

ABSTRACT OF THE DISCUSSION

Mr R. Frankland, F.I.A. (introducing the paper): It was about fifteen months ago that we were asked by the Life Practice Executive Committee to consider whether we could attempt to assess what sort of market movements could be demonstrated to be consistent with a test described as being based on 1 in 200 year events in the context of the ICAS regime.

I am not intending to summarise the paper or its conclusions but did want to make a few remarks about the timing of the paper.

There are those who might feel it appropriate that we should embark upon that work in time to report now, when some commentators believe we are experiencing just such a 1 in 200 year event.

Alas, the reality is that times of turmoil, or apparent turmoil, are not the times when it is easiest to consider, rationally, analysis based on long term experience. Also, our initial focus on movements in equity values and fixed interest rates to the exclusion of credit risk and market correlations, preferring to leave those other elements until we had established a suitable methodology in markets with better documented history, will prove disappointing to those in search of a topical paper explaining precisely how we managed to get into the mess we are in at the moment. However, of course, the reality is that today we need to deal with market conditions as they are, and knowing that the events of the last 12 months were more or less than one half per cent likely to happen is not really much use.

Further, any information which we could have derived about movements in credit spreads from data available to us over most of that period would probably not be considered particularly useful going forward. However, a number of members of the Working Party have indicated a desire to work together to explore credit further, along with certain other aspects of extreme event analysis.

Within the ICAS regime there is apparently considerable scope for individual companies to set their own interpretation of the economic scenarios which correspond to a 1 in 200 year event which define the level of tests which they perform to determine their capital requirements. However, much of this apparent freedom is removed by the ability of the regulator to provide guidance to companies on how its minimum requirements are to be met. This restricts companies' ability to minimise their capital needs by resorting to the most optimistic model relative to its own risks.

The original hope was that this Working Party could produce results which would input into any process of standardisation, whether that was to be achieved through more open FSA guidance, through BAS standards or through an IAN issued by the profession. In terms of the results of the work, the outcome has been somewhat different to that which was expected, but following discussions with the Life Research Committee it was agreed that it would be beneficial to the profession if the results of the research were shared, nonetheless.

The key issue highlighted in the paper, in relation to the work on equities, is that, fundamentally, there is not enough relevant data to derive probability distributions from historic data relying on a pure frequentist approach. If we had, say, 6,000 years of homogeneous, non-overlapping, year-on-year market movement data, one might deduce that the 30th worst fall represented the one in 200 year event. The key point is that if we are looking at year on year movements in markets then you need many years of data to estimate that 1 in 200 year event. This may seem obvious, but sadly many people have claimed to find solutions to the problem in the form of fitting assumed probability distributions. We have tried to pursue as many of those as reasonably possible and indeed most have worked, as shown by the relatively small 95% confidence intervals around the parameters. However, each solution appears to yield a different answer with non-overlapping confidence intervals.

We do identify other possible approaches including relying on "expert judgement" or "market prices". However, it should be noted that both of these take historic experience and use some form of "prior belief", albeit a more subjective one, to derive a distribution.

In effect our end conclusion, at least in relation to equities, was that, based on recent historical data, there is a seemingly inevitable need for the researcher to impose some sort of prior belief on the observed data to arrive at an answer, and that the choice of prior belief is critical to the result. Further, whatever prior beliefs are adopted, it does appear that there is an over-reliance on the magnitude of the one or two observed worst case events which dominate the results of the analysis. Hence from a practical point of view it appears that, in answering the 1 in 200 year question for equities, the key question often comes down to a case of “how do you treat 1974?” This is at the heart of my nervousness about presenting such a paper today. How will 2008 affect our view as to what is a 1 in 200 year event? What will our future prior beliefs say about 2008?

A further issue that arises following a large movement in markets, which is of particular relevance today, is that of the relationship between conditional and unconditional probabilities. In other words, though a fall of $x\%$ in equity values over a twelve month time period may be equivalent to an unconditional probability of 1 in 200, what is the fall corresponding to a 1 in 200 probability given we have just experienced a rise or fall of $y\%$ over the previous year? In essence, our conclusion is that given there is insufficient data to derive unconditional 1 in 200 year falls without imposing prior beliefs, then there is certainly insufficient data to derive conditional probabilities and, consequently, the role of prior beliefs would be even stronger.

Many would argue that what is needed here is a pragmatic approach. We would agree, but would suggest that there is no consensus, actuarial or otherwise, on what pragmatic equates to in terms of market fall. We touch on this point in section 2.4.3 of the paper. In practice it may be desirable for the Regulator to temporarily reduce the probability level of survival for a period to allow companies to rebuild capital.

On one thing the working party is agreed: it is undesirable for individual entities to be able to interpret the 1 in 200 year standard in different ways and so to effectively set their own minimum capital requirements. However, we believe that we have demonstrated that tying standard levels of a test to a specific probability level is only achievable with the imposition of significant, and probably material, prior beliefs.

As a young actuarial trainee many years ago I remember hearing a general insurance actuary ask how you put a price on insuring the launch of a satellite when following just a few years of claim free experience the last two launches had resulted in failure and consequential enormous claims. I fear that, potentially, we may have a very similar problem in assimilating 2008 into our experience.

