2005 Valuation Actuary Symposium*

Orlando, Fla.

September 22–23, 2005

Session 36PD
Applied Modeling Concepts—Product Guarantees

Moderator: James Stoltzfus

Panelists: Dennis Stanley
Don Wilson

Summary: Implied or explicit product guarantees on life and annuity products add additional complexity to the modeling and valuation process. This session provides an introduction to building models that appropriately reflect these guarantees. Panelists discuss the modeling issues related to these features, including capturing "in-the-moneyness" in the model point selection; simulating future policyholder utilization behavior; fund mapping; model granularity; reflecting the impact of risk mitigation strategies; and simulating future policyholder fund selection and premium patterns

MR. JAMES STOLTZFUS: I'm a consultant with Milliman. On our panel today we have two distinguished gentlemen, Don Wilson and Denny Stanley. Don is a senior manager in the actuarial and insurance solutions practice of Deloitte and is based in the Hartford office. He's a U.K.-qualified actuary who moved to the United States just under two years ago, transferring from Deloitte's London office. Don's role focuses on asset-liability modeling (ALM), stochastic modeling and the applications that require these techniques. Since arriving in the United States, Don has been involved in a number of a client engagements relating to the audit review and implementation of Financial Accounting Standard (FAS) 133, Standard of Practice (SOP) 03-01 and risk-based capital (RBC) C-3 Phase 2 models. He has also become a regular speaker at actuarial meetings here in the United States.

Our second speaker will be Denny Stanley. He is a consultant at Milliman in the Seattle office. His areas of expertise are life insurance, product development and financial management. He's assisted life insurance company clients in the design

NOTE: The chart(s) referred to in the text can be found at http://handouts.soa.org/conted/cearchive/valact05/036bk.pdf and development of intersensitive life and annuity products, projections of financial results, allocation of investment income and recognition of federal income tax in the pricing of life products. He's also assisted employers in the analysis of life insurance to fund executive benefit programs. Since 1995 Denny has headed up the MG-ALFA development team and has overseen the market introduction and its growth to become one of the premiere life insurance pricing and projection systems in the market today. He's worked extensively in Korea, Japan and China at overseeing MG-ALFA development efforts there, as well as in the United States. With that I'm going to turn it over to Don.

MR. DON WILSON: The subject on hand for this morning goes under the grand title of "Applied Modeling Concepts—Product Guarantees," and it seems to be one of a series of presentations in these two days. We start with applied modeling concepts and cover various aspects, but I translate that to data modeling—creating a data model of the business, and that's the aspect that we're here to talk about today.

The way we're operating this session is that I'm going to review the various principles and processes more generally of creating a data model and then Denny is going to dig down in a bit more detail on the practical aspects of product guarantees.

What is data modeling? Basically it's the process for creating model cells to use in your projection system. In the United Kingdom, we tend to call them model points, and they're individual data records that you will run through your projection system instead of using the seriatim data. They represent a particular product line and a particular business. You would use this process when policy computations are too slow, unnecessary or inappropriate. It's always tempting to use policy-by-policy calculations. One reason for that is it's one less thing to persuade the auditors to sign off on. If it's seriatim, there's one less variability, one less area of issues, but not appropriate all of the time.

That reduces the number of records that you process, but also, when you're dealing with products with equity guarantees and separate account business, it's not practical to model every account separately. You may have a range of several hundred separate accounts that policyholders invest in. It would be inappropriate to do a stochastic projection of each of those separately, so there needs to be a process to distill that into a smaller number that you can then process.

Underpinning this process of data modeling is the question of materiality. We have to satisfy ourselves that the model that is coming out is materially accurate and that it doesn't skew the results or distort them and that it gives a reasonable representation of the results. That underpins the tone of the presentation this morning.

Why do we need data modeling? Historically, it might have been used in cash-flow testing and ALM-type projections, but today's focus is for products that have investment or interest-rate guarantees. Here a raft of regulations have come out in recent years that require stochastic projections: FAS 133, SOP 03-1, RBC C-3 Phase 2, and there are more coming. When I first came to the United States from the United Kingdom, I was befuddled by all the terminology and shortened names. Now I just reel them off. Those particular product lines with their valuations and their hedging needs all potentially require data modeling, but more generally any stochastic model may require data modeling.

As those companies do more around economic capital, for instance, there's a need for it, because if you have a certain amount of time that you can use for your projections, I would generally say that rather than using twice as many model points, it would be better to do half as many and do two runs on two assumptions that will give you a much better flavor and understanding of your data seeing two sets of results on two assumptions than just one more accurate number. Even if you have computer power to do things, increasing accurately, there's still value to be added in reducing the number of cells so you can do more ad hoc projections to see how your business changes under different circumstances and therefore more effectively manage it.

I've laid the process for data modeling out in five steps. Step 1 is the big one. Step 1 is where we're going to be concentrating most of our time. We need to determine the structure of the cells, which, in simple terms, might be saying we're working in five-year age bands. We're going to work in durations of 10, 15, 20 and 25 policy terms, and we're going to work in the money by 10 percent, 20 percent and 30 percent—that sort of concept. You have to determine which of the things we want to show in the cell dimensions: age, policy term, duration and how many cells in each dimension, so you end up with a multidimensional grid. It's not symmetrically multidimensional, but it can be thought of as a multidimensional grid. We're going to come back to more detail of that, but I thought it was helpful to lay out the total process before we concentrate on particular aspects.

