Informal Discussion Transcript
Concurrent Session 3B: Comparing Mortality of Different Groups

Presented at the Living to 100 Symposium
Orlando, Fla.
January 4–6, 2017

JOSEPH LU: Thank you, Kai, for your comprehensive review and thought-provoking questions, and I really thank the three authors as well for your very intellectually rigorous papers. For each of them, it must have taken several person-years to write them, and we get exposed to them in just a short several minutes. Yeah, let's give them a clap and thank them please. [*Applause*]

You can't relax yet, because Kai has asked each of you some questions. Can I have a show of hands if any of you have any further questions? Right—one, two—right. I'll let them answer Kai's questions, and we'll come back to you. Can each of you take five minutes to answer Kai's questions, please, if that's all right? Thank you. Kenneth.

KENNETH ZHOU: When we estimate the model, we did use very a long period of data from the EW [England and Wales] population. Actually, we used 90 years of data from 1921 to 2010 or 2011. I believe, if we chop off the very first part—say, we chop off 40 years or 50 years—then the model won't show up well fitted, and the parameters of the model won't show up significant. That means vega won't be defined, because there is no GARCH [generalized autoregressive conditional heteroskedasticity] effect or ARCH [autoregressive conditional heteroskedasticity] effect. The parameters *a* and *b* are close to 0, or they don't exist at all. But nevertheless, delta and gamma can still be defined, because they are based on the period effect, and vega won't be defined. If we use a very short period of time—let's say 40 years—then there are two ways to look at this. One way is saying that there is definitely no GARCH effect in the data we observed, or the other way, we can say that maybe we haven't observed the data long enough to actually see the GARCH effect in it. Is it one way or the other; we don't know.

There are also other papers showing that there are GARCH effects in many national populations, and I think some of you should have seen those papers or have heard of papers saying that the GARCH effect does exist in many populations. So it depends on how you look at this. It's really do you believe that there is GARCH effect, and do you want to use vega as one of your hedging methods. If yes, then you may have to consider a longer period of data, or the other way, just use the short data, and believe that there's no GARCH effect in it at all.

For the other question, it's talking about the Lee-Carter model, which our method is based on. The method is a model-based hedging strategy. That means it's based on a model, and as mentioned by Kai, it's possible to use a non-model-based hedging strategy, where you calculate the sensitivity with respect to the underlying mortality rates, like the central mortality rate or the probability of death. There are papers showing that it's possible to do it. For example, I think Johnny Li proposed a method they call key q-duration, where they calculated the sensitivity with respect to the probability of death, q. A problem with doing that is it requires many q-forwards to achieve a good hedge. I think in their paper, they use five q-forwards, and the benefit of using a model is that you can simplify the problem of having many different mortality rates for different ages and different time points to a single, one-variable underlying value. So in our method, only one q-forward is required to do a delta hedge, and in an early market, I think having one q-forward is more likely than to have five q-forwards at the same time. That's the benefit of doing this.

This paper or this method is demonstrated on the Lee-Carter model, and of course, I think it can be done using any other models that contain period effects, because the method is still based on taking the sensitivity with respect to underlying period effects. With the CBD [Cairns-Blake-Dowd] model, of course, I think it's possible to do it, and there might be other benefits to it, because it contains two period effects, and we can have many different kinds of combinations of Greeks to develop the hedge.

Yes, that will be all my response, and I think I have my time.

JACK YUE: First of all, I apologize for my paper. I think the paper was written in a very short time, so I didn't put a lot of details inside, so I apologize. For the slides, I add some details. For example, the first question regarding a wider Lee-Carter model, you would have biased estimates. For example, here. Sorry. I guess because of—for the Lee-Carter model, first of all, I need to mention the biased estimates always occurred at middle age. The data considered are for ages ranging between 0 and 84, for both male and female. The larger bias appears at the middle ages.

To me, I think this has to do with the setting of the Lee-Carter model. It's just, the Lee-Carter model to me is just like a bunch of linear equations for different ages, and we try to combine different—the regression question for different ages, using the same regressor (time). So this is a very strong assumption.

I think that for the regression models, the center points and their estimates usually would have the smallest variances. But, because the setting of the Lee-Carter model is similar to the regression model, I expect that the estimates of the Lee-Carter model would behave similarly. I will check if my assumption is correct and probably add more results, too. I'll modify the results here. This is the first question.