Turning to the fixed interest part of the paper, our main area of focus has been to explore possible solutions of how to model yield curve movements in a way that does not open up the possibility of hiding the imperfections of a simplified approach to duration matching. Data imperfections involved in the production of historic yield curves mean that the sort of detailed analysis we have been able to do on equity movements have, at best, been more difficult to achieve, and we have made limited progress in that area. Again, this is possibly an area that, if the profession wishes, members of the working party may be prepared to take forward.

Before handing over to general discussion, I would like to make the point that although our paper is clearly focused on the needs of the ICA regime, the methodologies and analyses have application elsewhere in risk assessment, including potentially in Solvency II, as well as in the assessment of market risks outside of life insurance.

Mr A. N. Hitchcox, F.I.A.: I liked this paper, not just because it gives us great insights into the long term behaviour of financial markets, but also because it reminds us what a good piece of actuarial work should look like: it avoids making too many model assumptions, instead it found out what the data itself is telling us; and it lays out enough of its workings so that the independent reader could form his own view of its conclusions. In addition, I am, personally, a great believer in using graphs to make points, and not just showing tables of figures.

Financial economic theories are good at:

- modelling near the centre of the distribution; and
- modelling smooth behaviour.

They use data series based upon 200 or 500 trading days, but do not seem to look at data series based upon 50 or a 100 years of price levels.

As soon as you examine the longer term data, you have to allow for:

- fat tails;
- jumps in price levels, i.e. you cannot assume continuous functions; and
- clustering of volatility.

When you insist on market consistent approaches, you must model the behaviour of markets:

- markets are not physical systems, they are people systems, driven by beliefs and anticipations as well as actuality;
- they demonstrate turbulent and rough behaviour; and
- the fractal modelling ideas of Mandelbrot can contribute to understanding of these factors.

Insurance companies have developed a lot of expertise in the modelling of rare events, i.e. those that can contribute to risk measures such as:

- 99.5% VaR (ICA level);
- maybe even 99.9% VaR (A-rating level), 99.97% VaR (AA-rating level), and 99.99% VaR (AAA-rating level); and
- 99% TVaR: although apparently a nearer term risk measure, because only 1/100, not 1/200 or 1/1,000, it does, of course, need to take the average of all rare events with a return period of 1/100 and upwards.

One of our chief jobs as actuaries is to use Economic Scenario Generators (ESGs) in our ICA workings, to give advice on the 1 in 200 outcomes of our insurance companies.

ESGs are good at:

- linking assets and liabilities together;
- modelling duration risk and the net unhedged systematic market risk in the balance sheet; and
- modelling the embedded options and guarantees that we have sold.

But we must now take them to the next step:

- they have to start allowing for the turbulent behaviour that is such a feature of modern global markets; and
- fat tails, price jumps, investment bubbles, clustering of volatility: these must all become the regular language of the outputs when we are advising Boards of directors.

My challenge to the authors of this paper is as follows. You have done the job you set out to do. You have been data driven, with no prior model assumptions. Now you must move on to the more difficult jobs: the so-called intractable problems, harder to model, often with less data than you would like. Yes, you will likely have to start making model assumptions again. But these are the real value added areas of work. They represent the true cost of rare tail risk, apparently sometimes overlooked in some of the banking world's financial models. Do not be afraid to say that it is "speculative modelling", to distinguish it from the more soundly based data driven work.

The challenge is for you to bring together three things. First, the best insights of financial economic explanations, for example, setting the discount rates for inflation and equity risk by reference to external prices; and understanding that one of the chief concerns of the investors who own our shares is the correlation with the rest of their portfolio.

Second, add to this our own insurance and actuarial modelling insights, into for example Dynamic Financial Analysis (DFA) modelling, Asset Liability Modelling (ALM), and the modelling of rare tail events over long time horizons.

Third, add onto it the newly-emerging insights into the modelling of turbulent market behaviour, as mentioned before, and, when you combine all of these features, great advances can be made.

Remember that models are not just predictors of future outcomes; they are also stress testers to apply to different strategies, to make sure that you are not under-capitalised for the rare events.

Regarding model risk itself, if you meet a problem with your models, you can respond either by making them more and more complex, i.e. more difficult to understand, or you can be more sophisticated by changing the way you use the models. More sophisticated modelling does not mean harder and harder mathematics; instead it means piecing together more and more bits of structure, each bit of which is reasonably transparent in itself.

So, I say to the authors: a job well done. But, we need you to take it to the next step of the harder problems.

The Chair (Mrs S. Bridgeland, F.I.A.): Is there anyone else who would like to add something else to the shopping list for this working party who thought that they had just finished doing their work? Or is there somebody who wants to take a contrary view about the limitations of models and what we can and cannot do with them?

Mr P. J. Sweeting, F.I.A.: I note the purpose of the paper is to look at asset modelling in the context of the ICAS regime and similar sorts of regimes. I wonder if the authors have any views on the validity and the structure of those regimes having looked at the models used, and any insights as to whether those regimes could, or should, be changed in the light of what they have found.