Once you've determined that cell structure, the next step is a more mechanical process of putting each policy into the cell in the grid, so it's a question of running a database process or another calculation process that looks at each policy's characteristics and populates its interparticular cell. Having done that, the next step would be to calculate the characteristics of the cell. You may say you already know that because that was in your cell structure, but it's better to determine the characteristics of the cell based on the actual policies in the cell. If you're doing five-year age bands, keeping it simple, it may be that the average age is not the middle age in the five-year band; it might be skewed, and it's going to be better to take that skewed age. It will give you more accurate results in the model, and the same would apply for in-the-moneyness or any of the other factors. It will give you a better result.

Step 4 is perhaps a new one to some of you. It's one that's emerged like second- or third-generation systems. Having gotten policies into cells, you may find one cell with 10,000 policies in it, and the neighboring one has only one. Each of those will take just as long to process, but it's crazy to have one with one and one with 10,000. You won't get any proper value for the one with one, so it would be better to move it into a neighboring cell. There are processes around for doing that. Come to think of it, maybe I have this the wrong way around just thinking about it. You should determine the characteristics of each cell after having moved the small cells into a bigger cell. They have some effect on the results.

The last step, by far one of the most important ones, is you need to test that the model's okay; that you can validate it. If need be, you need to refine it, and we'll come back to that.

The big thing is creating the cells. Here, the first challenge is to look at your products and plans and reduce them into a practical number that you can reduce. There's no point in just doing policies within plans. Your minor plans ought to be aggregated with others. There's a host of factors that you need to consider as to what you can merge, because you need to keep the main financial aspect separate, so plans with front-end loads or back-end loads would need to be kept separate. You need to work through all you have and see what you feel you can merge together. Some things will be relatively easy to determine, and you could group them together. Bit by bit, it gets harder and harder as you get to the gray level of whether this is significant or not. There, I'm afraid, you don't have any simple answers. This is the point at which you stop. This is okay. You have to look at it all for yourselves and your own circumstances and how far you feel you need to go in terms of cutting down the number of cells.

Structure variations are important, but there are things you can do in the way you construct the model to reduce the number of cells. You can look at the mortality and expense (M&E) charges—if it's just the rate that changes and now the incidence that changes across the plans. You can group policies with different M&E charges together as long as in the cell calculation you work out the average M&E charge that is applied. A consequence of that may be that the cells you end up with don't represent any policy that you actually sell, but that doesn't matter if it's the best way of getting to the accurate answer. It doesn't need necessarily to represent a policy that you actually sell. Maybe you have to be careful if you have some table look-ups. You have to make sure that the cells do have entry points in the tables, so if it's something like an average age, you have to be careful that you don't go too far and then find that you haven't done your data model. It doesn't map onto the projection model.

Obviously you can include the commission payments themselves into the data model and make sure that each cell has in it the outstanding commission payments relating to the policies that are in that cell, so that you no longer need to make that part of the cell grid structure. You can just move it. The important thing is to make sure that any future commission payments are properly modeled and not lost. More important to today's session are the product guarantees. You need to make sure that in the data modeling process, you do not lose the dynamics of those guarantees and, most important, if you have different types of guarantees or different combinations of guarantees, almost certainly you're going to need to keep them as separate cells.

Having done that initial look at your products and assessed those groupings, you then need to look at the data the next level down. There are three key factors, products or guarantees that you need to consider: in-the-moneyness, the volatility of funds of modeling account values and policyholder behavior. We're going to come back and go through those in more detail in a moment.

There are other factors you need to consider. We've already covered product and plan variance. For reinsurance arrangements, if they're different you need to keep the cells separate. Attained age is important, particularly for a guaranteed minimum death benefit (GMDB) to capture the mortality payments that may fall due. Policy duration is important for things like surrender charges. Things to consider that are not necessarily so obvious include the timing to any trigger points where there might be policyholder behavior things like ratchets or resets. You need to make sure you're not losing those in the modeling. Keep them separate. I've already covered commission.

Here's a quote. There is little formal guidance on data modeling, but the C-3 Phase 2 working party did try to produce some guidance in that. I've just taken here a couple of quotes from the 150-page final report that it put together for RBC C-3 Phase 2 and anyone who is involved with that. It sounds more daunting than it is because a good proportion is on the alternative methodology, which few companies, if any, are adopting in practice. If you print it all out, you then take 50 pages out in the middle and shred them, and you're left with the bits that are more relevant. It doesn't say much more than I've already said. There are lots of things to be considered, and there's a need for validation at the end, as I've already said.

How do you define in-the-moneyness? Here I define it as the amount of the guarantee minus the account value, but it could be divided by the account value. It's a comparison of what the guarantee is against what the policyholder thinks his current value is. Once that guarantee gets greater than the account value, an aware policyholder would potentially change any decisions that he might make about the policy compared to what he would do if it weren't in the money. There's a need to be careful not to lose the in-the-moneyness in the data modeling process. Typically, the behavior and the financial effect go up disproportionately, almost exponentially of increasing in-the-moneyness. The more that the policies are in the money, the finer the grouping of cells. It might even get down to seriatim level for the ones that are most in the money to make sure that accuracy is captured.