The second question is about the CBD model. I mentioned that the CBD model doesn't suffer from the small population size. To me, the Lee-Carter model is a bunch of different linear equations, and we try to combine them together. The CBD model is a little bit different. It is not a relational model, and we use different ages of mortality rates to fit for

$K_t^{(1)}$ and $K_t^{(2)}$. Given the $K_t^{(1)}$ and $K_t^{(2)}$, it is just like solving linear equations for different ages at the same year, and thus it belongs to the group of relational models. So I guess this is one of the reasons that it is pretty stable, and this is my explanation for the CBD model being quite stable.

The third question is about why choosing different graduation methods. I think first of all, the spline method, for example, has been used quite intensely by the U.K., Australia and New Zealand. For all these models, I think that they are all model dependent. In this study, I tried to avoid the influence of model selection, and tried—Kai also mentioned that—to keep the method as simple as possible. I don't want to use a strange—maybe I should say, or a more complicated—graduation method—for example, in addition to spline, another method such as the functional smoothing method, [which] appeared in 2007, and quite a lot of, I think, researchers choose this type of model. But to me, I would choose more intuitive methods like life table construction, which most countries are still using. I mentioned that the spline method is used by the U.K. nations, and most countries are still using very simple methods, such as moving average. I think for most Asian countries, they're still using some special version of moving average. For example, Greville's formula you can check on the web. It was founded in 1970 but is still popular in Asian countries. For the Canadian and U.S. life tables, they would choose Whittaker and something, for example, like the kernel graduation method. In general, they use a very simple formula to avoid the problem of model dependence.

So I will go back and check, and see if different graduation methods would matter. And I agree with Kai that the model with fewer parameters is more attractable, so this is also my philosophy. I tried to use a simple model, as I don't want to use a complicated model. Thank you.

MARTIN GENZ: First of all, Kai, thank you for your discussion and for the flowers. Well, of course, it would be interesting to find proof or data issues with my means of our framework, but in this application, we have focused on—well, we just used data where we have no indication for some data issues, and did look this up in the data protocol for the HMD [Human Mortality Database]. Where they alerted to bad data, we did cut that time period off. So, for example, in the Sweden example, I showed that we cut it off from 1751 or something like that the data starts, and we just took the data off to 1861. So because of the time period between 1750 and 1860, it was some buggy data in the collection in that time, and so we did not use this data.

You mentioned the example from Russia, the Eastern European population where we saw this outlying thing, and I don't know exactly which picture you meant. Was it this picture here?

KAI KAUFHOLD: That picture shows up in every one.

MARTIN GENZ: Yeah, of course. Well, this might also be a data issue. Of course, after 1990, there was a change of the system in this country, and the data collection system also changed probably. We observed some catch-up effects in this country, so the life, especially life expectancy, for example, was lower before 1990 compared to the Western European countries, and they caught up after 1990. So, of course, there was a change, and this was an eventual change.

What was an issue in our application, in our methods is this thing here. We have a fit of our time series, and we did that under the assumption of steady time series, and we detected a little bit later, after the paper was already written, we detected that there are jumps in the data, especially for these Eastern European populations. In 1990, there were some significant jumps downward, upward, whatever which we cannot fit with a continuous line. So we have to do a combination between continuous fitting where we have real, say, some trend changes, which is something like that where we have increasing and a little less increasing trends, but we have also to pay attention to some jumps in the data, and this is what we're still researching in this method here. We are searching, as I said, for a fitting method which can combine these two issues. That's it.

JOSEPH LU: Thank you. We'd like to invite people for questions. There are some mikes there. Would you please use them and briefly introduce yourselves, so that our speakers can reply to you.

MICHEL FOURNIER: So my question is for Kenneth. Sorry. My name is Michel Fournier from National Life. If you think of delta hedging as an option in the financial markets, you would be going in and out of the stock or futures. So my question goes to Kai's comment to some extent about the underlying: How do you view your delta hedging of annuities or pension plans from a practical approach? How could this be implemented?

KENNETH ZHOU: Yes. In a financial market, you will trade a delta quite frequently. That will turn into a dynamic hedge, where in this paper, it is a static hedge. We purchased a *q*-forward at time 0, and we keep it at that point and then see how it performs at the end of the hedge period. It's a one-time purchase. So at time 0, we purchase some *q*-forwards—actually one *q*-forward, that will be enough—and we don't have to go into a market again. If you want to go into it again, then it would turn into a dynamic hedge, and there are papers working on how dynamic hedging could be possible with delta hedging as well.

MICHEL FOURNIER: So what would be the underlying in that case?

KENNETH ZHOU: Oh, you mean there is no underlying asset, right? There is no underlying in that case because the underlying is not observable. We cannot accurately trade the underlying. That's a problem with it, yes. So we can only develop the hedge at time 0 and then purchase the q-forwards, and that's it. You try to hedge at time 0, and that's it.