Mr A. D. Smith (Student): We could take it a step further. There is a fair number of insurers and banks which have produced statements something along lines of, "We have calculated our capital as a 1 in 3,000 year event. We think what we need is x . As you will notice from our balance sheet, we have $x + 2$." Therefore you are supposed to understand that they will be good for at least the next 3,000 years.

Some of those organisations have recently been nationalised because, unfortunately, their 1 in 3,000 year events happened almost before the ink was dry on the paper.

So, of course, anybody involved in quantitative work is going to expect some challenge over those sorts of numbers. I do not think we are arguing against the risk based supervisory approach; what we are highlighting is the extent of model choice in fixing some of those numbers. So, for example, under Solvency I we had a 4% of liabilities solvency margin requirement. I think it was pretty much universally acknowledged that someone had pulled that 4% more or less out of thin air. I do not think anybody claims to have quantitative backing for it.

Now we have a 40% equity stress test. If we are claiming this is not pulled out of thin air, we may be kidding ourselves. We have a more complex regime, one which is arguably more sensitive to risks. But I think we have to be quite careful about the claims we make for this being driven by unanimous scientific assent, because that is not quite what we found.

The Chair: Are there any other views about either regulatory information or maybe even management information? There has been a lot of criticism of banks, and other financial institutions, suggesting an over-reliance on technical people understanding products and models rather than the senior people understanding them themselves. Are there any senior members of the profession who have worries about how these models are being used?

Mr R. H. Johnson, F.I.A.: I am sorry but I have only had a chance to scan this paper. But I did read the first part of it, and a lot of the interesting conclusions are in the early part, particularly the remarks about subjectiveness, the choice of modelling being what matters so much. I assume that the later, more technical sections are correct and have no comment on these.

The point about subjectiveness being a part of the modelling suits my intuition, and I really welcome it from that point of view. What we seem to have seen in the asset markets is people believing that the models were producing objective statements about probabilities. Certainly, in general insurance, good actuaries construct good models if they can. But they do not, if they are

good actuaries, believe them in the objective sense. Rather models are regarded as tools to help people to ask the right questions and to throw challenges back to colleagues, because the reality is we do not know what is going to happen in the future.

What we can do is help people to understand the decisions they are making now in the light of what is known now, and to that end it is much better to do some good calculations and then look at them critically than it is simply to say, "We cannot do any calculations. Let us have a 4% solvency margin; let us have a 40% capital requirement."

I have thought for some time that the calibration of the ICAS regime at 1 in 200 probability of ruin should be interpreted as a set of tests that people are going to do, in good spirit, which produces an answer which is consistent between companies. It could not be expected to tell you objectively what will happen with a probability of 1 in 200.

A probability of 1 in 200 is well beyond anybody's practical experience. When I have asked very numerate people, they, generally speaking, do not have an intuition about events that have that level of probability. The models we construct can easily be used to identify modelled events with 1 in 200 probabilities within those models and a spurious sense can build up that these model results are the same as the probabilities in the real world.

When the subjectiveness of modelling is acknowledged and embraced, there should be better subjective judgements and a wider understanding of the limitations of the models.

That does not mean that ICAS is not a good thing. At Lloyd's, we had a risk-based capital system which I thought was very good — I would think that because I was responsible for it. But ICAS is better. ICAS is better because it requires every agent at Lloyd's to go through a detailed exercise confronting the risks that face them, and then we look at that exercise. If we do not think that they have been honest in that assessment, we say so. The benefit of ICAS is in the system, and in the activities that people go through in order to produce the numbers.

The number itself, people should realise, is the consistency test. It is not a 1 in 200 test.

I think the paper is very good and I also support the suggestion of Mr Hitchcox about taking it on to the next level, and trying to bring in disparate insights.

The Chair: Any further comments in that vein or any other questions for the authors? Does anyone have any other comments about models which have not catered for extreme events in the ways they would have expected, about how they have needed to adapt models and whether this kind of approach, and this paper, would be helpful in their work?

Prof. A. D. Wilkie, C.B.E., F.F.A., F.I.A.: On the basis of what has just been said, I agree that any safety criterion that one chooses, a 5% significance test that many statisticians use, or a 1 in 200 test, is, in itself, arbitrary. But, as the previous speaker has said, at least it gives a sort of consistency among different firms and in different circumstances.

It is very difficult estimating a 1 in 200 level with rare data. In a general insurance context, I had the job of estimating a 1 in 200 event of the worst possible shipping accident that might happen in the next 200 years. I had no idea what the answer was, and the distributions I was using sometimes gave it as zero and sometimes gave it as about \$3 billion. So there was quite a difference in the result.

The fact is that nobody has any idea what the worst extreme events would be of that type.

General insurance has been mentioned, and general insurance uses the concept of realistic disaster scenarios, putting in some of the events that might have happened but have not; the events that might have happened to other companies but have not happened to yours.

I cannot think of how to do it easily within the investment world, but that sounds the right type of approach.

As somebody has mentioned, make sure your data series go back as far as 1929. It is a bit difficult with equities. The old Actuaries' Investment Indices were produced just at the end of 1929 and missed out the drop, which was a problem. They only captured the subsequent drop.

