# RECORD, Volume 27, No. 3*

New Orleans Annual Meeting
October 21-24, 2001

## Session 101PD
## Regulation XXX Update

**Track:** Product Development

**Moderator:** SHELDON D. SUMMERS
**Panelists:** DOUGLAS L. ROBBINS
MICHAEL S. SMITH
MIKE ECKMAN
SHELDON D. SUMMERS

*Summary: Regulation XXX has been enacted in a majority of states, yet there is still plenty of uncertainty.*

*Panelists present:*

- *Current product development efforts—term and universal life*
- *Current regulatory guidance*
- *X-Factor certification experiences*
- *Potential impact of the 2000 CSO table*

**MR. SHELDON D. SUMMERS:** The first panelist today is Doug Robbins. Doug is a consultant at Tillinghast-Towers Perrin's Atlanta office. He practices primarily in the area of individual life insurance and annuities, including pricing, evaluation, embedded value, and asset-liability matching. He was a member of the task force that developed Actuarial Standard of Practice (ASOP) 40.

To his left is Michael S. Smith, vice president and appointed actuary of Lincoln National Life. Mike is a chief actuarial officer for Lincoln National and is responsible

for the actuarial opinion and memorandum, surplus management, asset-liability management, and other related matters.

Mike Eckman is second vice-president and appointed actuary of ING ReliaStar. In addition to his appointed actuary responsibilities, Mike works with various units within ING's life group to write and implement corporate standards with respect to the pricing, accounting, statutory reserving, and tax reserving of products. He has also worked on acquisitions of life insurance companies, including the acquisition of ReliaStar by ING.

**MR. DOUGLAS L. ROBBINS:** I am excited to talk to you today about some of things that have been happening with XXX. In particular, I'm going to cover X-factor certification experiences and XXX product development efforts, although there's probably more to say about the first than the second on a recent basis.

I'm going to begin by talking about a few different perspectives on XXX certification experiences. I've spoken to several practitioners in the industry to get the company perspective, a handful of regulators to get the regulator perspective, and I'm going to give you a few of my own thoughts. I will address recent product development efforts, and then I'll summarize.

When I spoke to industry practitioners, I asked them about their certification experiences: what they did that went well, and, what they did that didn't go so well. I asked them a few pat questions, but mostly I just let the conversation flow. I'm going to relate to you what I learned.

I'm sure all of you are familiar with the fact that when you do your certification of your anticipated mortality, you do it first for each of your X-factor classes and then you do it for the company and aggregate. Most companies' X-factor classes were very much in line with their underwriting classes. In other words, they were grouping across age, gender, duration, and things like that. They were mostly differentiating the testing they did only by their underwriting classes—smokers, nonsmokers, preferred, super-preferred, etc.

I learned that Monte Carlo testing was definitely the favored method of certifying the X factors. We know that X-factor certification at the end of the year really involves two steps. You first certify your anticipated mortality rates using Monte Carlo or some other method. Then you certify using a couple of deterministic tests that ensure your X-factor driven mortality is properly related to anticipated. The Monte Carlo methods are favored for certifying the anticipated mortality and establishing that it hasn't changed from what we thought it would be at the beginning of the year.

There were other similarities. For most people, this past year-end was the first year that they were dealing with policies that were subject to XXX. They were mostly starting with their pricing mortality, so this year was an assessment of pricing

mortality. Some people found that their pricing mortality was conservative, a few found out not so much. The people found that their mortality was either adequate or more than adequate—it had sufficiencies in it.

There were other techniques employed for certifying anticipated mortality. The next most common was the Harry Panjer's method, and finally, several companies just made a straight comparison of actual to expected mortality. It's especially true, obviously, when actual is less than expected, because if that's true, you don't need to know a lot about the right tail of your potential distribution of claims, given your anticipated mortality. If your actual for each class and as a whole is less than your expected, you know you're in the left tail; therefore, you know that you're basically sufficient.

The methodologies the different companies used for Monte Carlo were also similar. One thousand was by far the most popular number of trials that others in my firm and I have heard; 300 was the next most popular. A few companies tried 10,000 but found that in most cases that took too long for what they needed. It was overkill.

Why would 10,000 be overkill? If you are under any time pressure, the only time you would need to have that much detail would be if you were very, very close to the dividing line that you had set when you decided what confidence range you were going to use for testing your anticipated mortality. If you decided you were going to reject anything over the 95th percentile and your result was 60, you're not going to need 10,000 scenarios. You might not need 1,000; you might be able to tell with absolute certainty with 300.

You might need 10,000, on the other hand, if you end up at the 74.5 percentile and you were going to reject at 75. With 1,000 scenarios, you might be only 85 or 90 percent certain that you're on the correct side of the line. To be certain that you're on the correct side of the line, the 75th percentile in this example, you might need 10,000.

There were other similarities in Monte Carlo testing. Most people did 1,000 or N, new trials for each cell. No one talked about doing random sampling; they pretty much certified using their entire database of lives over the course of XXX sales. They also generally constructed confidence intervals around our expected result with the intent of not adjusting, if actual results were reasonably in line with expectations. In some of the literature I've read recently, the 95th percentile is specified as where you would reject. A lot of the people that I talked to were rejecting from as low as the 75th percentile.