MICHEL FOURNIER: So in financial markets, traders were practically doing delta hedging on the underlying without really knowing how delta was computed. Whereas in this case, it's the opposite: We're defining delta before knowing the underlying.

KENNETH ZHOU: Yes.

MICHEL FOURNIER: Okay, thanks.

TOM GETZEN: Hi. I'm Tom Getzen from Temple University, and I thought these were fascinating presentations, but I'm assuming that we're all trying to do the same thing, which is predict future mortality, but in different situations. And one of the questions that links all these three papers is, When can I use this data to help me do this? So I wanted to ask Martin. First, just a technical question: When you did the polynomial trend, how do we determine the degree of polynomial? I would look at something like that, and I would go, "Hmm, maybe I see a trend here and then another trend here, and then I see a couple of things here, one of which coincides with World War II, and I'm not sure I'd want to use that for prediction in the future. It's got some kind of a little funny kind of a characteristic there, and that really affects all these things." Where do we say this period is so different from this, it does become a classification problem? Like when Kenneth was talking about when we take one set of models, if we cut off the data between 1950, the difference disappears, and the apparent relationships are radically different. I wanted to ask you, particularly Martin, do we have real differences here, or can we characterize them in different ways? You said that the Eastern Europeans fall into a particular group. It would also seem to me that things before 1960 or 1950 fall into almost a different group. My last comment is that you might want to be cautious about using the word "worldwide evolution." Okay, it's a limitation of the HMD, but if we throw out China, India, Africa, and Latin America, it gets a little tough to call it worldwide.

MARTIN GENZ: Thank you for your comment. To answer the technical question, we used a method which was proposed by Muggeo, which is actually for continuous fitting of such time series. If you look in the paper, and the reference there, the paper is cited, so if you're interested in that, I recommend you read it. It's actually quite helpful. Of course, cutting off the data would lead to slightly different results, because, for example, if you cut off, say, two years before a trend change in our example here, the trend change would not be apparent. That is, the slope off the following subsequent period would be a slightly different trait, and this is what I also mentioned on the comment of Kai. This is also a problem with finding trend changes or some outliers, and this is always a bit arbitrarily you know, that you have to choose significance levels, you have to choose optimum criteria and so on, and of course, yeah, you're right, in a certain sense it's arbitrary .

KAI KAUFHOLD: I just wanted to maybe clarify one thing. It's a polygonal fit, not a polynomial fit. So you've got just straight lines stuck together. So it's zigzagging, and I assume you can't use it for a projection; it's just to classify what happened, where your comment is absolutely right. So you can't just go, "Oh, it will keep on going like that."

MARTIN GENZ: Actually, some of our collaborators are researching on a model to it to protect things like that, but that's all open research.

KAI KAUFHOLD: It's not that easy, Martin.

MARTIN GENZ: Not that easy, yeah.

DAN SHUMAN: Dan Shuman from Nationwide. I'm just really curious about something. In a lot of the presentations over the past couple days, it's led off with Swedish mortality experience. It seems to be a common example, and I'm wondering if that's just a coincidence, or if there's a reason for that. Why is Swedish mortality often used as an example?

JACK YUE: Actually it was—well, also I've been trying to choose Sweden, but it's one of the few populations which has considerably long data history in the HMD [Human Mortality Database]. In addition to Sweden, perhaps France and Switzerland [and a few other countries] are good alternatives for longer data history. But Sweden is one of the longest and most reliable data in the HMD.

DAN SHUMAN: Okay, that makes sense, thanks.

DAVID WATKINS: David Watkins, Rothesay Life. My question is for Martin. So is there a common path followed by the four statistics relative to each other, and does this give us a clue for the less developed countries as to which of those statistics might be developing faster—say, for example, if they lie behind on the common path?

MARTIN GENZ: So the question to me, and I did not understand anything, since it was a little bit low volume.

DAVID WATKINS: I asked, is there a common path followed by the four statistics relative to each other, and could we use this as a clue as to which of those four statistics might be developing faster in a country with higher mortality rates, by looking at a country with low mortality rates?

MARTIN GENZ: Interesting question. I can only guess. We did not really look at this

JOCHEN RUSS: Jochen Russ. I am one of the coauthors of Martin's previous paper, where the framework was being suggested. Since Martin didn't answer the first part of your question: A clear no—you cannot say that one of the four components depends on the other three. That's why we used four components in the first place, and that's why we claim that it's not sufficient to only talk about expansion or to only talk about rectangularization, because it's four different things. They may happen at the same time; they may not happen at the same time. But they are definitely not mutually exclusive. That's why you need to look at all of them to get a full picture of what's going on.