Volatility of fund returns is an independent area of this calculation. We need to get from the separate accounts that may be used by the policies that you offer to the classes modeled by the scenario generator. There's a bit of give and take there because you also need to think about which classes need to be modeled by the scenario generator to fully reflect that range of separate accounts. There are a number of moving parts here, and you have to end up settling them all down. There are different ways you can go through this process, but a typical process to Step 1 would be to go through your list of separate accounts available and group them into what are called proxy funds. There's typically a many-to-one relationship, although some people use the term "proxy fund" in a slightly different way in this mechanism.

You would then end up with a smaller number—let's say out of 200, you're down to 10 to 20 funds, but with your scenario generator, maybe you have only six funds. You might have a NASDAQ, S&P, an international, a money market, a bond and another. You now need to get from your 10 to 20 down to the six. The typical approach there would be to regard each of the 10 to 20 as proportions of the underlying six, and you would do that by looking at the volatility and returns expected in each of them and map them across to the six by taking blended proportions. You have to be careful in this process. If there are any rules that limit the allocation, if the policy has limits on the spread of investments that it can hold, if they want the guarantee to apply, they need to have some sort of balanced allocation. Then you have to be aware of that in the mapping so that you don't end up with a mapping that steps outside that requirement.

Having done those two bits, in the projection model that you run, that's going to project each of the asset classes. Your generator will do a separate roll forward of the accounts based on those, so you will distill the blended proportions as if each cell holds proportions in each of those underlying six or however many accounts that you're modeling in a stochastic generator. There's a process to work through in that data modeling.

Obviously, the C-3 Phase 2 working party has something to say about that. The concept of materiality is acceptable to group if you don't materially reduce capital requirements, because that was the business issue it was addressing.

I like this second quote. "The actuary must carefully and deliberately map each variable account to an appropriately crafted proxy fund." You can hear a group of actuaries writing that, can't you? It's so actuarial! It doesn't help you do it. It just says you have to do it properly. There are a number of things you can use to help with the proxy in terms of things to look at, and the group's come up with a guideline: the portfolio objectives, the investment managers being asked to do for each of those separate accounts, the Morningstar classification, the actual assets held in each separate account, the historical returns (I guess that would include the historical volatilities on those funds), any performance benchmarks, the market on the funds and any AG 34 classifications. There is a host of data available to help you with that mapping.

Here's another nice quote: "It would be imprudent to ignore the concept of market efficiency in establishing the proxy funds and the associated model parameters used to generate the investment return scenarios." Market efficiency is the concept that the greater the return you get, the greater riskiness you'd expect that. You'd expect that any fund with a higher expected return would also have a higher volatility associated with it, so you would expect to see that pattern across all the funds that you end up with, and you need to take that into account.

A factor to be careful with in the correlation of the funds is to make sure that the correlations are sufficiently high, because while you could say in day-to-day markets, the S&P index and the NASDAQ might bob around, and you think they're moving fairly independently, but when something bad happens like a passing hurricane, everything tends to move in the same direction. It's those bad events that drive the high payouts in these products or guarantees, and so you need to make sure that you are properly modeling that and therefore the correlations need to be high to allow for the way the markets tend to move together at the points when things are bad.

There's no easy answer to policyholder behavior. It's a big question and a big issue as to how to model it. There are lots of things to consider in terms of driving policyholder behavior, who or how the policies were sold, what was said in the material, geography (maybe the people on the East Coast behave differently from people on the West Coast), and timing of elective benefits would obviously be important. Where things are most sensitive, you obviously need more cells.

You need to know your policyholders. There are a number of techniques for doing that in terms of exploring into the data. Some of the things that are available in data mining, going into the database of policies that you have in force or have had in force in the past, are techniques for connecting that up with external databases and the credit rating databases to give you a better understanding of your policyholders or people who might buy policies in general and therefore have a better handle on how they might behave in different circumstances and how they might act with this policy. This is something that might develop over time as this type of business becomes more mature.

One technique that we adopted for at least one exercise was to see the difference between an active behavior and a passive behavior. Passive is someone who did nothing based on in-the-moneyness and sat there with the static-based assumptions. The active one was the policyholder who knew everything and therefore was going to manipulate the policy to get the best guarantees possible. We did this mostly for the RBC requirement, but in the passive one there was no capital requirement at all, and the active one was off the scale.

The reality is somewhere in the middle, and it's a challenge, but it's important to the policyholder. It shows how sensitive the policyholder behavior rules can be, so not only the main focus that we're seeing is on the calculation end of how you would adopt the policyholder behavior, but what we're saying is you also need to consider it in the cells because the calculations will only be as good as the data that you feed into them. You need to make sure that the cells don't lose something that will drive behavior.

It's possible that you will feel that the behavior can't be trapped for any one cell in a single rule, in which case it's possible that you'll decide to do two or three calculations of three different behavioral rules for every cell you end up with. The ultimate of that is this data modeling process. Instead of necessarily reducing from seriatim to a smaller number of cells, you could end up with a larger number of cells to be calculated than the number of policies you have in force. If you do run multiple behavioral patterns for each cell, you need to adjust the results by taking proportions so that you get the weighted average for each behavior.