It's important to understand that the implications of those two numbers are the reverse of their usual implications. If I'm testing reserve adequacy and want to be 75 percent confident that my reserves are adequate, if I pass 75 of my scenarios—my cash flow testing scenarios or whatever—then I'm going to be happy. If I said

95th, I'm going to be happy if I pass 95. When you do the X-factor testing, it's important to understand that if you set a parameter of the 75th percentile, you're going to reject the X factors that you've set more often, which means you're being more conservative. If you reject at the 95th percentile, you reject results that would be correct one time in 20, but you're also going to reject incorrect results less often. So you're being more aggressive. We'll discuss some of the statistics later on.

From talking to practitioners, I learned that some of the valuation actuaries wished that they had talked to their pricing actuaries more before they started the process for the year. The pricing guys thought anticipated was what had always been used and it would be fine. The valuation guys actually knew something had changed, that there was a likelihood of higher mortality for a certain cell, etcetera.

Some people that I talked to saw dramatic improvements in the quality of their systems' output over the course of a year because they knew XXX requirements were coming. This is obviously a very good thing. A lot of companies have not had excellent means of gathering data to do mortality experience studies. Now XXX mandates that you do an experience study every year. It creates a situation in which we're going to have better data quality.

Some of the smaller companies that I talked to had no claims yet within the entire XXX compliant block. They could use the actual less than expected test that I talked about, but some of them decided that to pave the way for future reporting, they'd report to the regulators in terms of their "if" statements. "If we had actually had claims of more than $15 million, then we might have rejected X factors for this X-factor class, which would require us to reset them."

A couple of questions are unanswered in my presentation. What if underwriting classes change in the middle of the year? How do you manage your data so that you get the right information on the right expected classes? Another question is, what do I do in a situation in which I have one huge claim that distorted all my results on a dollar basis? We're supposed to use total face amount, not net of reinsurance, but what if someone died who had a $15 million policy? We're going to talk a little bit about that later. That wraps up what I know about the company perspective.

The regulator perspective was also interesting. I talked to a few regulators who agreed to speak with me on the condition that they remain anonymous. Most of them were willing to talk quite a bit. Regulators have been demanding about certain aspects of XXX work from companies, as you've seen and as you will see, when my two colleagues speak. So I was wondering if they might think people were cheating on X factors and were dissatisfied with what they were getting.

In fact, that was not true at all. Most of the regulators I talked to seemed either unconcerned—in other words, they felt like they had bigger fish to fry—or they

were impressed with industry efforts. One regulator I talked to was very, very diligent, detail-oriented, and actually verified Monte Carlo results independently. But he was also satisfied with the results he got from the companies he talked to. Overall, I'd say regulators are apt to let the actuaries' judgment stand unless experience proves that the actuaries' judgment needs to be refined. So I think overall, the industry has done a very good job so far. I think regulators expect to see improvement every year, because we're in a learning curve, but that's to be expected.

One regulator did tell me that I could make a specific request. . He said it would be very helpful if you would include mortality rates, line by line, for each record when you send in seriatim data, so regulators could check your work on certification of anticipated mortality. The person did his own Monte Carlo testing and said that it took more time for him to build in look-up tables of all the different anticipated mortality rates, than it took to do the Monte Carlo run. Building both anticipated and X-factor-adjusted mortality would actually be of interest.

I was a statistics person coming out of school. That was my area of concentration. I think that a lot of what's behind the ASOP and a lot of the anticipated mortality testing are driven by statistics, so I wanted to talk about a few of my views. One is that it's very important to set your confidence level or rejection level before you start testing. From a stochastic perspective, the confidence level or the rejection level doesn't have nearly as much meaning if you let your results influence it.

By that I'm not saying that somebody started out at 85, ran their tests, erased the 85, put in a 95, and said, "We pass everything." What I'm saying is that even if you look from class to class, you start with preliminary X-factor classes and then decide halfway through testing that because of credibility or anything else, you really need your rejection level somewhere else. You've kind of rigged the test. Kind of like rolling a dice, you need to know what you think the truth is before you start. You can set the parameters such that you'll reject at a 95 percent confidence level, and then you either do or you don't and then you've created a situation where your Type I area is what you say it is, 5 percent. At the aggregate level, of course, failure at the level you've chosen ahead of time should in most cases lead to action. So that's the other reason it's important. You will want to make sure what you've set is appropriate for your company and then you will want to act on it.

Having said that, what if the X-factor class level, what if my confidence level is 95 percent? In other words, my Type I error is 5 percent. When I do my hypothesis test, I want to reject my null hypothesis at a 5 percent level. I've said that three different ways, but the point is, that's how I set up my test. So I have a rejection criterion. I have 40 expected classes. One of my classes fails; does that mean I need to adjust that class? Well, let's think about what this means.

By setting my confidence level at 95 percent, I expect to be wrong one time in 20 when I make a decision. If all my X factors are correct and I test 20 independent

classes, I should expect one of them to fail. If I've tested 40 independent classes,, I should expect two of them to fail. So if only one of them fails, I may not have any reason to adjust it. I think you want to base this decision-making on where you come out in aggregate. If you come out in aggregate up in your upper tail somewhere, but not at the rejection criteria, you might want to adjust this class to affect the whole test. You might start to be suspicious that you might have anticipated mortality to be too low. However, if you were right around the median with your expected testing and you have one class failing, you can certainly chalk that up to just statistical randomness, overlook it, and just wait for more data for that class.

