Session 3A Discussant Comments

Johnny Li

Presented at the Living to 100 Symposium

Orlando, Fla.

January 8–10, 2014

JOHNNY LI: First of all, I would like to say that I truly enjoyed reading the three papers and the presentations. I learned a great deal from the presenters. It's certainly my pleasure to be the discussant for this session. I am going to divide roughly my discussion into three parts. First of all, I'm going to discuss the papers on old age mortality, the paper by Natalia [Gavrilova] and the paper by Nadine [Ouellette], and then I'm going to discuss the paper by Matthias [Borger], which is on the coherent mortality projections for dependent populations. Then I'm going to look at a bigger picture, trying to draw relationships between these three papers.

First of all, discussion on the paper by Natalia. Well, in this paper the authors considered data from three sources. They are the Human Mortality Database, the U.S. Social Security Administration's Death Master File and also the railroad retirees. I think the last data source was downplayed in the presentation, but it was mentioned in the paper. The authors firstly compare the Gompertz and Kannisto models, and they also consider the life table aging rate, and on the basis of these measures and the datasets that they considered, they conclude that mortality deceleration is measurable up to age 106, and beyond that, because there is a lack of data, they didn't draw any conclusion.

In the paper by Nadine, they considered a dataset on French Canadian centenarians. The dataset was meticulously validated. In terms of methodologies, they calculated the observed death rates among these centenarians for various age intervals, and they also repeat the same thing for different groups of birth cohorts. They came up with a rather different conclusion. They conclude that our mortality deceleration is observed for the age range that they considered, and the observation is still true no matter if we consider different age intervals or different groups of birth cohorts.

So the conclusions in these two papers are somewhat contradictory. Nadine mentioned that this may be due to a methodological issue, but it may also be due to the differences in the age ranges that the two authors have considered. I recall that Nadine considered an age range of 100 to 115, while the other author considered an age range of 80 to 106, but the differences in age ranges may lead to a rather different conclusion, and of course, the other reason may be due to the difference in data size. One is from Canada and the others are drawn from the U.S. So at this stage I'm not too sure yet that it's solely because of a methodological issue or it is somewhat due to the data issues as well. Maybe the presenters can elaborate a bit more on these two points.

Now, I’m not going to take any side of the debate. (Laughter) So to avoid taking any side of the debate, I would like to instead draw some connections between these two papers to actuarial practice, to real actuarial work.

In practice it may be true that there are not many individuals in our portfolios who can survive to ages beyond 100, but then if they do, then say pension payouts, annuity payouts, to these people can be large, so I would consider our old age mortality as a source for what we call low frequency and high severity losses and that constitutes a risk, and because of this it matters to capital requirements and is something that we need to pay attention to.

Old age mortality is of course particularly important to pension derisking businesses that include, for example, longevity swaps, pension buy-ins and pension buy-outs. For these businesses, the portfolios are mostly concentrated on people at higher ages. So the two papers are actually very relevant to real actuarial practice.

Now, I am going to further connect the two papers to practice by using this diagram. Now, I think the two papers are mostly focused on the upper two rectangular data analysis. Now, this stuff is important and is nontrivial because in our own portfolio there are not many survivors to age 100, so we need to rely on larger scale meticulously verified datasets to draw patterns regarding old age mortality, so the two papers presented by Nadine and Natalia are crucially important to us.

To make data analysis useful, we need probability modeling. We heard some in today's presentations that includes, for example, the Gompertz curves and the Kannisto model, and then after probability modeling, we would like to take the results to actuarial practice. In particular we need to know values of q_x, so that we can fit the values of q_x in the systems for valuations and pricing purposes. In the rest of my discussion I am going to add a bit more on the lower two rectangles.

I would like to use a slightly different, a more actuarial angle. Now, in the presentations, the authors are focused on the force of mortality, hazard rates and central death rates but in practice as a natural, we care about q_x. Our systems are based on q_x, so I would like to change the lingo a little bit to a more actuarial q_x language, and I would like to borrow a few paragraphs from a paper by Roger Thatcher published in the *Journal of the Royal Statistical Society: Series A* in 1999. So in that paper Roger Thatcher said that in general there are three hypothesis or three schools of thought about q_x. The first one is that q_x tends to one as x tends to infinity, and this corresponds to a situation where there is no deceleration in the force of mortality. The other school of thought is that q_x tends to a value that is slightly less than one as x tends to infinity, and this corresponds to a situation where there is a deceleration in the force of mortality.

Finally, in the last hypothesis, q_x is precisely one at a finite age. Now, some people don’t like these hypotheses, because they say if we say that q_x equals one at age 115,that means that when a person has attained age 115, we are sure that the person will die within one year. So that doesn’t seem to be very intuitive. Nevertheless, the assumption of q_x equals one, i.e. closing the life table at a limiting age omega, is quite commonly used in actuarial practice.

