* to refer to the process of determining the fair value of a block of business in the sense of what a buyer would pay a seller. What's it really worth? Under risk management, there are several specific applications. These include liquidity analysis; reserve and asset adequacy analyses; interest rate risk (C-3) factor determinations; and determination of value at risk. Under both of these topics, there is investigation of the impacts of alternative management strategies (investment, disinvestment, liability management, crediting rate management) on either the risk posture or the valuation of the block of business or entire company.

I will make some sidebar comments with regard to liquidity analysis. What I mean by a liquidity analysis is the "safe-harbor" method in New York Regulation 126, i.e., the computation of the excess (we hope) of the market value of assets less the market value of liabilities. Here the market value of liabilities is taken to be the cash-surrender value for surrenderable liabilities and the present value of future obligation and related expense cash flows for other products. This is really a liquidation test, not a reserve and asset adequacy analysis. I don't see this as being equivalent to reserve and asset adequacy, and so I split that out as a separate risk management analysis. An appointed actuary may choose the liquidity analysis as his or her definition of reserve and asset adequacy. That is up to the individual appointed actuary as the regulation allows the appointed actuary to adopt his or her own basis for analysis.

I do not believe that a reserve and asset adequacy analysis is equivalent to a liquidity analysis. Furthermore, I have concerns about the credibility that should be placed on a liquidity analysis. For example, how much credibility can one place in the assessment of the market value of assets ten years out in a projection? What really are the market values of real estate, commercial mortgages, and equities? Even for bonds, will the spreads bonds be the same in ten years' time? If you look at bond spreads over the last 20 years, you'll see that they have varied on intermediate term bonds from 40 or 50 basis points up to 200 basis points. Projecting for ten or 20 years and computing market values results in a number; but how credible is that number? This is not to say that we shouldn't consider liquidity analyses; but one must carefully adjudicate the results in arriving at conclusions. There is a dangerous tendency to simply follow the path set by New York Regulation 126 and not apply the critical light of reason. If we as a profession do this, then we are just "vice presidents of formula reserves and regulatory reports."

I want to follow up on the issue of deterministic versus stochastic scenarios. If one adopts the New York seven approach to reserve an asset adequacy, you're basically performing a deterministic analysis. If you invent 20 scenarios by designing specific interest rate paths, you're still following a deterministic approach.

What is good about deterministic scenarios is that you get to design them. You get to specify exactly how interest rates are going to evolve. Then you perform a very extensive path specific analysis regarding exactly what happened and why. What turned out well and what turned out badly. For stress testing, this approach is excellent.

But if your mission is to make a robust decision about how to manage your business or what your risk posture really is, what does a single scenario mean? How do you use it or even a collection of deterministic scenarios to form a basis for action? With stochastic scenarios, you can manage the business based on the entire probability distribution of results; but you'd better know your interest rate model and that gets us back to the issue of science.

The recommendation I would make is that stochastic models are required for robust risk management and valuation; but you have to choose a model that generates credible and diverse scenarios, and you must assign parameters to it correctly. You must avoid arbitrage in the scenarios. This doesn't mean that the scenarios are "arbitrage-free," but that, along each interest rate path, the yield curves do not admit arbitrage, i.e., no negative spot rates or negative implied forward rates. You must use an adequate number of paths, and you must match the interest rate model to the application.

There are three characteristics of interest rate models: model factors, model architecture, and model calibration. The factors of an interest rate model are the random variables in the model. One might characterize models in this way as one factor or two factor. Typically, the factors represent the short rate, the long rate, the volatility of the short rate or the long run mean reversion target. As Peter discussed, one needs sufficiently many factors to capture the richness in movement of the yield curve over time, but needs to avoid too many factors that would allow implausible yield curves.

Model architecture refers to whether the interest rate model is designed to follow risk-neutral valuation principles or is designed to produce realistic interest rates. Risk-neutral valuation is the basis for computation in option pricing models. It provides a special frame of reference in which it is straightforward to compute the price of a security. This price is the probability weighted present value of the security's cash flows. Risk neutrality can be achieved by either generating equiprobable risk-neutral interest rate paths or by generating realistic interest rate paths with special risk-neutral probabilities. But risk-neutral interest rate paths typically result in paths whose interest rates deviate significantly from historical interest rates. Deviations may be seen, for example, in the absolute level of rates, the frequency in which rates are abnormally high or low, the frequency of significant rate increases or decreases, and in the frequency and durations of inversions.