A couple of little points on indices which are relevant. In A.7.1 there is a very kind reference to the Heriot-Watt Gilt Database that Andrew Cairns and I produced. It says: "Yields and yield indices are available monthly since November 1998." In fact, the yields from the FTSE

Actuaries indices, are available monthly from December 1976. It is a much longer series. But between 1976 and the mid-nineties the tax system was different from what it became after the late nineties.

The indices changed in 1998. The system was that interest was taxed and capital was not and therefore pension funds, which were tax-free, tended to buy high coupon bonds, and individuals, who were subject to very high rates of tax at that time, bought low coupon bonds. There was not a yield curve but a yield surface because the yields were quite a bit different, depending on the coupon level. So you needed a high coupon yield, a medium coupon yield and a low coupon yield which represented three bits of that surface, and they had quite different yields in them.

If you were interested in pension fund work or a gross fund of some kind, then the high coupon ones were the right ones to use. For a life insurance company, possibly, the medium coupon ones were the right ones to use.

If you now take the present database, going back over the yields for that time and just pick up whatever has been produced by somebody on the basis of all the prices at that time and fit the yield curve to them, you will get unsatisfactory answers before the mid-1990s unless the coupon effect is actually measured correctly. So, that is one sort of problem with indices.

The second little problem was the one mentioned. The German markets appeared to drop in 1948 by 80%. I remember, I believe, that there was a 1 for 10 currency revaluation. I just wonder whether what started at 100 at the beginning of the year was revalued to 10 and increased to 20, or what started at 100 was revalued to 10 and then dropped to 2. Had the constructors of that index got the currency revaluation right? That certainly would be my question, seeing an outlier like that. Most outliers in data prove to be errors the first time through. Not in 1974-75. You can check; the extreme values were there.

The authors recommend the MSCI indices and mention a paper of mine recommending them. Actually, in my paper I did mention them but warned about them. In the series I looked at there are price indices and total return indices. These are inconsistent. You should be able to derive the dividend each month, and the dividend yields, from those.

Sometimes the apparent dividend is negative, which is clearly wrong. Sometimes it is clearly inconsistent. The apparent dividend in October 1987 in the UK dropped by 30%. Anybody who was working at that time knows that the dividend did not drop by 30% at all. Share prices did, but dividends were almost unchanged.

So there are some problems with those series. Personally, I would rather use the FTSE Actuaries World Index, although that is only available from 1987 onwards.

Obviously, it is very important which model you choose. One needs to have a flexible mind as to how many different models it is worthwhile reviewing.

The authors seem rather keen on using simple returns on equities rather than log returns. They do say there are problems with simple returns because they can go negative. That means that a simple return model is clearly wrong. There is no point, to my mind, using it at all.

If you are looking at individual shares, which can go to zero, and you use log returns, there is a positive probability of zero, which would not be available within the lognormal distribution. The authors point out that lognormal does not fit very well — that is, normal is a worse fit to log returns than to simple returns. Do not use a normal distribution but use one of the other distributions that are mentioned. There are quite a lot of them.

A fairly easy distribution to fit and to simulate for the return in the period is using one lognormal minus another lognormal. That can be pretty fat tailed, skew if you wish. You can standardise it to $(0,1)$, mean zero, unit standard deviation. Or you can take a Pareto minus a Pareto, or a Pareto minus a lognormal, or any of the popular distributions that general insurance people use and take the difference between them. You get really quite good fits that way.

The authors have appropriately looked at share prices over a long period. They have fallen a little bit into the trap into which the short-term financial economists have fallen (there are long-term financial economists too). Option pricing has become so important that the short-term seems to have dominated thought for a while.

In the short run, dividends and earnings do not change much and therefore changes in

dividend yields, or in P/E ratios are effectively directly related to changes in share prices. In the long run the mean reversion that is mentioned means reversion to a sort of middle level of P/E or a middle level of dividend yield. You cannot ignore dividends and earnings in relation to the share market or individual share prices.

There are companies which say “We will never ever — and our constitution prevents us — declare any dividends whatever.” There are many tens of thousands of these companies. Almost all companies limited by guarantee have that concept. They cannot pay dividends and their shares are non-existent and worth nothing. So if a company says, “Our constitution prevents us ever giving any returns to shareholders at all” then it is not worth anything. So you buy shares ultimately, in the long run, to get dividends. So one may as well look at dividend yields.

This is where the conditional versus unconditional argument matters. If share prices drop over a period because earnings and dividends have dropped, or seem very likely to drop in the near future, which certainly might well be the case at the moment, then it is not surprising that share prices go down. If, as in the summer months of 1987, share prices went soaring up, with no particular change in dividends or earnings or any particular prospect of change, it is not surprising that they then come back down again. What is surprising is that they did it all on one day in October 1987, rather than over some months.

So, if at the end of the year one was having to do a stress test, and share prices have dropped considerably, and one’s forecast of dividends is that they will not be affected at all, then it seems reasonable to use a rather less strong stress test. If, on the other hand, dividends will come down, then there may be perfectly good justification for the share prices to come down. For example, we know with banks at present that dividends are bound to come down, because the banks are borrowing money from the government and are being obliged to stop paying dividends for a period. This does not mean that profits will be banned, though they might be a lot lower — but it does mean dividends will not be paid for a period.