What if many of my expected classes have less than one expected death? It's almost the same issue. If that's true, you might be able to make a case for saying you have too many X-factor classes. If you feel like you need all of those classes, because it's just a super-duper-preferred class and you just don't have a lot of lives there yet, then you're saying you don't have credible data. If you don't have credible data, you really can't do much with that class yet. You need to aggregate your data and base your decision on other things. That particularly applies in the case of the single $15 million death claim that I mentioned earlier. Chances are, somebody with a $15 million face amount is going to die, and it may impact an X-factor class really badly. Again, you probably need to base a decision on that class by how you're doing in the aggregate.

It's also important to know what the implications for Type II errors are if you set your confidence limits too high. I do not have a definition for what I mean by too high. I like the 95 percent number because I like permitting the actuary to use judgment, since I trust all of us to follow professional standards and act professionally. This means trusting the actuary's judgment and permitting a rejection criterion like one out of 20. So we allow enough time for experience to unfold before we start making decisions. But, as I said before, we need to know what we're doing when we set it at 95 percent.

This is not a realistic example, but I just want to use it to make it clear. Let's assume we have one (rather old) life in a class, whose anticipated mortality is 0.050. You could do a one-person hypothesis test and use a 95 percent confidence level very easily. If my null hypothesis is true, and his anticipated mortality is 5 percent, then he should only die one time in 20. So if he dies, I failed the test at my 95 percent confidence level. What if he's four times as likely to die? His anticipated mortality would increase to 0.200. In reality, I still have an 80 percent chance that I will make the incorrect decision and not raise my X factor. So that's what I'm getting at with Type II error.

My example is silly because it has only one life. It doesn't change. If I change my limit to the 75th percentile, I'd have to reinvent my example. But you need to realize that when you make your rejection criterion higher, you are putting in more Type II error, and you can't actually know what your Type II error is without

knowing the truth. Nobody can know that; that's what we're getting at. I think everyone has done a good job, as we said earlier, but it's important that everybody goes into testing every year realizing this. While the X-factor class level may not, at times, be so important due to insufficiently credible data, at the aggregate company level (where hopefully you're credible enough), you have enough lives such that the Type II error can be kept low, as well as the Type I error. But if your actual mortality starts to creep up into your upper percentile, you might want to at least note the situation, even if you're not going to reject and raise X factors immediately.

One more observation I want to make involves run time on Monte Carlo testing. Monte Carlo is presented by a lot of the regulators and practitioners as a very handy way to get a feel for the entire distribution of potential claims if our anticipated mortality is correct. However, even without doing convolutions, which would be the way to get the exact distribution, but would be very, very involved and take forever, you could run into run time problems. Is there a way to speed up run time? I think there is a neat way to do it. I'm going to give credit for the thought to Ed Robbins of Zurich Life, who is also my dad. He, in turn, would credit a *Transactions of the Society of Actuaries* article dating back 20-plus years. I didn't actually find the article, but the development of the theory behind this is pretty easy. At some point, you may have a close call, the 74.5 or the 94.5 percentile that we talked about earlier. You may need to do extensive numbers of trials. It could happen even in aggregate. Let's say you're going to do 10,000 trials for 10,000 policies. This involves creation of a grid of 100 million ones or zeroes, where one equals a death, and a zero equals no death. For the ones, you're going to say that amount of face amount died, and you're going to tally down the column for the first Monte Carlo trial, and the second Monte Carlo trial, etcetera. One hundred million random number generations is a pretty big number. With some of your computers, that would take awhile. Some of you may have more lives. Even with a faster computer running, there are run-time problems.

I'll go back to one point—the typical company methodology arrived at the 100,000 entries in my example, one random number at a time. But it's not the only stochastically correct way to fill in the grid. If I have any sequence of independent trials, each with the same chance of success, there's a probability distribution function (PDF) for the trial number upon which I get my first success. For instance, I'm going to roll a single die until I get a six. The chance that it happens on my first roll is obviously one in six. The chance that it happens precisely on my second roll is five in six for missing it on the first roll times one in six for getting it on the second roll, etcetera. For the third roll it's five in six squared times one in six.

Each of my records reading across my first life, let's say has a 1 percent probability of death. I could fill in my row of 10,000 ones or zeroes by having each cell randomly generate a value between zero and one. If it generates a 0.01 or less, I put in a one. If it generates anything else, I put in a zero. The other way I can do it is to create a cumulative distribution function (CDF) for the geometric distribution.

So in other words, if it's a 0.01, if it's a 1 percent probability of mortality, then if I get a number less then 0.01, I do get a one in the first cell; otherwise I get a zero. But if I get something between 0.01 and almost 0.02, I fill in my first one in the second cell. If it's the CDF variable for the third, for getting a success on the third trial, I throw it in on the third cell, etcetera. I basically find out where my little N is that I drop in the distribution, and that's where I fill in my first one. I fill in N minus one zero and then a one.

Let's say I have 1 percent mortality anticipated for that cell. If I have 10,000 trials to fill in, then on average, I'm going to do it by drawing only 100 random numbers instead of 10,000 random numbers. So I'm going to speed things up by a factor of 100. I'm going to draw a new random number, in other words, each time I do an old random number and fill in a number of zeroes and then a one.

Clearly I do have to start my test over at each new row. So, if I come to the end of a row and I'm on zeroes, and it tells me to put a one somewhere off of the grid, I'm just going to forget that and start my test over at the new row. I can't run down to the next row and keep going because my mortality rate is going to change. It changes the whole distribution.

Once the grid is filled in, then the ones and zeroes are multiplied by the face amount and they're tallied down the columns, just as before. This is not an approximation. The stochastic implications of doing it this way are identical.