Realistic rates are created from paths where the interest rate process is designed to track certain global characteristics of historic interest rates. Paths generated in this way are appropriate for simulating the performance of a block or a firm over time.

Model calibration refers to whether the interest rate model is calibrated to reproduce a specific set of security prices on a given day ("arbitrage-free") or is calibrated to reflect the characteristics of historic interest rate movements (equilibrium). The first approach is used in option pricing and the pricing of fixed income securities with embedded options where relative pricing on a consistent basis is needed. But the price of every security in the underlying set imposes a condition that translates into the value of a parameter in the interest rate model. The resulting parameterization of the model is sensitive to the security prices of that set. As market values change, the parameterization also changes. Such changes are typically volatile as markets fluctuate with bid/ask spreads widening or narrowing, volume of trading activity in various portions of the term structure and macroeconomic forces. The parameterizations of arbitrage-free models are unstable over time.

Equilibrium models attempt to capture the underlying dynamics of the interest rate process over time. They involve relatively few parameters. As a result, they will not price securities precisely at any given moment, but will provide reasonably accurate values and provide such reasonably accurate values over an extended period of time.

It is important to match the application with the interest rate model. If you're performing risk management exercises, then you should use a realistic interest rate path generator and equilibrium pricing because you're projecting over a long period of time and you are attempting to simulate experience in the real world. In the real world, there are term premia for investing longer.

If you're performing valuation exercises -- for example, the price of a block of business -- you should use a risk neutral model, and you could choose to use either an arbitrage-free or equilibrium calibration. Again, if you choose an arbitrage-free model, you may be overspecifying the situation and giving too much weight to a specific set of security prices today; whereas if you use a risk-neutral and equilibrium model, you would achieve more balance over the period of the financial projection.

The example I will use is essentially the same example used by Doug. It is a $100 million block of SPDA business. These are generic SPDAs with a low guaranteed minimum crediting rate and surrender at book value less a surrender charge that grades to zero over seven years. The block represents three years of issues. The supporting assets consist of noncallable bonds, simple collateralized mortgage obligations (CMOs) and mortgage pass-throughs. The study is based on the interest rate environment as of the end of March 1996.

Whereas Doug's illustration focused on risk management, my example will focus on valuation, i.e., the fair value of the SPDA block to the insurance company. This fair value is measured by the shareholder dividends that could be paid. This quantity is called distributable earnings. Distributable earnings will be discounted at 700 basis points over the one-year risk-free rate. These 700 basis points are the assessment for risks other than the interest rate risk in the process, and the risk that the interest rate process is misspecified.

Both one factor and two factor interest rate models will be considered. The one factor model is the Tilley model that appeared in the *Transactions* a few years ago. This model is risk-neutral. It will be implemented in both an arbitrage-free and not arbitrage-free formats. The two factor interest rate model is the Cope 2000 interest rate model. It is based on the double mean reverting process developed by Mark Tenney. It will be implemented in two formats: risk-neutral/equilibrium and realistic/equilibrium.

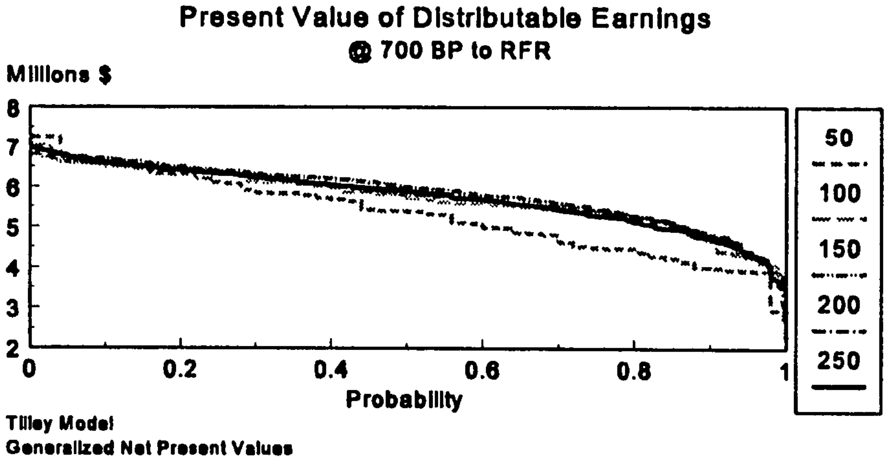
One must be cautious when comparing two or more interest rate models. The risk can be illustrated using a one factor model. Consider a one factor risk-neutral and arbitrage-free interest rate model. Suppose it assumes a 15% volatility assumption for the short rate and has been calibrated to a given set of security prices. Alternatively, set the volatility to 30% for the short rate, and recalibrate the model to the same set of security prices. If the two interest rate models are applied to securities outside the set with such securities having embedded options, then the prices of those securities outside the set will likely differ. Due to the significant difference in volatility inputs, they may differ by a large margin. The danger in comparing two models is that, if care isn't taken to parameterize them reasonably consistently, then the differences that you see might be due to the mis-parameterization, not to differences in the fundamental characteristics of the two models.