There are two sides to validation: static validation and dynamic validation. Static means you need to check that when you've done your data modeling and look at the summary of your cells, it should equate to the original seriatim in force, and this would be a check that you should make with any calculations to make sure that you've picked up all the data. You would be looking at the initial account values, the guarantees and the number of policies in force to be sure they are the totals that you thought they should be. If your data modeling process is accurate, that should come out correctly.

This point perhaps doesn't work with variable annuities (VAs) as much as with some other things, but if your system is calculating liabilities, but let's say if you have to do an AG34 calculation of a data model, you should check that the AG34 result was similar to what it would be on a seriatim basis. That would be proof that the data model was reasonably accurate.

Dynamic validation is looking at the projection and making sure that the values are reasonable, so that's a tougher one to do, but you can check that it's extrapolating from recent history in a reasonable way, and if you've picked up, you've not made a mistake with the margins and policies or something like that. Depending on your modeling system, it might be possible to back-cost the projections, see what with that data the results would have been in the prior year instance and then compare that with actuals and that would provide a check. Certainly, it would be, if possible, a good idea to test the data model against the seriatim run. For that test, if it's a stochastic projection, you might only want to do a single scenario, and you would pick an adverse scenario so that you can make sure that you're tracking the guarantees and check that it's not distorting the results.

In those comparisons, typically you set up a spreadsheet with two feeds with the actuals. One is seriatim and one is with the data models, and you set each cash flow against the two cash flows from the two feeds in two successive rows and maybe you color them differently. Then you do a visual check and maybe do ratio checks below that to check that they seem to be coming out reasonably and then you can do present values of those cash flows, because the present values should be pretty close. It may be that you grouped the timing of maturities so that there may be a big difference because maturities have jumped from one year into the next year when you did cash flows, but the present values would not be significantly changed as a result of that.

You need to monitor this whole thing going forward, so as with any other modeling, you need to be comparing any projected values and what you expected to happen with what actually happened. If you have a data model, you have another source of difference than you would have otherwise had. Maybe because there will always be the question, assumptions can't happen in practice. Your projection models have approximations, and now you have a data model in there and that may also be a source of error. The analysis of actual to expected needs to be fed back into the data modeling rule so it's continually refined. That means that with any experience analysis work you do, typically the lapses, surrenders, deaths, in this world we're now moving into, you may need to become more sophisticated and look at some of the behavioral things that are happening so you can test whether your assumptions are working in a reasonable model of what's happening and correlates any exposure to risk factors with what's happening there.

We mentioned generalized linear modeling. It's something that the property and casualty (P&C) people use a lot in determining what the significant factors are in the experience that drive behavior. In this case it's a statistical approach to determining which of the significant, multiple factors are the ones that are mathematically significant in driving behavior so you could do that sort of analysis.

Last, having set up some initial rules, you would expect to refine that over time as you look at the emerging results. Also it's not just your own results, but industry experience will emerge and maybe best guidelines and best practice will emerge to help and feed into the data modeling in the future. That's my general overview of the data modeling.

MR. DENNIS STANLEY: I have to admit it was interesting to listen to Don speak. I've known Don for 10 years or so, and he has a British background. Ten years ago, I couldn't have ever imagined being on a panel with Don talking about data modeling because I think he's a seriatim guy. I think when you start thinking about your benefit modeling, your natural tendency would be that it would be nice to be a seriatim guy, but because of practical constraints with the number of scenarios you're running, it just is not practical. You have to start getting a balance between the two.

I'm pretty much going to follow the agenda that was set up for us by whoever planned this session. There's going to be some overlap with Don's comments. I'm going to try to be a little more example-oriented and won't be quoting from the 120-page variable annuity (VA) working group document.

Let's start out with a little bit on capturing in the money. Historically when I built data models for a life model, it's not your exposure that's your concern as you're building your model; you're trying to deal with the mortality rate, so it's how you group policies with respect to age. When you're dealing with benefit guarantees, particularly guarantees that go in and out of the money based upon the equity returns, it's an exposure issue, not a mortality rate, that you're trying to capture.

As I said before, seriatim is a nice way of going about it because then you don't have to worry about it. If you have the ability to project your policies on a seriatim basis, you can forget about how much they are in the money, and you don't have to come up with the criteria for grouping. What you'll generally find as you adopt some type of policy grouping algorithm is you're going to start understating the value of your embedded options. I'm going to go through an example so you can better understand why that happens. Just as a general approach, the more grouping you do, the higher likelihood, I think, you're going to have that you're understating the value of the guarantees.

As I said before, you pretty much conclude because of your run-time constraints how much computing capacity you have, the number of policies, the size of your model, and whether you get results in real time. My criterion tends to be the model is not useful unless you can get results on an overnight basis. If it takes days or weeks to get results, that's too long. For some reason, my staff and I never have run a model once. It seems that we make a few mistakes along the way, so we have to rerun them several times. If your turnaround time becomes several days, it is not reasonable to get a decent turnaround, and many times sensitivities are more important than the absolute value of knowing the relative value.