That concludes my thoughts on the different perspectives of X-factor certification. I'd like to talk a little bit about recent product development. For at least several months, and maybe stretching back to the beginning of this year, I haven't seen much. Frankly, the companies and other consultants that I've talked to have not seen many new ways to try to develop a product under XXX. The term products that I've seen that are being developed mostly have lower guarantees. There are some that are still putting up with surplus and going longer—sometimes with neat solutions involving reinsurance. But by and large, even if you look on an Internet Web site, which I did to look at term plans, most of them for a 30-year term are not going to be guaranteed for 30.

Most of the good quotes that you get are going to be with the "trust me" approach. Some do have other means of trying to dodge XXX reserves, or don't set up a humpback reserve, but the preferred guideline that we're going to talk about would severely impact those. Similarly for UL, there are ways within the product to try to get a lower reserve while setting up a longer guarantee. Shadow fund products, particularly for the very long guarantees, are the primary way of doing that.

So to summarize, XXX X-factor certification seems to be doing well across the industry. Practitioners are happy with what they've done, regulators are happy with what they've done, and there have not been many surprises. I've offered a few thoughts that I hope can help improve the process a bit for some of you, and

product development efforts appear to be more or less on hold, at least for now.

**MR. MICHAEL V. ECKMAN:** I'm going to talk about developing an actuarial guideline for XXX. I'm going to give some review here so that we all have the same terminology. Regulation XXX is more properly known as the Valuation of Life Insurance Policy Model Regulation. It defines statutory reserves for term and term-like products with guaranteed premiums.

Because there was concern with the adequacy of unitary reserves for these products, the regulation requires the greater of the unitary reserve and an alternative humpback reserve be held. There was also some concern that the UL model law might not provide adequate reserves for UL products with secondary guarantees. These secondary guarantees provide death benefits, even if the contract's account or cash value is negative as long as other conditions are met.

As I'll describe later, these other conditions may be based on premium payments or an alternative shadow or ghost account value. In addition to the basic reserve, Regulation XXX defines a minimum reserve. The difference between the two is the deficiency reserve. Because the valuation mortality table was known to be redundant, the X factors were developed so that a company could take its own experience into account, and Doug has addressed those issues.

The regulation became effective in many states on January 1, 2000. During 2000, the new Y2K problem began—how to apply XXX to products that appear to have features specifically designed to provide guarantees that would not be subject to XXX. In addition to product features, companies appeared to be using variations of reinsurance and combinations of contracts issued by more then one company to provide guarantees. For example, a life company would offer a term contract with no guarantee premiums. A sister casualty company, however, would then offer a contract that essentially guaranteed the life insurance company's premiums. Although shadow account UL products predate the regulation, the wording of the regulation did not give specific guidance on the reserving for this product. The guarantee provided by the shadow account is dependent on the level of a shadow or ghost account that is calculated only for the purpose of determining whether the guarantee is in effect. The regulation used wording that considered more traditional measures, such as premiums.

A UL contract that provides a secondary guarantee through the premium criteria may have a catch-up adjustment. For example, the contract holder may have to pay $10 a year to qualify for the ten-year guarantee. As long as the $10 per year is paid, the contract remains in force during the first 10 years, regardless of the amount of account value or cash value. But if the contract holder has paid only $8 per year for six years, and the contract is in force, a catch-up provision would allow them to pay the missing $12 from the first six years and qualify for the guarantee. As long as the contract is in in force, the contract holder can pay additional premiums to catch up to a guaranteed period. As is indicated by the length of time

I've spent on these various topics, I'll be concentrating on UL products.

The status of the regulation has prompted some to conclude that an actuarial guideline is required. First, not all the states have adopted the regulation. That by itself, because we've had this situation with other regulations, doesn't necessarily call for a guideline. But as I'll point out later, this in combination with some other developments has caused some to ask for a guideline. Second, some of the states that have adopted the regulation have used wording that differs from that in the model. This isn't unique to Regulation XXX, but when you add all these points together, some problems can develop. Third, since the regulation cannot cover every possible type of contract, the wording changes have led to different interpretations.

For example, the regulation doesn't mention the catch-up adjustment that I described earlier. And just how should reserves for those products be calculated? Finally, codification has added a little bit of confusion. Many of us saw an e-mail message earlier this year that implied that all states had, in fact, adopted XXX, because it was part of codification. There was discussion later that this was just an overstatement.

As we look forward, a company does have to report Regulation XXX reserves to the states that have adopted it. A company must disclose the difference, if any, between the reserves it holds and Regulation XXX reserves to a state that has adopted codification. To report these reserves and prepare the disclosure, of course, you're going to have to calculate the reserves. The products on which you calculate the reserves are going to have shadow accounts and catch-up provisions. Earlier this year the NAIC surveyed companies and regulators, and the findings indicated differences of opinion as to the applicability of XXX and how to apply it to various contacts. These differences of opinion by themselves may indicate the need for a guideline. To this end, the guideline dated July 12, 2001, was written. Basically it has six numbered points at the beginning that cover various new products that appear to be intended to avoid XXX reserves and two additional points, seven and eight, that deal with universal life products.

When we put together an actuarial guideline, we look for certain characteristics. The primary characteristic is that the guideline clarifies the application of a law or regulation. Of course, we want the guideline to reflect the intent of that law or regulation and to respond to developments that have occurred since the adoption of the law and regulation. The guideline, however, should not change the law and regulation—guidelines should not be used as a shortcut method to develop new laws or new regulations.