For these examples the several interest rate models were calibrated so that all of them reasonably reproduced the prices of a set of interest rate caps. Therefore, the results we get when we apply them to our annuity block will be attributable to real differences between one factor/two factor, risk-neutral/realistic or arbitrage-free/not arbitrage-free.

Chart 13 displays the cumulative distribution function of the present value of distributable earnings for the one factor, not arbitrage-free model. Five sets of paths were tested with 50, 100, 150, 200, and 250 paths, respectively. It happens in this model that the 50 path distribution is quite a bit different from the distributions for larger numbers of paths; but when at 100 paths or higher, the plots of the distributions tend to be very close together. Visually, the distributions tend to converge to the 250-path distribution.

CHART 13

One Factor/Not Arbitrage-Free



One item to note is that the behavior of the distributions in the tails, i.e., the initial 10% and the last 10%, seem relatively stable for this interest rate model once the number of paths exceeds 50.

I performed statistical tests on this interest rate model and found that for three different tests the 50 path distribution and the 100 and higher path distributions are different. The significance level chosen a priori for all of the tests is alpha = 5%. The first test is the Kolmogorov-Smirnov test, which tests if two distributions are identical. This test is rejected for the 50 path distribution and the higher number of path distributions. However, the test isn't rejected at the 100 path and higher paths. The second test was for the equality of the means of the distributions, and the results are similar. For the 50 path distribution and any of the higher number of paths distributions, the hypothesis that the means are the same can be rejected, but not so for the distributions of 100 or more paths. The third test was a nonparametric test for the equality of the medians of two distributions, the Mann-Whitney U test. Results are similar to the prior two tests. The medians for the 50 path distribution and distributions for 100 or more paths are not the same; but there is not statistical difference for the 100 path and higher distributions.

Chart 14 displays the cumulative distributions for the present values of distributable earnings for the one factor, risk-neutral arbitrage-free model. It may just be coincidence, but this particular sample of 50, 100, 150, 200, and 250 paths did not exhibit the same 50 path versus over 50 path behavior on the statistical tests that the one factor, risk-neutral not arbitrage-free model did. In this case the tests did not disclose any statistically significant difference between the distributions for the various numbers of paths. Examining the 250 path distribution, it is seen that both the tails are relatively smooth.