To give you an example of how I've started understanding how in-the-money grouping can affect your results, consider a simple situation. You have a VA. You have 10 policies. Of those policies, the first one has an account value of 90; the 10th one has an account value of 99; and the GMDB benefit on each of those policies is 100. If you go through the sum of the digits math, you'll come out with an average in-the-moneyness for those 10 policies of 5.5 for each policy.

For these 10 policies, we have an in-the-moneyness of 5.5 for each policy, but it's a distribution. We know that if the stock market goes up 10 units (it's not quite that way, but assume that each of the policies goes up by 10 units, so our 90 policy went up to 100 and our policy with 99 went up to 109), all of our policies are now out of the money, and we have an exposure amount of zero.

Let's consider how we're going to take these 10 policies, and on a monthly basis we're going to transition from being in the money to out of the money. Think about doing it using the seriatim, you might say, of those 10 policies, a model of which you're grouping into two groups of policies: those that are one to five in the money, and those that are six to 10 in the money, or just do a single cell using the average of 5.5. You'll see in the first row of my table (Stanley Slide 5), all 10 of our models, whether we're using one, two or 10, we have an exposure amount of 5.5, and after an interval increase for the account value, up by 10, we have a zero net amount at risk. Our endpoints, whether we're using a one-cell or a 10-cell model, have the same exposure amount, but it's the transition as your policies are going from in the money, in this case, to out of the money.

The same thing can be said when you're going from transitioning from out of the money into in the money. If you look at those averages at the bottom of the page, you can start inferring that if you use a 5 percent interval for this increase in the stock market, I think that's about a 5 percent distortion in your net amount of risk. Similarly, if you use 10 percent brackets, you have a considerably larger distortion (the 1.8 versus the 2.2), so that starts to give you a sense of what's really driving this in-the-moneyness. As I characterize it, it's not the policies that are way in the money or the policies that are way out of the money. It's through your scenarios as policies go through the money that your group model gives you distortions because you don't accurately reflect how it goes from in the money to out of the money.

I'll give you some examples now. Stanley Slide 6 is a population of business running it on a seriatim basis versus using 5 percent intervals for in-the-moneyness looking at several statistics: the mean, standard deviation, a couple of conditional tail expectation (CTE) amounts and then looking at the low and the high scenario. As you look across that, if you calculate some ratios, you'll see there's about a 2 percent differential across that, so that's how much you've lost because of the policy grouping with 5 percent brackets. I would generally characterize this as satisfactory. If you want to narrow the market intervals, you can get those numbers down or the deviations down further. A lot of times if you start looking at your overall projections, while you may be misstating some of your guaranteed benefits, under a wide range of scenarios, there are still a lot of other cash flows that your data model is doing a good job of capturing, so this 2 percent error is only a 2 percent with respect to one element of a projection. I think these are 500 scenarios that ran that had this 2 percent variance.

Another thing I like to do to deal with in-the-moneyness is set up your model so you understand and can check at a high level and in an aggregate way of: Are the policies moving in and out of the money given a scenario in a way that makes sense? For this same population of policies, I looked at scenario one and scenario five (Stanley Slide 7). For scenario one, it was basically a monotonically increasing scenario except for year two, which had a dip. These are total return indices, so they would have price and dividend returns built into them. For scenario five, I intentionally selected it so you could see that that's a trough scenario. The market goes down, you stay down for five years, and then somewhere between years five and 10 the market rebounds. You would expect just looking at those that for scenario one, we should start seeing the business go in the money further in year two, relative to how it started at the beginning of the projection. Then for scenario five, you would expect to see the policies move in the money, stay in the money for a period of time and then come out of the money in the later years of your projection.

We won't go into too much detail, but if there is one number right in the section here, you can see that five bounce up (Stanley Slide 8). In scenario one, you'll see that. This table shows the distribution of the policies at the beginning of the projection, what the distribution was in-the-moneyness, and then the rise represents time of how the in-the-moneyness changed. As you see numbers move to the right in this report, that means the policies are going in the money; as the policies move to the left, more and more of your account value is moving out of the money and you'll see that five in row two, which means we have some policies that bumped up; they're 150 percent in the money, whereas the prior year we only had two. That correlates back to this 973 equity return in year two.

Similarly you'll see in scenario five (Stanley Slide 9) a lot more shifting of your business to the right for a few years and then it starts moving back. What I find is that the actual numbers of this report are not important, but this kind of report is useful, so you can check your model and see how your different types of policies are moving around. If you had run this type of report on a block of VA business and separated your policies between return-of-premium business, your ratchet business and your roll-up business, you'll see a lot of different characteristics of how the in-the-moneyness will move around for different types of policies. I find this kind of report useful to see if everything seems to make sense, if you take some bad scenarios and do your policies and move in and out of the money, as you would expect given the product design.

Don was mentioning maybe there are some variances between policyholder behavior on the West Coast and on the East Coast. It's a subjective issue. If you go back 15 to 18 years ago when we had started doing cash-flow testing in the United States and the high interest-rate environment and all of the high lapse rates that we saw in disintermediation particularly in the single-premium deferred annuity (SPDA) business, we all had a sense that antiselection was going on in the policyholders. It was quite pronounced, and we were faced with building behavior models that had to reflect the expectation that your policies would disintermediate as interest rates rose.