Sheldon is now going to talk about some of the considerations that went into the first six points of the July 12 proposal. After Sheldon finishes, Mike Smith is going to talk about shadow account products. Then I'll return to talk about catch-up products.

**MR. SUMMERS:** As Mike mentioned, there was a survey that was sent to regulators in October 2000. There was also a fact sheet that the Life and Health Actuarial Task Force (LHATF) sent in December 2000 to let the state insurance departments know about the issues that have come up with certain policy designs. The proposed guideline lists some of those products. I'm just going to go through a few of them.

The first one involves a policy that has an initial level-premium guarantee for ten years that's followed by increased guaranteed premiums. But the contract says that the company can't increase the premiums after the first ten years unless some specified event occurs. One example of this is if certain Treasury rates went below 3 percent, or something to that effect. After debating this, the LHATF concluded that the entire 30 years of level premiums should be considered as guaranteed because the ability of a company to raise premiums after the tenth year is limited.

The second policy type was a term policy that had illustrated level premiums for 30 years, the first ten of which were guaranteed. After the first ten years, if the company ever increased the premiums, a refund would be provided to the policyholder. The participants debated whether different calculations should be done. Should it be looked at as an endowment or considered an entire-period guarantee? The lesser of these calculations? The greater? Again, the conclusion was that the entire 30 years should be considered as guaranteed because if the premiums are raised, there's a consequence to the company.

Another example is a policy that has relatively high guaranteed premiums, but it also has either guaranteed dividends or guaranteed refunds scheduled; or, by some other means, guarantees of low net cost to the policyholder. In this case, the task force concluded that the net amount of premiums—in other words, the gross premiums less dividends or refunds—should be used in the reserve calculation.

Another example is a re-entry term product that has an initial rate guarantee for ten years, but there's loose or nonexistent re-entry underwriting, allowing the policyholder to re-enter for an additional 20 years at specified favorable rates. And another example is a universal life policy with a guarantee that a substitute policy will be issued in the event the policy value falls below zero but stipulated premiums had been paid.

That's really a continuation of the guarantee. So for both these examples, the initial reserve segment applicable to the original 30 years would apply if the guaranteed re-entry premium or the guaranteed premium for the substitute policy was not high enough to trigger a new reserve segment.

There have been a couple of comments regarding these six policy designs that have been incorporated in the proposed guideline. So these really haven't been controversial. But the seventh one deals with the design that has the catch-up provision, and the last one deals with the one with the shadow or ghost account.

Regarding those two, there has been some disagreement among regulators and the industry.

Fortunately, when I became moderator and needed to recruit additional speakers to complete our panel Mike Eckman and Mike Smith, who have voiced differing views on how the reserves should be calculated, agreed to participate. Mike Smith is now going to express his views, and Mike Eckman will follow.

**MR. MICHAEL S. SMITH:** Thanks, Sheldon. I'm going to talk about the issues surrounding Example 8, which are products that allow prefunding of guarantees in a UL contract. Shadow accounts are one type. There are other types that don't necessarily involve a shadow account but would simply take accumulation of the premium paid to date and compare that against the accumulation of the required.

I think this is unique among panel discussions at SOA meetings in that this proposed guideline is being developed as we speak. There's going to be a call in two weeks, I think, to discuss the direction of this actuarial guideline. One of the goals for Sheldon, Mike, and me on the panel was to encourage other members of our profession to get involved.

What I'll provide is a perspective that's probably slanted toward my view, but I need to distance myself a little bit from what Sheldon said. Also, I'm not presenting Lincoln's view here. I've tried to step back a bit and include all of the comments that have been provided by industry, and to a certain extent by regulators, as it relates to the guideline, as applicable to example eight. Those of you who have been following this carefully may hear me say things that I haven't said in previous meetings.

I'd like to provide a little background on shadow accounts. As I said before, there are other kinds of products that would be affected by the shadow accounts. I'm going to focus on shadow accounts, because they're what I'm most familiar with and they're what Lincoln has offered for several years now.

It's based on the UL product chassis. It looks like a normal UL, except that there's a memorandum account that keeps track of whether the policy is in force if the cash value becomes negative. That's what is called the shadow account or memorandum account. It functions essentially like an account value. Actual premiums paid are tracked, charges are deducted, and interest is credited, based on the terms of the contract.

The shadow account design provides a lot of flexibility for policyholders to pay on various schedules and still ensures that their coverage remains in effect. This flexibility, I think, drives the confusion over exactly how to reserve for this in an XXX context. The particular difficulty is that there's no specified premium at issue. And XXX was really designed, I think, to do that.

As Mike alluded, shadow account products were sold before XXX became effective. We started selling ours in 1998. They weren't specifically mentioned in the regulation. I've talked to folks who were involved in XXX's development, and they said they really didn't think about it or that it wasn't part of what they were contemplating, although there is language we'll talk about in a minute that you could interpret as being applicable to shadow account products. But the real question here is, what's the premium that's the source of the debate? XXX seems to anticipate that a specified premium will be in the product. There's no specific mention of what to do if you can prepay that specified premium and how that affects it.

In 1999, when Lincoln was looking at how to apply XXX to this product, my predecessor spent a lot of time working with the regulation, with the product folks, and talking with regulators. We have this regulation, and we have this product. How do they fit? We came up with a method that we think is consistent with what the regulation defined as a method to calculate a minimum premium, which then goes on to determine your deficiency reserves.