CHART 14

One Factor/Arbitrage Free

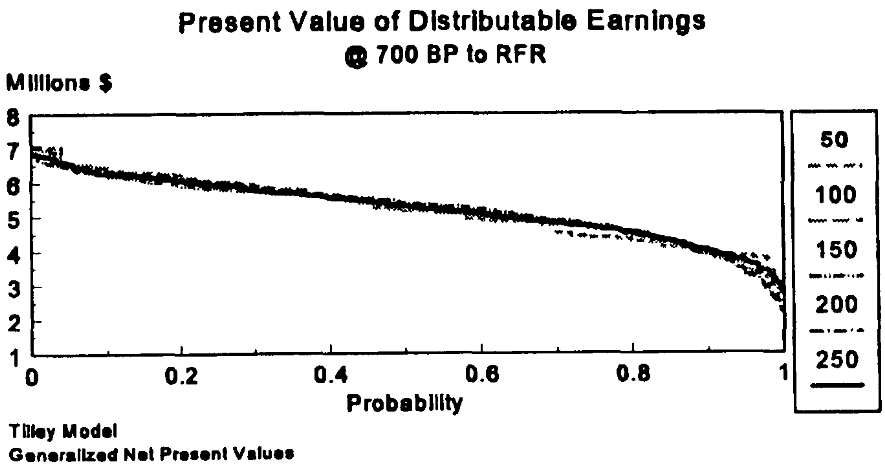


Chart 15 displays the cumulative distributions for the 50, 100, 150, 200, and 250 path samples from the two factor, realistic equilibrium model. Chart 16 displays similar information for the two factor, risk-neutral equilibrium model. For each of these two models all the cumulative distributions are relatively close. The statistical tests fail to differentiate between the distributions for the various numbers of paths. Unlike the one factor models, however, for both of the two factor models the right-hand tail is erratic even for the 250 path distribution. The 250 path distribution is relatively smooth until at approximately the 95th percentile. For smaller numbers of paths the distributions become erratic at lower percentiles, respectively.

CHART 15

Two Factor/Realistic

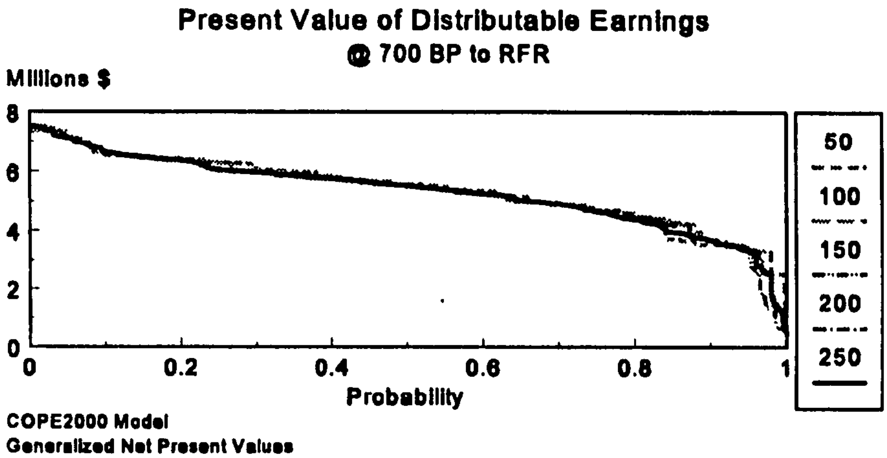
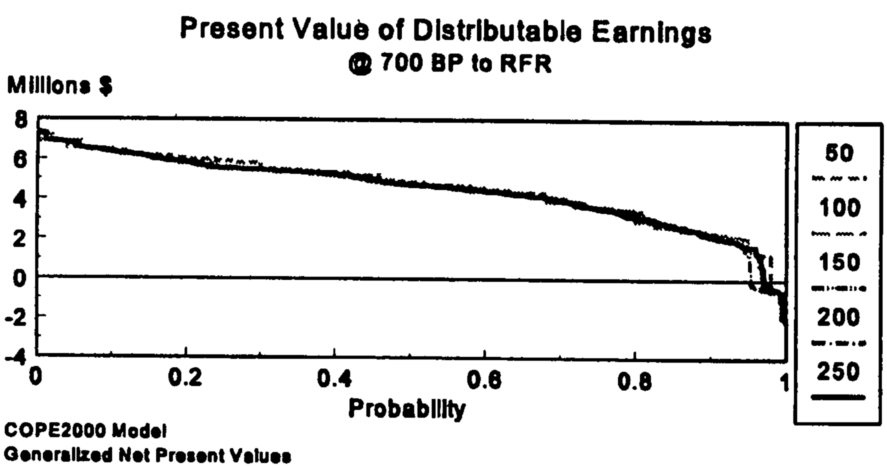


CHART 16

Two Factor/Risk Neutral



Because of their extra degree of freedom, which translates into more potential shapes of the yield curve over time, both two factor models uncover more interest rate sensitivity than the one factor models uncover. The fact that the extreme right hand tail is not stable indicates that, for a very detailed analysis of behavior in the tail, a large number of paths would be required. But if our concern is not a detailed analysis of the tail, then 250 paths provides credible information.

Table 7 presents the percentile values of the distributions of present value of distributable earnings for a large choice of percentages for the 50, 100, 150, 200, and 250 path results for the two factor, risk-neutral equilibrium model. This provides insight about the shape of the entire distribution and not simply the mean or median. Examining the table reveals that if the 1st, 5th, 95th, and 99th percentiles are ignored, then the remaining percentile values are stable at 150 and higher numbers of paths. In fact, the 1st and the 5th percentile values are very close; it is only at the 95th and 99th percentiles (the right hand tail) where there is significant instability. This provides quantitative support for the same conclusions stated earlier based solely on the visual inspection of the plots of the distributions.