Another thing is that the authors have preferred in some cases to use overlapping annual returns over successive twelve month periods starting at the beginning or the end of each month rather than looking at monthly returns. They say that it is unsatisfactory looking at returns monthly because of factors such as volatility clustering.

Another approach is to look at the monthly returns or rather the monthly dividend yield series. Allow for the auto-regression and the volatility variation in the series and then look at the distribution of residuals from that process.

This procedure is trying to get independent monthly returns. While I think it is correct that twelve separate months do not tell you much more about the annual data than one year in some respects, in other respects, they do tell you quite a lot more about it. You can do bootstrapping methods; that is, simulation having fitted your monthly model and see whether yearly simulated values represent the sort of yearly data that might be experienced.

The authors mention simulation but do not actually give any simulations. One of the awkward things about the distributions that they mention (T, hyperbolic and Pearson Type IV), is that in order to make use of them in simulations, you need to be able to simulate them. Student-T can be simulated without too much difficulty. I have not had the chance to look up and to see how one would simulate the others.

The authors, in ¶6.3.8, mention the stable family. This is a theoretically very nice set of distributions because when you add them together they remain stable and do not tend to normality. But they do also have considerable disadvantages like infinite moments, and ill-defined or infinite means.

Stable distributions produce problems. For example, option prices are far too cheap to buy and far too dear to write if you believe that a stable distribution fits the data. The stable distribution does not seem right. There is evidence that the kurtosis is high for returns on a daily or weekly basis, as rightly mentioned in the paper, but as you add the returns together and get monthly or yearly data, the kurtosis reduces quite a lot. That would not be true for stable distributions. The kurtosis would stay high. There are probably distributions with finite variance even if the higher moments are infinite. Infinite kurtosis is a bit of a nuisance. But there are

plenty of distributions that have infinite skewness and kurtosis. Pareto minus Pareto probably would have, depending on the parameters.

I hope those comments are of some use and maybe encourage other people to say some more as well on the technical aspects as well as the general principles in this very good paper.

The Chair: As the authors have only a few minutes to respond at the end of any discussion, I would like to give them the opportunity to respond to any of those particular points while they are fresh in our minds.

Mr Smith: Many of those points are well made and we would, I think, agree with most of them.

In terms of the MSCI data, we are aware of some of those apparent imperfections of negative dividend yields. You could actually have that happening if there were more rights issues than dividends paid. But we have not checked whether that has actually occurred during the periods in question or whether it is just a data error.

One of the things about the MSCI data is that they have been used, more or less since their inception, by investment managers as benchmarks. That actually helps a lot. It means if you get something in the data that fouls up their performance measures the managers will complain like mad and the compilers of the index seek to fix it. In many of the longer data series what has actually happened is somebody has said: "Would it not be nice to have a data series going back to 1800? Let us go and find some old newspapers."

It has not been compiled contemporaneously. So you can find slip-ups like, for example, railway shares that did not exist at some particular period when they compiled the list of companies, and that sort of thing. There are a lot of biases that come from constructing an index with hindsight after the event has happened.

Having investment managers scrutinising your index is not totally proof against that sort of error, but I think it does help. That is another reason why we looked at the MSCI, which was the ability to include the 87 data which is difficult to do in the FTA All-Share Index. But I take your point: none of these data series are perfect. That is one of the facts of life that you have to live with. The further back you go, potentially, the more useful long-term information you can obtain, but also the more you struggle with data inadequacies.

You made some points about modelling total returns or modelling prices and dividends separately. Many people in this hall will be aware of Professor Wilkie's own work in this area. We have some sympathy with that view. The trouble is that when you try to build models that mean-revert to a particular level of dividend yields, dividend yields have this stubborn habit of not doing what you want.

One example of that would be the MFR, introduced in 1997 with the hard-coded number that a lot of us mischievously called Wilson's constant of 4.25 dividend yield which had been apparently the mean reversion level for the previous umpteen years but then it stubbornly went down and never came up again.

In the resilience test there have been various adjustments to cope, supposedly, with mean reversion in relation to dividend yields and P/E ratios. In Figure 1, Section 2.4.4, we show how an attempt to reflect that in a resilience test produces some rather odd effects. I am not saying it cannot be done, but trying to construct tests from a bivariate series of dividend yields and dividends seems intrinsically harder and potentially more error-prone than constructing one based only on total returns. It is not impossible, but there do not seem to be a lot of successful attempts out there that we were able to pick up.

Your point about overlapping annual intervals is one we did agonise over in the working party. In some ways it would have been a lot neater to use non-overlapping intervals. Let me first of all say that the reason we used overlapping intervals, rather than purely year-end numbers, was because things like the October 1987 crash do not show up at all on year-end numbers. You either decide you are arbitrarily going to fit October to October or alternatively you are going to throw in all the month ends. That was what we did.

So why not look at just monthly data and then allow for the auto-regression or the volatility

clustering? That can be done. The problem is how do you test the way that you are modelling that auto-regression? How do you test the way you are aggregating those monthly returns to get annual returns?

Just about the only way we could think of to test it was to look at the data for annual returns which was kind of what we have done anyway.

The difficulty in modelling monthly returns, is that the way in which you then aggregate them to annual returns, becomes very model-dependent.