The regulation says—I'm not going to quote it verbatim—where a premium isn't specified, use the minimum premium that keeps the policy in force. Then there's further language in Section 7 of the model regulation that says the minimum premium is the premium you pay, so that if you start with a zero account value, you wind up with a zero account value. So that's the interpretation that we've read into XXX and what we think makes sense in this context.

The premiums that emerge are an ART scale, as you'd expect following the cost if insurance (COI) pattern predominantly. There have been arguments—and I think this happened mainly in the middle to end of last year—that the premiums should be leveled in some way. I think there was a consensus that emerged in the task force that XXX just didn't allow you to do that. So we have moved on to a different point in which we're trying to develop a guideline that doesn't level the premium but provides another way to reflect the accumulations of prepayments of guarantee.

The concern that I've heard from regulators and others, as it relates to XXX and how it's applied to these products is: Was an ART-type of minimum premium scale contemplated at the time that the regulation was developed and were the implications thoroughly considered? Does the language really apply? The answer in my mind is XXX clearly applies to shadow account products. I think the main concern is, does it create some sort of loophole? Could a company or a product be designed that allowed for a shadow account design? You could go back to low-low premium term, a 10- or 20-year term product, and avoid the humpback reserves. To the best of my knowledge, that isn't being done in the current market, perhaps because of the concern over where the guidelines are going. So these concerns lead to the guideline, as both Sheldon and Mike talked about.

As I said, this is real time, so the language is continuing to evolve. But basically, Example 8 says to add the shadow account value to your XXX basic reserve, and that's your new reserve. That is capped by a net single premium ceiling, which is calculated on a secondary guarantee basis. The concept there, I think, is that the shadow account represents excess funding that you need to hold as part of your reserve to be used later.

Now we'll get on to the summary of the concerns that have been raised with that example. I was able to detect six different themes throughout the concerns that have been raised by industry mainly. Some regulators have expressed some of these hesitations as well. I'll just handle each one in turn.

The first is what I call anomalous results of the method. I think of this as a test of the method at the extremes. Assume a UL policy has secondary guarantee features that are identical to the basic guarantee. I'll say it's a 4 percent interest rate and ADCSL mortality, your generic UL basic guarantee. We applied the logic in Example 8 to a policy with that structure. As I said, the secondary guarantee and the primary guarantee are identical. There's no additional economic benefit to the consumer provided by the secondary guarantee. But you get different answers if you apply the method. You get significantly higher reserves in the early duration if you apply this method. I think our feeling was, as I've learned over the years in actuarial work, that when you test things at the extremes, you learn when they break down. And our feeling was that this identifies some sort of basic flaw in the methodology that was being used, or at least it creates the sense that there's a problem here.

The next comment is that the method produces reserves that aren't consistent with the risks that are being assumed. At the margin, the secondary guarantee provides or creates economic risk for the company when the policy is not well funded. In other words, if a policy is well funded early on, in order for the secondary guarantee to come into effect, actual charges and credits would have to be less than the secondary guarantee level for an extended period of time. As that period of time gets longer the higher the policy value is in the early durations. What the proposal here does is the exact opposite. If a policy is well funded, it creates that much more additional reserve for a time, above and beyond the reserve you would hold for a product without a secondary guarantee. If it's not well funded, it would produce a reserve probably less than you would hold for a product without a secondary guarantee. So in the minds of many, there's sort of a risk mismatch.

Another objection that's been offered—and it's been offered at least in part as a possible explanation for those anomalous results that we saw earlier—is that there doesn't seem to be any commissioners reserve valuation method (CRVM) expense allowance recognition in the excess funding addition that you're putting into the reserves. I don't think anyone has figured out how to fix that or exactly what the problem is; there's kind of a sense, though, that it's not consistent with CRVM. There haven't been any publicly proposed methods to correct that so far.

Another objection was that this guideline was not consistent with the intent of XXX,, that XXX was intended to serve as a floor for UL in situations in which reserves were getting too low. The person who made this argument said you could see $\frac{1}{2}c_X$ reserves on a UL product at times that work into the UL model. The feeling was that by changing XXX in this way, it really sought to redefine CRVM for UL.

Now, I want to get back to Mike's point about what a guideline can do. The point could be made that XXX is clearly focused on calculations of reserves and premiums as of the date of issue. All of the definitions are set at issue; you do your segmentation as of the issue date, and all your reserves are known at issue if they're XXX. This is an issue that Sheldon has asked the task force to reconsider in its call in two weeks. The question is whether a guideline can recognize premiums that are paid after issue, a guideline that interprets XXX when there doesn't seem to be any allowance for that in the language of the regulation. That, I think, raises a basic question: Can guidelines go beyond what the regulation says? Is this beyond what the regulation says to do?

Finally, I think the last issue that was on my list was that of retroactive application. Guidelines typically are applied retroactively on the basis that they are indeed interpretations of the law. It's simply what the law said, and we're now clarifying exactly what you should do within the context of the law. Given the arguments that others have raised, I think there's at least a question as to whether this is interpretation or maybe more of a fix or a correction. If it's a correction or it doesn't seem to be within the scope of XXX, I think it's a legitimate question to ask, and it's been raised by many. Should this be applied retroactively? LHATF has offered that they would consider whether this should be applied prospectively only. There are precedents both ways. There's a guideline under consideration right now that is only prospective, but typically that's pretty unusual.

The arguments in favor of the proposal are that it's more consistent with the intent of other reserves methods. I think another way to say that is, it would close what's perceived to be a loophole that could be exploited. Also, this method does produce additional reserves, and it's possible to design a product with shadow accounts that doesn't produce additional reserves beyond what the UL model would already produce.