I noted earlier that my presentation would focus on valuation or the fair value of the block of business, not risk management. For valuation purposes the recommendation is that risk-neutral paths be used. Chart 17 displays the cumulative distribution functions for each of the four interest rate models using 250 paths. All four models have similar results in the low percentiles, where results are the highest. At the 50th percentile, there is some divergence in results. But as the percentile values increase, i.e., as results get worse, the difference in the distribution functions of the four interest rate models becomes very marked.

In a valuation exercise the choice of one or two factors and risk-neutral or realistic make a significant difference in the results. Clearly, the two factor models uncover more sensitivity to interest rate risk in the block of liabilities and supporting assets. The difference between the two factor/realistic and two factor/risk-neutral demonstrates the differences that can arise from the choice of the realistic versus the risk-neutral decision.

TABLE 7

Risk-Neutral Percentile Rankings of Distributable Earnings by Paths

Percentile	50	100	150	200	250
1st	6.9	7.1	7.2	7.1	7.0
5th	6.5	6.7	6.7	6.8	6.7
8th	6.4	6.4	6.4	6.5	6.5
10th	6.3	6.3	6.3	6.4	6.4
20th	6.0	5.9	5.8	5.8	5.8
30th	5.5	5.5	5.4	5.4	5.5
40th	5.3	5.2	5.2	5.2	5.2
50th	4.8	4.7	4.7	4.7	4.7
60th	4.5	4.4	4.4	4.4	4.4
70th	3.8	4.0	3.9	4.0	3.9
80th	2.8	3.2	3.3	3.1	3.1
90th	2.2	2.3	2.3	2.1	2.1
92nd	1.9	2.2	2.1	2.0	2.0
95th	1.3	1.5	-0.2	1.5	1.7
99th	-1.9	-1.2	-1.9	-1.2	-0.6

Table 8 displays the percentile results for the four interest rate models for the 250-path cumulative distributions.

CHART 17

Interest Rate Model Results

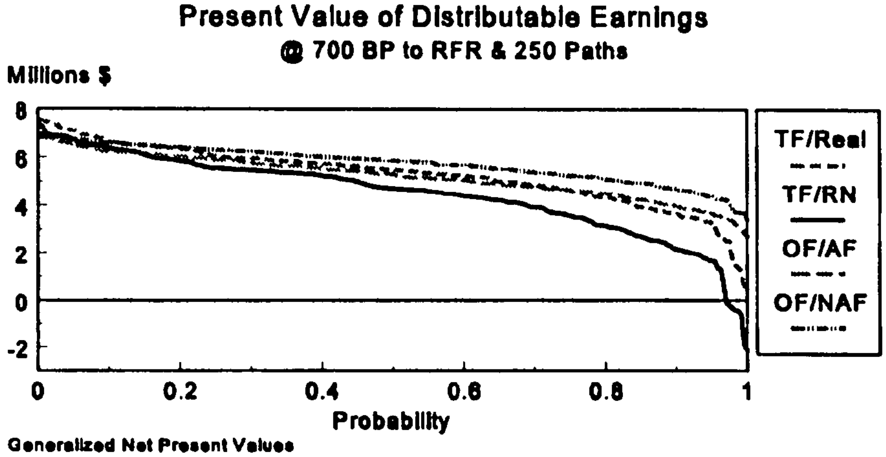


TABLE 8

Percentile Rankings of Distributable Earnings (250 paths) by Model

Percentile	OF/NAF	OF/AF	TF/Real	TF/RN
1st	6.9	6.8	7.5	7.0
5th	6.7	6.5	7.1	6.7
8th	6.7	6.3	6.8	6.5
10th	6.6	6.2	6.6	6.4
20th	6.4	6	6.3	5.8
30th	6.2	5.8	6	5.5
40th	6	5.5	5.8	5.2
50th	5.9	5.3	5.5	4.7
60th	5.7	5.1	5.2	4.4
70th	5.4	4.8	4.9	3.9
80th	5.1	4.5	4.3	3.1
90th	4.7	4	3.6	2.1
92nd	4.6	3.9	3.5	2.0
95th	4.4	3.6	3.3	1.7
99th	3.7	3.0	1.2	-0.6