You see that on a bigger scale for the kind of asset liability studies that some pension funds carry out where you are projecting 20 or 30 years ahead based on compounding annual data. Mean reversion parameters that are quite difficult to estimate suddenly turn out to be pretty critical for how you project those longer term horizons.

We do not have a magic solution to this problem. We felt that you are more or less forced back to the annual return data anyway in order to validate your monthly series. That does not mean that there could not be some extra insight from monthly data. I am very happy to hear suggestions for how that might be done better.

Mr T. W. Hewitson, F.F.A.: One of the points that Mr Smith just made and is of interest to a number of people at the moment is cyclicity, when some people are saying that if markets are very high, then perhaps you can anticipate a bigger fall in equities, and conversely at the moment, when markets are relatively low, perhaps you could anticipate some lower fall when setting capital requirements.

I fully agree that a very simple type of approach to allow for such possible cyclicity, such as the one described in Section 2 of the paper, is unlikely to work sensibly in all conditions. However, it would be an interesting challenge for the working party to see if they could come up with some better method, maybe based not on two variables but on three or four, or whatever reasonable number it takes to produce an approach that also looks more meaningful and plausible. This could then potentially dampen the effect on the capital requirements for firms of market excesses in either direction.

Mr Frankland, in some of his earlier remarks, made the comment: where does the 40% figure for the assumed equity fall come from? I think the answer is there has been a lot of research by various people, including the authors of this paper, which suggests that something close to 40% probably is the central estimate for a 99.5th percentile fall in a typical index of equity market values, if you ignore any potential cyclicity that exists in the data, albeit that there is of course still a significant level of uncertainty around that estimate.

Mrs K. A. Morgan, F.I.A.: I should like to suggest that people read the paper alongside the recent FSA discussion paper "Insurance Risk Management — the Path to Solvency II".

That paper sets out all of the tests that are currently detailed in the Level 1 framework directive for Solvency II for approving the use of internal models. It is probably worth comparing some of the work done in today's paper with those tests: the use test, so how these models are used, and calibration, statistical quality, documentation, validation, and so on.

Solvency II is a risk-based framework and it is very important to understand risk, which is what the paper is trying to describe, as I understood it. Solvency II is also about unpacking assumptions so that management understand what is going on in the firm, documenting things well, and understanding the drawbacks of models, as the authors have done in their paper with the different models that they have used.

The process by which risk assessments are made is also important. This also links to the upcoming Board for Actuarial Standards consultation paper on models which makes clear that actuaries need to communicate well about the models that they use, about the data that they use and the choice of model, which again are all clearly explained in today's paper.

I think the paper highlights an area where the profession could help supervisors in expanding the Solvency II framework. For example, in explaining how expert judgment is used, and how supervisors can assess it, particularly in the context of market risk, as covered in this paper, but also in other risks as well. I think David Wilkie's point about indices and how they are

constructed and used is relevant here. Understanding how those indices are used, and the drawbacks of that is very important.

The discussion paper that I mentioned recommends that the industry comes up with better data to underpin internal models. I think the profession could help on that. Linking back to what we were saying about indices, the profession could advise on how those indices should be constructed and what kind of data could be collected.

I was at a session last week looking at the data that is used to help governments decide on policy, and one of the directors there from the Government mentioned that he has a 20 year plan, so he is getting intelligent people together now to think about what data will be needed in 20 years' time so it can start to be collected now. Maybe that is something which we, as a profession, should be doing for the insurance industry.

My own day-to-day job means that I see Solvency II everywhere. Mr Smith mentioned the open forum on Mandelbrot and other new methods which I obviously saw completely in the context of Solvency II. But this paper, and papers of this quality, really help to add value to the thought processes as we put more flesh on the bones of the Solvency II regime. I plan to send this paper to my colleagues in CEIOPS and hope that they will find it useful as we develop our thinking.

Dr L. M. Pryor, F.I.A.: I should like to reiterate some of the points people have already made.

First, let me welcome this paper which talks about a really important aspect of modelling in today's climate. Modelling is important and it is fundamental to a lot of what actuaries do.

As Mrs Morgan said, the BAS will be publishing a consultation paper for its generic standard on modelling later this month. In that consultation paper we will stress the fact that models are generally used in the production of information that is then used as the basis for decisions. It is important that the people making the decisions are aware of both the strengths and the limitations of the models that have been used to produce the information. This paper will help to focus the minds of the people who are preparing that information, who are often actuaries, on both the power and the limitations of what they are doing, and help them communicate them better to the users.

So I am very pleased to see this paper and I hope that many, many people read it. And, as Mrs Morgan said, read it in conjunction with our consultation paper.

Mr A. J. Wells, F.I.A.: I should just like to pick up briefly on three points from the conclusion of this very important and, dare I say, timely paper.

One is from ¶8.2.1, where the authors say: "It is a common belief ... that equity returns show a negative skew ... and leptokurtosis" I have certainly treated this as almost perceived wisdom. I daresay I am not the only actuary here who did so. The authors rather give the lie to this in the rest of the paragraph, and I think this is a very important lesson for us to learn and an area where more research may be helpful.