Two observations have been interesting. If you follow the VA business, there was a debate between the industry and the regulators a couple of years ago about dollar-for-dollar withdrawal business. If you look at it on the surface, that's a rich benefit, but to date the antiselection of that business has been modest, which is surprising as an actuary when you sit back and think about the value of that benefit. The other thing that is interesting is that as annuity business comes out of the surrender charge, similar to fixed annuity business, you do see a spike in lapse rates, quite high lapse rates, which we know is driven by age of compensation issues as opposed to value to the policyholder. When you step back from that, the dollar-for-dollar business hasn't shown extreme levels of utilization that you might expect, and the high lapse rates that we're still seeing through a relatively tough stock market the last several years make you think how sophisticated policyholders are.

If you look at your business, if you look at cohorts of when policies were issued and the type of guarantees you provide, you've had a few experiences over the past five years of the stock market declines. We had the dot-com decline in the stock market. We had the decline in the stock market right after 9/11 and then as the Gulf War began, so you can follow some of your utilization of benefits and the lapse rates and get a sense of how sophisticated your policyholders are. As I said before, the observation I have is a lot more policyholders than you would expect seem to be throwing away valuable in-the-money options on their policies. At this point I'd say that the utilization of valuable options has not been as high as I would have expected from the relatively small experience I have seen.

If you're going to deal though with policyholder behavior, which I think you should in your model, you need to decide how to go about it. This is a model structure for a project we've been working on recently. It's a block of VA business. You can characterize it has a rich GMDB and a guaranteed minimum income benefit (GMIB) benefit. We set up a model to have utilization rates, specifically as policies went in the money further, that the lapse rates would go down, and as policies went into the money, the utilization of the GMIB benefit would increase. We had a core model associated with that. Thinking about the policyholders, we then layered on top of those assumptions their level of sophistication being low, medium and high. Those are qualitative terms as opposed to quantitative, but we started thinking in those terms and assumed that we have lower levels of utilization or antiselection for the low sophisticated policyholders and the medium and the high. The 50/30/20 were arbitrary selections on our part, but at least it gave us a framework for understanding the variance and the benefit cost for those types of policyholders.

The other important factor is the type of scenarios we were looking at. We ran this model using two types of scenarios; neither one was intended to be risk-neutral. The first was calibrated to historic returns. When you look over the past 20 to 30 years in the U.S. stock market, you see the return of volatility in the U.S. stock market. We were looking at expected returns in the 12 percent range. Also, we reduced that prospective view of expected returns from 12 percent down to 8 percent, similar to the VA working group and its recent modifications of the Academy scenarios to use a lower expected return rather than the historic.

Stanley Slide 12 gives the summary of our three assumed cohorts of policyholders—the low, medium and high sophistication—and then the middle column is the benefit cost. These are all end-benefit costs for GMDB and GMIB expressed in basis points for those policyholders using historic scenario sets for a composite of 25 basis points. You have about a 50 percent range that the low sophisticated costs are 21 and the high are 33, and the weights that I used were pretty close to the low end of the spectrum. The 50 percent range is a nontrivial variance of benefit cost. It demonstrates the uncertainty as actuaries we had as to what our assumptions are, because you could characterize the cost of this benefit. We may have a 50 percent uncertainty in the benefit cost because of the uncertainty of the sophistication of the policyholders.

As you move to the next column though, that's probably even more striking. You'll see that those costs are about four times the middle cost, so it's the nature of these scenarios that we're looking at. We have the same level of volatility but much lower returns, and to some degree that's even a more important assumption in your model—the nature of the scenarios you're looking at as opposed to policyholder behavior—but again you still have that 50 percent range: the 87 up to the 122 for the low to high sophistication, so it's still pretty much a material assumption.

I won't go into too much detail, but Stanley Slide 13 drills down and gives you more statistics that correspond to the middle column: the historical calibrations and scenario sets. The three middle columns—the low, medium and high—are the policyholder sophistication. The right hand is the composite, so you can start seeing the mean standard deviation and some of your CTE numbers. You're not going to get a lot of new information. It just gives you a bigger picture of the results.

If you look at the mean over the average, the 25 basis points, that's your expected cost if you were using historical scenarios, although for the VA carbon you would not be using historical scenarios now for the 90 CTE and the RBC C3 calculations. You can see some big jumps from the 25 up to the 130. That's not telling you a reserve number, but it gives you an indication of the average cost of that tail. There is similar information for the lower sets of returns. Again, these were four times more severe as far as benefit costs for different cohorts of policies. That's all I need to say about that. It's not important to go through the numbers, but I thought it would be helpful for you to see some of the distributions.