First of all, one factor model percentiles are not quite stable at the 250 path level. Second, two factor model percentiles from the 1st to the 92nd appear to be stable once you equal or exceed 150 paths. The left tail, which represents the most favorable results, is stable. The right tail, representing the worst results, is not stable beyond the 92nd percentile. If a detailed analysis of the right tail is necessary, you'll need substantially more paths. But for making robust decisions 250 paths is adequate. Third, all distributions are negatively skewed and all have high positive kurtosis, i.e., there is more probability in the tails than would be found in a normal distribution and more probability close to the mean. This means that, when big deviations occur, they occur more frequently than one would find in a normal distribution, and they tend to be negative deviations. This is consistent with earlier studies showing that for interest rate risk, there is more downside risk than upside potential.

Fourth, consider the means of the four distributions:

- one factor/not arbitrage-free $5.7 million
- one factor/arbitrage-free $5.3 million
- two factor/realistic $5.3 million
- two factor/risk-neutral $4.5 million

The differences between any pair of these means is statistically significant. The relative decline in the mean resulting from using a two factor/risk-neutral interest model instead of a two factor/realistic model is 18%.

Fifth, clearly two factor models reveal more interest rate risk than one factor models.

FROM THE FLOOR: I have a question for Dave Becker. When you were talking about 150 paths, are those randomly generated paths?

MR. BECKER: All sets of paths were randomly generated.

MR. MATEJA: I have a question for the panel. Is it likely that, as a result of using the stochastic scenario approach to testing reserve adequacy, I will change the results that I otherwise would render as to the adequacy of the reserve? For example, if I pass the seven scenarios and I was thus prepared to render a favorable opinion, how could I relate that to the results of David's stochastic tests? Let's assume that 90% of the stochastic tests are adequate. How do you interpret those results?

MR. BECKER: The question is difficult, Mike, because what you're asking is, what if I run the seven New York scenarios and I passed them, and then I run a set of stochastic scenarios and didn't pass on a high percentage? Under the regulation, the definition of adequacy is up to the individual appointed actuary. And he or she chooses his or her own definition of adequacy. There is no absolute. But if a stochastic approach is being used, then it is not unreasonable that the passing percentage for adequacy should exceed 50%, but how much more than 50% isn't clear. What if you had a block of business that passed all the New York seven scenarios, but on a stochastic basis, it only passed 45% or 50%? Are the reserves and supporting assets adequate? I would say no.

MR. MATEJA: Perhaps that's a better way to state the issue.

MR. BECKER: It is only fair that the appointed actuary chooses the passing criteria before performing the test.

MR. MATEJA: Doug, what are your thoughts on this issue?

MR. GEORGE: First, let me say I agree with Dave. If you think about it, the guts of the whole cash-flow-testing movement and the actuarial opinion and memorandum regulation is to try not to make a strict test. The approach is not to define a minimum that must be "passed," or a threshold percentage that you have to meet -- e.g., passing at least six out of the seven or at least 80% of your stochastic scenarios. The whole movement is really toward more judgment.

I've worked with a number of different companies and found some disparity as to what is acceptable. I've talked to people who think that 75% is acceptable. I have another client that wants to pass 95% of its stochastic scenarios to really be comfortable and confident that its reserves are adequate, so I think there's a disparity. I would certainly agree that the percentage is greater than 50%, but I think it's well greater than 50%. I don't think I would be comfortable with 50% on just about anything. Reserves are supposed to be conservative by definition. Passing at 50% would seem to indicate that you pass on an expected basis, but not on a conservative basis.

The other point to make is that it really depends on your generator and the values that you put into it and the type of rates that are coming out. A couple of my charts showed that you can create some scenarios if you want to that you're going to fail. That's not very hard. You can also create scenarios that you're going to pass. You can put a very small volatility assumption into your scenarios or a big mean reversion assumption and create stochastic scenarios that all hover within a couple of percentage points of your starting yield curve, and you're going to pass those scenarios.

I think you also need to consider what kind of rates are coming out. When you see a stochastic scenario like the one I showed with rates that go up to 20% in three years and then three years later it's back down to 5% again -- quite frankly I'm not sure I care much about failing that scenario because I just don't feel that it's very realistic and, if it is, the whole industry is in pretty big trouble. Our reserves -- at least our current thinking as to where our reserves need to be set -- don't cover a scenario that is that extreme. I don't think we're really saying the reserves need to be that safe.