In ¶8.2.2 they say at the end: "We have estimated a 95% confidence interval of 30-40% using bootstrapping." I rather think my company's ICA might look rather at an equity fall of 25% than it does at 55%. Again, I do not think we will be the only ones.

The third is in ¶8.3.2. This is a lesson I have already learned as I have attended a presentation on it by one member of the working party. It says: "... assets and liabilities have been matched against parallel shifts in yields curves may show little capital requirement under a one factor model, but a larger requirement under a two factor model that includes the possibility of a change of slope." I think that is a very important lesson to companies who have not already learned it. They need to pay more attention to it.

Mr D. I. W. Reynolds, F.I.A.: I am going to take a slightly different approach to this debate. Despite a M.Sc. in Statistics I cannot keep up with the statistics used in this paper.

I had an interesting weekend. On Friday evening I read the paper. On Sunday I completed

my reading of the *Misbehaviour of Markets* by Mandelbrot and Hudson (2005). Both cover similar things. They are both saying that a lot of the methodologies that are being used do not work because they do not allow for the fat tails.

There are a couple of other things in Mandelbrot and Hudson that are perhaps relevant to this debate. To put them in context, just remember that the way in which the FSA looks at things is it looks both at the probability and the impact of events happening. Mandelbrot and Hudson separate these out too. Their example goes back to a man called Hurst in the 19th-century who investigated the flood levels of the Nile. We all know about the seven good years and the seven fallow years from the Bible. What he found was that the distribution of floods could be explained by a normal distribution.

What could not be explained was what is probably referred to earlier by clustering or volatility; the fact that floods would tend to happen in runs. So you would either have a run of large floods or a run of low floods. So we actually have to look at not only what happens in the tails but also the clustering of events.

Maybe that is a way in which future work, not necessarily of this working party but of this profession, should go.

Some of what we should look at in modelling relates rather closer to what is called operations research rather than pure statistical analysis.

Another of the comments in Mandelbrot and Hudson is the extent to which there are long term correlations, that is correlation between events that are not immediately contiguous in time. I want to give a couple of examples of events that occur and which do have such long term correlations or impacts.

Let us start with that first budget of Gordon Brown's and his raid on pension funds. We know what some of the effects were. The pension funds had less money. The employers had to put more money in. But this was happening at a time when a separate driver, improving mortality, was taking place. We all know that a very large percentage of defined benefit pension funds have closed down over a long period, and it is still happening.

So one event has this long run impact. I suspect that it is having another impact which is that the pension funds and their trustees change their investment mix. So the event did not only effect which funds remain open but also on how they invest and therefore what happens in the investment markets.

So I think we have to take events and look at how they run forward and not necessarily just analyse things in our models on a statistical basis.

The second event I am going to use is a bit light-hearted. It is the impact of CPD on actuaries. Meetings in this hall are now fuller than they used to be, but we do not get any more speakers because people come here to get their CPD and to listen.

Maybe the approach to these meetings should not be to get rid of the opener but actually we should have an opener, a first, a second and a third speaker, and we should actually plan the discussion much more. That is light-hearted but I hope the President and the Council will consider that.

The final example is much more recent. We know that one of the things that happened is that the FSA told banks that they needed more capital. If you have actually lost capital and you still have the same volume of business, then, sensibly, you need more capital.

But if they are actually saying you need much more capital — to cover the risks you have just suffered from, and it is greater than the amount you have lost, then the banks have two choices. They can find the capital, and in extremis they may have no choice but to find the capital and accept the government's money and bear the 12% after-tax return on preference shares and no dividends for five years.

But they can take another action. They can actually reduce the amount of risk they are taking on. I think there is a danger that that is what is happening. — The Times said this morning that it is becoming very difficult to get letters of credit and, as a result, some world trade is not taking place because the traders cannot shift the goods around the world. That could be a serious consequence way beyond "Are the banks solvent?"

This is an excellent paper, an excellent statistical analysis. It undermines a lot of what has

been done in the ICA calculations, and people have to go back and review their simulations and their stochastic modelling.

However, I think that there is more to be done which is actually to look at some of the drivers that are taking place. I think the more that the profession, the leaders of the various financial industries and government and regulators look at the drivers, and those economic or operational research models, the better off we will be.

The Chair: If you are here without having read the paper, then do feel free to ask the authors about the practical implications of their work for your day job.

As the trustee of a significant pension scheme with an equity bias, which has its actuarial valuation at the end of this year, my question for the authors is: what does this mean for the modelling work that I should undertake to understand extreme events? What are the limitations of the models that most consultants will be using? What can a pension fund trustee do to try and get round this problem of the unconditional vs the conditional future?

Do we model where we are starting from, given where we have been recently, or do we try and look at the average year, as the authors call it in ¶3.1.1?

That is my challenge in my day job. While it is constructive to devise a shopping list for next year's working party, I have to sign off an actuarial valuation within the next year and review my funding and investment risks. So I have a practical problem which I am expecting my advisers to be able to help me solve. Like Dr Pryor's and BAS's work on the standards on modelling, it is not just about the technical construction aspects, but how these things are going to be used. What have I got to look out for when I am being advised using the models that are already out there?