With fund mapping, as Don indicated, because of the number of proprietary and other types of funds that are available in the VA and the variable universal life business, it's not practical to consider a seriatim approach to fund modeling. I would characterize that there are two approaches that are in practice today using the AG34 assignments or one that we've used considerably within our firm. We're doing a regression analysis of the historic returns for your funds relative to your index funds that you're using and doing a one-to-many mapping, so a specific fund is mapped to your index funds. Some of the common funds that we would use include three domestic equity funds, an S&P, a Russell 2000, a NASDAQ, an international equity fund, a bond fund and a money market fund. Those types of funds feel right as the types of funds you would have to have on your variable side of the business. We're building a model for the general account and the separate account business, besides those six funds, of which we would need to make sure we've calculated the weighted average management fee and the gross management fee that's being charged by the funds so that our model index funds appropriately have the returns reduced for the management fees, and then what portion of those gross management fees are ultimately profit through the revenue-sharing arrangements with the fund management.

On the general account, depending upon the type of business, we might end up with three funds for a model point. If there is a dollar cost averaging for the newer policies, we might end up with a balance of the dollar cost averaging amount, the interest rate and the number of months remaining and then have a relatively simple model that that balance will transfer from the general account into some distribution of your funds over the remaining period. Similarly, if you have a modified-guarantee-type product structure on your general accounts side, we might calculate the weighted average interest rate and the weighted average period remaining for each one of our data points and then treat that as a single bucket of money that will accumulate to the maturity date and then transfer it appropriately into some mix of general account and separate account funds and then finally the other general account, which typically would have some type of declared rate and a general account interest crediting strategy.

The subject of model granularity is a bit of a repeat from the in-the-moneyness, but it's good to step back and think about the level of model granularity you'd like to have. It's good to think about the purpose of the model as you're setting it up, as Don indicated. Having a seriatim model is desirable for a variety of reasons. It's useful for validating your group policy model, and it appears to be becoming a necessity for the movement for reserving in the United States using stochastic processes that the regulators are more comfortable of having the standard scenario with some type of a seriatim process applied to it. If you need to have a seriatim model because of regulatory considerations, it's also useful to have that same seriatim model so you can get a sense of how your group policy and your seriatim model connect to each other.

As you're designing the size of your model and the granularity, you need to think about what type of access to grid computing or computing horsepower you have, how many scenarios you're going to want to have in your typical run and how long it might take to run each scenario. Coming back to my earlier benchmark, overnight runs are desirable. We're seeing that companies are moving from trying to run their models on a handful of machines, say one to five, to developing relatively large high-density computing access and considering having 100 machines available for an overnight run. As you move from one machine to 100, you pick up two orders of magnitude with respect to the number of model points you can put in your models. As you have access to that kind of computing capacity, it changes the way you think about the number of scenarios and the number of model points.

Here are some numbers that I put together from a large block of VA business. It was less than one million policies, but it was approaching a million policies on a seriatim basis. We were able to use some fairly granular criteria, I would characterize, for grouping the policies and getting tight static and dynamic validation statistics as Don described, using 50,000 data points. For doing high-performance stochastic runs, we found that was a suitable model for doing business planning projections. If you're looking at precision in your earnings over a three- to five-year horizon, I'm not sure whether we had those data points on a quarterly basis. I think they were probably more or less annual data points, so if you were looking at annual earnings, they were a good representation of the business. If you were looking at quarterly earnings, we may have wanted to boost that number of data points up to more like 200,000 if you're looking at quarterly earnings.

We felt that for stochastic analysis, we could collapse that down to about 25,000 model points and get reasonable overnight runs using access to a grid. Then when we moved into doing stochastic–in-stochastic or nested stochastic projections, in which not only did we want to calculate the embedded value across 100 scenarios, but we also wanted to reflect the RBC C-3 Phase 2 calculations across those scenarios, not only did we have 100 stochastic scenarios, we wanted to calculate the capital requirements across the scenarios. We reduced the model size even further because of the intensity of those calculations. As Don said, though, it's good as you start building those models that you understand your purpose of being able to reconcile them back to the higher level. When you're working with the 50,000 cell business plan model, compare results to the seriatim model, understand your differences and become comfortable with them, as you move to the 25,000 or the 5,000 minimodels and see how those results compare to your 50,000 business plan model, again, understand the relationships and be comfortable that you're satisfied that what you have lost because of compacting your model is acceptable and that you understand.

Some of the criteria that we would typically use in getting down to that 50,000 cell model of group issue years—issue ages somewhere in the five- to 10-year range—depending on whether we're looking at annual earnings versus quarterly earnings, we'd group issue dates that are either annual or quarterly. You need to look at all of your guaranteed benefit combinations, separating your roll-up business from your ratchet business, the ages that the benefits cease and the various surrender charge combinations. If you think in terms of a traditional life model having a set of plan codes representing all the combination of benefit structures, hopefully you would have that kind of a broad data model from your valuation system that will allow you to separate the policies out. Thinking about when benefits cease, if your roll-up benefits stop at age 80 or your ratchet benefits stop at age 80, reflect that in your separation of policies, the policies that have reinsurance coverage and different types of reinsurance coverage, in-the-money brackets as we talked about before.

One criterion that we often put on top of the model after we considered all of the other criteria above is that for each one of those, what would have been model points, we separated into three, and the three separation is to separate out the policies that have high exposure to equity returns versus low and medium exposure to equity. We had to get the high-risk policies separated from the lower risk policies, so we separated out and captured that volatility. The way you do that is, for each of your model funds, if you have a volatility index associated with it, you can calculate the weighted average volatility of a policy based upon its distribution of account value across the six separate account funds. As I said before, separate all the policies that are being grouped together into low, medium and high because you can rank them and put one-third of the policies into the three brackets so you'll have more granularity in your model.