This debate has been going on since October 1999, at least. Through that period several other solutions have been offered. I'll just throw them out there so that you're aware of them. The first is what came to be known as Potter II, a second attempt at a clarifying guideline written by Bob Potter from North Carolina. The second version sought to very clearly stay within what XXX says to do and define the method that essentially I laid out in the very beginning of my presentation.

Another option, which Sheldon proposed at the beginning of this year, was that perhaps a review of the UL model regulation would be appropriate, that a secondary guarantee should be considered as part of the UL model. That had really

wide-ranging implications and I think LHATF has put that on the back burner for now. I'm not sure whether that will come up again in the future.

In conclusion, LHATF is continuing to review this, and it's very important for industry to comment on this and other issues. If you're seeing those guidelines for the first time and you have products that will be affected, there's very little time for you to comment. The next call is November 6; it's public for those of you who don't follow regulatory matters. LHATF conference calls are always public. You can dial in, listen, and you're often given an opportunity to comment. Then there's a meeting in December.

**MR. ECKMAN:** I'm going to talk primarily about the application of XXX to UL with a catch-up adjustment. I'll be repeating some of what Mike has said about shadow accounts as a way to contrast and compare. So I apologize for that ahead of time.

To review, the catch-up adjustment allows the policyholder to qualify for a secondary guarantee by paying additional premiums, so that the cumulative premiums paid to date attain a certain level. The regulation calls for the actuary to reserve for the product using the longest guarantee. But if a contract holder can catch up, what is the longest guarantee? Take, for example, a $10 premium for a 10-year guarantee or a $20 premium for a 20-year guarantee. As long as the $10 is paid, as I said before, the policyholder qualifies for the 10-year guarantee. But since the policy is in force, the contract holder could make up any missing premiums and qualify for the 20-year guarantee at any time. So in effect, the policyholder has the option for that 20-year guarantee.

In addition, if the guarantee is provided for by premium requirement, the contract holder could prepay the guarantee**.** My example is for a $400 single premium paid at issue or 20 years multiplied by $20. The contract holder would lock in at 20-year guarantee, as long as no surrenders or partial surrenders were made**.** So what does this mean for the reserve calculation? Of course, the UL model reserve will be higher than those for a contract paying the annual $20 premium, but will it be high enough for the guarantee provided? Does the regulation intend an additional reserve in this situation?

On the other hand, if we must reserve for the longest guarantee and the contract holder has paid less than that required to qualify for that guarantee, is the Regulation XXX reserve too high? After all, the contract holder will have to pay a high premium in one or more years in the future to catch up, to qualify for that longest guarantee. So, should Regulation XXX reserves be reduced by this future catch-up premium that has to be paid?

In applying XXX to UL with catch-up, we can consider some options. One is that we can consider the actual premium payment. Because the future premiums required to qualify for a secondary guarantee depend on what the contract holder has paid in the past, the reserving method could take prior premium payments into account.

Again, if a contract holder pays more than required, he could be considered to have extended his guarantee. This may be considered as changing the nature of the contract. I will assume that at most, the contract holder would guarantee coverage for life.

On the other hand, if the account value of the contract is low or even near zero, and the contract offers a catch-up, the contract holder would have to pay a large premium. In valuation, this could be recognized by adjusting the valuation premium similar to what is done for graded premium whole life, so that the present value future net premiums are high and the reserve is low. An alternative could be to subtract an amount representing the catch-up from the level of premium reserve. There should be a maximum reserve, and as I said, I'm assuming that the contract holder cannot purchase more then a single premium life contract, so that would be an effective maximum.

But as Mike has pointed out, we need to take into account a CRVM allowance. In a sense, it's still a premium-paying contract. If the contract is prepaid, then we can argue that there are no future premiums required and there can't be any future premium deficiencies. Therefore, there should be no deficiency reserve. Similarly, to the extent that prepayment reduces but does not eliminate the need for future premiums, the prepayment should first reduce any deficiency reserves.

Finally, there should be a minimum reserve. All of statutory valuation is based on minimum reserve standards, so we need a minimum reserve for the benefits provided. Mike has talked about shadow accounts and has pointed out that even though they existed before Regulation XXX became effective, there were no specific provisions made for them in the regulation. In particular, the valuation premium needs to be determined. One option for a valuation premium is to use a level premium just as would be done for a level premium specified premium product. This would require a contract that can be funded and the secondary guarantees qualified for using an increasing premium to bear the burden of level premium reserves. Since the shadow account is based on the development of an account value involving COIs, expense charges, and interest credits, the contract holder can meet the minimums using an increasing, ART-like premium. Using the option of the ART-like premium as the valuation premium produces the lowest reserves.

Just as the actual specified premium payments could be considered in the calculations reserve for the specified-premium product, the level of the shadow account could be considered as indicating the degree of pre-funding in the shadow account product. A large shadow account implies that the contract holder has prepaid for coverage, as premiums could be suspended, and the coverage and guarantee would remain in force as long as the shadow account remained above zero. Because the shadow account cannot be underfunded—that is, allowed to go below zero—there is no direct parallel with the requirement of the specified-premium contract. If you're told premiums to date drop below, you don't have the guarantee. If the specified-premium product has a catch-up option, however, the

contract could be kept in force by maintaining the account value or the cash value right above zero and then making that catch-up premium payment.