So I am going to add a bit more technical stuff to my presentation. I would like to see how we may borrow extreme value theory in the analysis of old age data. In particular, I would like to take an important theorem in extreme value theory. In this theorem, it says that the excess over a threshold must follow in general a generalized Pareto distribution as the threshold tends to infinity. In a mortality context, it means the start lifetime random variable must follow the generalized Pareto distribution as we move on to very advanced ages.

So what does that mean to actuarial practice? Well, as a matter of fact, the generalized Pareto distribution is very flexible, and it encompasses the three schools of thought regarding q_x. In particular, if we look at the distribution function again, it contains two parameters gamma and theta and different values of gamma imply different behaviors of q_x. In particular you have the estimated parameter gamma as greater than one. That implies that q_x tends to one and x tends to infinity. If gamma equals zero, then q_x would tend to a value that is in between zero and one, and finally, if gamma is less than zero, the distribution has a finite right-hand support, which means that q_x reaches one at a finite age. So by using the generalized Pareto distribution and extreme value theory, we can actually know quite a lot about how q_x will behave as x tends to infinity, and I would be interested to see what the results would look like if extreme value theory is applied to the datasets considered by Nadine and Natalia.

That's all I wanted to say about the papers on old age mortality, and next I'm going to move on to the paper by Matthias on coherent mortality projections for two related populations. That is a highly important paper. I think it gives us a very good demonstration on the balance between statistical rigor and expert judgment.

There are a few good things that I would like to reiterate. The model that they use is the age-period-cohort

(APC) model. It's slightly different from the usual APC model. In particular, the authors apply the APC model to the mortality reduction factor instead of the mortality rates themselves. I think that is quite original, so credit must be given to the authors, and there are additional statistical features including, for example, alternative distributions or assumptions including Student's t-distribution and lognormal. The use of alternative distributions are often ignored by researchers, and in this paper, this issue is explicitly addressed by the authors. Another issue is parameters moving may help us to obtain a more biological reasonable pattern of parameters, and these models are successfully applied to the populations of German males and German females. So again, this paper is original and highly important.

I think the presenter focused mainly on the methodological issues. So I would like to add how Matthias' work is linked to real actuarial practice. Two coherent population mortality models are useful in many aspects. First of all, we can, as the author demonstrates, use a two-population mortality model to model male and female mortality rates jointly. This is particularly important in jurisdictions in which gender neutral pricing is enforced. It is also useful in situations where an insurer or a pension plan sponsor wants to project the mortality experience of their own portfolio but at the same time examine data of a larger population like a national population. So a two-population model allows us to model the mortality of our own portfolio and a bigger population simultaneously.

Lastly, a two-population mortality model by design is very suitable for use in the quantification of population basis risk in longevity hedges. Population basis risk arises in situations when, for example, if a pension plan uses a security like a q-forward that links to the general U.S. population but then there are possible discrepancies between the mortality of the general population and the hedges own population.

Population basis risk can be quantified by a two-population mortality model that is fitted to the general population and the population being hedged. Two-population mortality modeling is a rather important topic in recent years. I would like to take this opportunity to review some of the recent developments in this area.

First of all, there are several models that incorporate cohort effects in a two-population setting. That includes the two-population age period cohort model by Andrew Cairns in 2011. This model is somewhat different from Matthias' model in that it is fitted to the mortality rates themselves, rather than the mortality reduction rates, and then the gravity model by Kevin Dowd and his coauthors. Now, in this model there is a gravitational force between a bigger population and a smaller population, so if there is a divergence between the mortality trajectories of the two populations, there exists a force that brings the mortality of the smaller population closer to that of the larger population, and then there is a generalized two-population version of the generalized CBD model that is recently developed by myself and my co-authors. So in this model it consists of a cohort effect that is applicable to both populations that are being considered. Then there are models that are based on advanced time series projection techniques. For example, two groups of authors conceded a special factor error correction models for projecting the period and cohort effects encompassed in a two-population model. The factor error correction model is adapted in such a way that the mortality rates of two populations under consideration were not diverged in the long run, and then there are models with jump effects that are due to, for example, wars and pandemics in the paper by Zhou et al. (2013). A two-population Lee-Carter model with trajectory jump effects are developed. They use a multinumber approach in which different possibilities are taken care of. For example, a jump may affect one population but not the other. A jump may affect both populations at the same time, at the same severity or at different severities.

Finally, I would like to draw a bigger picture trying to relate the three papers together. In my opinion, I think the three papers are nice complements of one another. In the model by Matthias, the model applies to a full age range, but the age range is finite. I recall that they apply the model to ages 0 to 100, but the model tells us nothing about mortality beyond age 100. So the mortality patterns for ages beyond 100 may be obtained from the patterns that are described in the other two studies.

In the other way around, a two-population set up would be useful for us to understand the relationships between U.S. and Canadian old age mortality, so it would be interesting to apply methods similar to the two-population model proposed by Matthias, to the datasets considered by Natalia and Nadine. So I think that is the end of my discussion.