Of course, as you group everything, as Don said, you need to then circle back and calculate all the appropriate weighted averages that you need for your group policies to deal with surrender charge schedules, your M&E fees, the charges for the guarantees and the policies that are grouped, but it's relatively straightforward. If you have a seriatim model and have all the data to support a seriatim model to appropriately calculate those weighted averages for the policies that are being grouped, it's a relatively straightforward exercise. The key is having the data to begin with to support your seriatim model.

Regarding risk mitigation strategies, reinsurance modeling is relatively straightforward. In terms of your cellular model, you can reflect the terms of your reinsurance agreements. It can also reflect the reinsurance premiums that you pay and how much you expect to get back, ignoring any inside limits in the reinsurance treaty on an aggregate basis and then across your projections for a scenario. If your treaty has aggregate limits, apply the aggregate limits on the rolled-up cash flow so you can reverse out if you're not going to receive all of those reinsurance credits or cash flows because of aggregate limits.

Dynamic hedging strategy, or any type of hedging strategy, whether you're using a systematic purchase of futures contracts to do delta hedging or purchasing actual options, becomes difficult. I'd say that the people are feeling a need to reflect their hedging strategy in their projection models, but I wouldn't characterize that overall people are there yet. There's some tendency initially to treat it as reinsurance, that they've considered what the costs and what the expected benefits of the hedging program are and to begin with some relatively simple assumptions in the projection model. I think over the next year we're going to see more movement toward building stochastic-in-stochastic models, where you're reflecting the periodic determination of the valuations of all the Greeks that you need to determine what your hedge position should be and then simulating your actual hedge strategy as you envision it, so it's a more explicit representation of your dynamic hedging strategy, if you have one. It's an evolving state that a number of companies have a dynamic hedging policy in place and certainly are intending to get more explicit recognition of how that policy is expected to function in their financial projections.

Now, we're down to the cats and dogs of fund rebalancing and renewal premium assumptions. I would characterize, at least for VA business, that people are generally not reflecting that in their models. If you're dealing with secondary guarantees on universal life and the variable universal life, certainly a renewal premium assumption would be appropriate because of the importance and the expectation that renewal premiums will flow.

If you think about some automatic fund rebalancing that may be in some of your contracts, for a natural hedge, there may be requirements. The account value to get the guarantees, some percentage of the account value must be in money market and general account or bond funds. You can put those types of assumptions under your model, but other than contractual limitations or requirements, I would characterize that people are not reflecting the systematic movement of policies or the assumption of how policyholders will behave in moving their account value among the funds. However, if you do have a view as to those assumptions, you can certainly reflect them.

For future premiums, as I said before, if you're modeling a block of life insurance business where you have a high degree of confidence that there will be renewal premium, you should reflect that. Also, if you have some type of a benefit structure where you know you have a rich benefit and expect there will be some utilization of that rich benefit through future premium deposits on a VA block, you may want to reflect that or other experience studies that suggest that occurs. Generally, I'd say it's easy to model the fund rebalance in the renewal premium and build it into the models, but, it’s more difficult is to come up with a view as to what the assumption would be.

MR. MICHAEL E. DUBOIS: This one is for Denny. Taking a look at the example on the policyholders' sophistication, would it be correct that the take-away there in taking a look at the difference between the low, medium and high (and there appears to be a significant difference) is that one of the key concerns is we have when taking a look at making some of our estimates as to margins, etc., is to have some awareness of the potential for the changing nature of the sophistication of our policyholders?

MR. STANLEY: Exactly. There are two things related to that. If you've done any work with the guaranteed minimum withdrawal benefit (GMWB) business, you know that the sophistication can have a big impact on the expected cost. With market pricing today, you may assume that your sophistication is relatively low, but if you have some type of a roll-up or a ratchet benefit five years from now, you need to consider whether the policyholders will be proportionately more sophisticated five years from now than they are today, or will it become vogue for life-settlement-type operations to start purchasing valuable products? Certainly those are two ways in which sophistication could improve in the future if you're thinking about that.

MR. MIKE Y. LEUNG: My question is related to selecting samples for analysis. We've been talking about doing seriatim valuation, and then you condense to smaller cells. My question is, Instead of picking representative cells, what do you think of the method of doing a random selection of policies, for example, you do a random draw of 1 percent of the policies and do an analysis of that sample? It does have the simplicity in terms of doing the selection, but I don't know how appropriate that you think this is for the kind of actuarial analysis.

MR. WILSON: We probably both have a comment to make on that. I would say that it's probably generally not appropriate to use it, but there's one particular circumstance when it would be: when time is of the essence, and you have to get some results out tomorrow. The only way of cutting down the run time is to do a sample because you can do that more quickly than building a data modeling set of algorithms, so it would be acceptable if you were hard-pressed on time, but that's the only circumstance that I can think of relative to doing a more scientific approach.

MR. STANLEY: I was going to say it may be possible. I've never tried it myself, though I think you need to go through the static and dynamic validations for your sample size and get comfortable that your random sample was capturing what you thought the key demographics of your business were.