The two different contracts look a little bit more similar in the use of increasing valuation premiums because a shadow account product is really comparable to subtracting a catch-up premium from a stipulated premium's level-premium reserve. For the shadow account, I have the same three points that I had for the catch-up product: a maximum reserve, a principle added—if it's been prepaid, you should reduce deficiency reserves first—and a minimum reserve.

Despite the differences between the specified-premium product with the catch-up option and the shadow account product, the reserving should be based on similar principles. Ideally, the reserving would not favor one product over the other. Even though the shadow account product may not be advertised as having a catch-up provision, it actually has one that may be more liberal than the specified-premium product. At each contract anniversary, the shadow account product has a catch-up option that requires a relatively small premium. The specified-premium product requires making up all prior premium shortfalls, but the shadow account product requires only enough premium to keep the shadow account above zero. This can be considered a continuous catch-up, while the specified-premium product has a discreet catch up.

Several guidelines have been suggested. Bob Potter had proposed one early in 2000 and revised it later in the year. That was the Potter II that Mike referred to. The first proposal dealt with all types of UL and suggested a level-valuation premium. After objections that this went beyond the provisions of Regulation XXX, Bob revised the guideline to apply to shadow accounts only and using increasing valuation premium.

ING ReliaStar has offered a couple of actuarial guidelines. The first attempted to cover all products. The key to the guideline was to take actual premium payments into account in determining the future valuation premium. Although the guideline did not propose level-valuation premiums for shadow account products, it did propose the use of a level percentage of an increasing premium scale. There are opinions that this went beyond the provision of the regulation.

The second proposal is a further development of the one suggested by the LHATF. An LHATF sub-group, chaired by Sheldon, has considered the various proposals and developed a few of its own. One incorporated some of the UL Model Law into Regulation XXX, resulting in the calculation of four reserves—the UL model and XXX reserves, with and without X factors. Other proposals were somewhat simpler and included either an increase or reduction to the XXX reserve, determined by the degree of overfunding or the existing to the catch-up provision.

As Mike has pointed out, despite our efforts, there always seems to be a set of circumstances that indicates the proposed guideline produces inappropriate

reserves. Small secondary guarantees produce large reserves. Large premium payments increase the reserve, when it was argued that the prepayment of premiums actually reduced the risk and therefore, the reserve should be lower. Resolving any of these anomalies requires long and complicated calculations. Although the calculations are not impossible to solve, they would be difficult to implement.

A major criticism that has been voiced is that the guidelines go beyond the wording of Regulation XXX. In particular, the use of a level-valuation premium in consideration of actual premium payments in the determination of valuation premium has been criticized. One of the greatest fears regarding the guideline, however, is that the more specific it is, the more abuse it will allow and even encourage. Because we cannot anticipate every product design in the guideline, there's a fear that listing only some designs would imply that the others are not subject to the guideline.

I have a couple of brief comments on formula reserves. Formula reserves, in a sense, give us a safe harbor. One of those points in our actuarial opinion is calculated in accordance with the law. On the other hand, the more specific they become, the more abuse that is feared. In the end, however, we do have to test reserves for adequacy, and this implies a final check. My personal opinion is that Regulation XXX reserves are redundant. The reserves could be less than those required by the spirit of the guideline, but still be adequate. This could lead to an unlevel playing field.

After working with this for more than a year, I put down what I consider my practical priorities. We would like the guideline to (1) be consistent with the intent of Regulation XXX, (2) be uniform among all states, (3) be able to be administered, and (4) provide a level playing field. Frankly, I've moved the level playing field to the bottom of this list from the top, where it was originally, because I'm willing to sacrifice that for certainty as to how we should reserve for these products, even though one product design may be favored over another.

As Mike has pointed out, the next steps require industry input. Mike mentioned the phone calls and the meeting in December. LHATF may consider what we have to say and possibly incorporate it into a guideline that it develops. The NAIC will review that guideline once adopted by LHATF. Again, industry input will be necessary at that time, so we're going to continue to work on the guideline. But in the meantime, product development does continue. We have to do something. We're going to have to report statutory reserves at year-end. As the amount of UL subject to XXX grows, the possibility of surprises in the future grows. I will just say in conclusion, get involved.

**MR. SUMMERS:** I would like to mention that Mike Batte, representing New Mexico, is the chair of LHATF, and that I, representing California, am the vice-chair. Not only do I encourage questions, but I also encourage you to express any opinions

that you might have on this issue of whether the actual premiums should be recognized in the reserve calculation. I'll start off with a question for Doug. On your proposed method for shortening the Monte Carlo simulation run time, isn't it true that by reducing the number of random numbers involved, you increase the variability in the result?

**MR. ROBBINS:** Actually no, and the reason for that is that by using Harry Panjer's method or convolutions, although it takes a lot longer then Monte Carlo, you can develop the actual distribution of a set of potential claims for the anticipated mortality distribution. For the same reasons that we get practical results, the central limit theorem tells us that as you take a sample from a distribution, that the distribution of the mean approaches a normal distribution. It's just distribution theory and the fact that a large set of trials, such as those you use to develop the matrix of ones and zeroes, does in fact approach the distribution of the first one of them to be a success. That's the same thing as the geometric distribution. So the stochastic outcome is identical.

**FROM THE FLOOR:** I know that you haven't even solved the current round of problems, but looking ahead to the 2001 CSO, what are considered to be smaller margins? Will we still have X factors?

**MR. SUMMERS:** I think we will, because you could still have different underwriting classes. I think that at an LHATF meeting, it was decided that the 20 percent would still apply as the limitation.