MR. MATEJA: Are there other questions?

MR. CHARLES V. FORD: I think Doug's analysis where he looked at the scenarios that failed is always useful. Such analysis reveals areas of concern such as really high interest rate scenarios, or the nasty inversions. We understand better what he failed. That's what lets him make a professional judgment about whether he passes in aggregate. My question actually is more on the software. Where do you get at this stuff? Is this built into TAS and PTS? Is it built into GAT, Intex, or CMS, and how do those relate? How do you link up the state-of-the-art asset model with the liability models of your consultants?

MR. MATEJA: The answer to that question, I think, is that there are commercially available interest rate model generators. There is in fact such a generator, a single factor model in PTS currently. SS&C/Chalke has developed COPE 2000, which was illustrated in some of our work that we've done here for the two factor model, which we think is pretty realistic. Cope 2000 is commercially available. But there are other commercially available models, and Doug, you have a couple of them, right?

MR. GEORGE: Yes, a lot of these models are being made more available, and in general we've done a lot of work, over the last few years, to integrate systems and models together taking "Wall Street" asset models and actuarial liability models and trying to take the best of both parts and putting them together. It can be done and some of it is rather simple. Some of it is just generating cash flows in one system and bringing them over to another. Some of it is more complex where you really want to get your systems talking to each other and feeding information back and forth in a more automated manner. It really depends on the type of analysis that you want to do, and what you want to accomplish. Some of it takes quite a bit of work and some resources, but it can be done, and it's being made easier. Even PTS has recently added an input, so it can accept some asset cash flows from external systems. For interest rate generators, in particular the ones I've used and I think some of the ones here the other guys have talked about, can be purchased separately as a stand-alone unit and fed into the different systems.

MR. BECKER: Actually, somewhat interestingly, in the last couple of years, both CMS and GAT, which historically did not make their interest rate paths available to the user, have changed their position. Formerly, you could use them to value securities, but you couldn't get your hand on the paths themselves. Now if you own GAT or CMS, you have the ability to use those systems to generate interest rate paths that you can export to other software systems. Those paths are relatively easy to bring into the major actuarial software packages. In addition, many of the insurance modeling software packages include interest rate generators. COPE 2000, which was designed for another purpose, can output paths directly to insurance software.

I want to follow up on Doug's comment earlier about pathological paths. That's a very interesting question that I researched some years ago. I did an analysis on the historic movements of U.S. Treasury rates and several of those articles have been published in *Risks and Rewards* and in the

Transactions. I wrote an unpublished document that I call, "Historical Stylized Facts," which identifies certain global characteristics of U.S. Treasury interest rates from the 1950s to the 1990s and presents statistical information about key maturities during that period. If you're trying to use a random generator in the realistic mode, which is what you would want to do for a risk management application, then you can take those output paths and subject them to the same types of statistical analysis and see how those compare to the stylized facts that are the one realization that we know of U.S. Treasury interest rates. For example, if you get interest rates shooting up to 40% and 50% or if you get rates above 30%, 20% of the time, if you have inversions more than about 15% of the time, if the inversions last more than about two years, then you can probably decide that this set of paths that you've generated is not reasonable. Using that type of an approach, you can calibrate an interest rate model to generate realistic paths and then validate that they do pass the "sniff" test by comparing them to these global stylized facts.

MR. FRANK G. BENSICS: I have a question for all three of the panelists, and that is whether they have a particular preference for generating forward, spot, or yield security.

MR. FITTON: Of course, if you generate a complete set of forward rates or a complete set of spot rates at all the maturities that are relevant at this point in time, they're equivalent to each other. You can derive one from the other. I tend to favor generating forward rates and deriving the others from them. Basically, forward rates are the most sensitive to numerical error because they're derivatives really rather than the original. They're in essence derivatives of spot rates. If you generate those first and then move onto spot rates and finally move onto yield, then you manage not to make your numerical errors build up on one another.

MR. MATEJA: I think with that I'll bring the session to a close. I want to first extend my thanks to the panelists for shedding light on what I believe is a new list of risks that we all have to worry about -- interest rate model risks. If there's one final thought I'd leave with you, it is this. It seems clear to me from these presentations that interest rate process is fundamental to the issue of value, and value is central to the responsibility of valuation actuaries. I think it's also clear that there's more challenge ahead. We're not to the end of the road where we can relax and be assured that there's nothing left to do.