Mr R. Kelsey, F.F.A.: I was going to agree with a point made by the Chair. We have had a bad year, it was a one in fifty year event. How much capital do we need at the end of a year? Are we allowed to fund the pension scheme or fund our own company through profits for another five years to get up to the next one in 200 year event?

My second point is are we trying to be too ambitious in estimating a 1 in 200 year event? Perhaps we should be a little bit normal and try to evaluate, say, two independent one in 15 year events.

Mr E. M. Varnell, F.I.A.: Maybe it is worth saying that as part of the reading that we did for this paper there was one book which some of you might have read by Riccardo Rebonato, entitled "Plight of the Fortune Tellers" (ISBN 978-0691133614) which offered an interesting insight.

One part of this book which struck us was research into the psychological difficulty of making rational management decisions very far in the tail of a distribution. When events are placed in the context of one in ten, one in five or one in 15, as has just been said, people tend to make more rational decisions.

Mr Smith: If I may make a general observation, it is nice to see so many general insurance people here. It is a shame there are not many pensions people. Perhaps that is because a current issues pensions seminar ran today and finished at ten minutes before five. I had a bicycle, but apparently I was the only one who did.

Comments from general insurance people are helpful. There is quite a contrast between the capital markets and general insurance markets. A well attested attribute of general insurance markets, following a large claim, one that dents the capacity in the market, is that rates subsequently rise, or rates tend to harden subsequently, partly as a result of reduced capacity in the market. It is a phenomenon that is sometimes called payback.

It is sometimes presented as reinsurers trying to claw back their losses. There are also rational forward-looking reasons for that to happen.

The corresponding effect in stock markets is much more controversial. Plainly, if you look back at historic low points in the stock market, subsequently the stock market rose because that

is what a low point means. It means that it is lower than it was a year later as well as lower than it was a year earlier. Currently, all we know is that it is lower than it was a year earlier. We do not have the second part of that. So we do not know that it is at a low.

Most of us in this room would love to be able to produce some evidence that says because the stock market has collapsed to such a terrible extent, it cannot go any further, there is some sort of bottom that is just below where we are now.

If we are intellectually honest and look at the data, we do not find a lot of evidence for that. So to my mind it is one thing for regulators, rating agencies or whatever to say, "We are going to use less onerous tests at this year end and we see it as a pragmatic fix because otherwise we would be opening our entire insurance industry to takeover by the French, and we do not want that", but it is another thing to represent that as being a 1 in 200 event, and that is where we would get quite uncomfortable from an intellectual point view.

Mr A. Coulter (a visitor): I have not gone through the paper in much detail but obviously everyone who helped author it has been very close to this data and I get the context that what has happened recently is not particularly included in it.

Given that we do now have another major negative event in the market, what is the feeling of the authors? Is 1 in 200 actually a reasonable capital level? Is it too stringent? We have had a couple of major disasters over the past 100 years. Are we actually seeing only 1 in 30 year events, and the 1 in 200 is something so much further off the scale that it is unreasonable to have a capital level that high? Or do you think that we are seeing 1 in 50, 1 in 70 events and we have just seen a few of them?

Mr Smith: We have been quite careful not to tell regulators how to regulate. We have set ourselves the problem of "If you want to measure a 1 in 200 event this is how you might go about it." One possible consequence of that is that people look at that and they do not like it.

It is helpful to hear that this will be provided to some of the CEIOPS as part of their deliberations.

We are not trying to demolish or indeed build up a regulatory regime. Other speakers have mentioned how it would be easier to measure a one in ten year event, and, of course, we agree with that. It would be an awful lot easier and it is a lot less dependent on the vagaries of data sets. We are trying to answer the question: "If you have to measure 1 in 200, how might you go about it? How reliable might the answer be?" while recognising that there are lots of other interesting questions that we have not answered.

Professor Wilkie: I was talking earlier about looking at auto-regression and mean reversion. I agree that is only one approach. One might look at plenty of other things.

Thinking back to 1974 and the events of that surprising year in the stock market, and 1975, looking at my data at about that time I discovered that the largest either seven or eight months of drop in share prices had taken place at the end of May 1974. You might have thought that that was an extreme event. The next seven months were a lot worse. So whatever you had, it could get worse.

I also remember writing an article — and I still have the document because I thought it was worth keeping — for the board of the company I was with at the time about the stock market in October 1974 when dividend yields had not quite got to their peak but were still up at about 8% or 9%. I observed that dividend yields were very high, and I thought that in a year's time they would be a lot lower, but I did not know whether dividends would halve or prices would double.

Those who remember the political situation at the time may remember that a Labour Government had been elected in February 1974 as a minority government and had an increased majority in about September or October, I cannot remember the exact date of the election. Part of its policy was extensive rationalisation of the commanding heights of the economy, etc.

The Conservatives, in 1972, had rescued Rolls-Royce, which had got into difficulties. Rolls-Royce recovered and has flourished since then. So it would not have been impossible for a

Labour Government to say, "We will help out all these bust companies simply by nationalising them."

That was one strand. The second strand was that there was a tax issue. There was very high inflation that year. The increase in value of stocks on a historic cost basis was brought into profits and was taxable. Companies that had, for example, huge reserves of oil were finding that they were going to have to pay huge amounts of tax when all they had was lots of oil to pay it with. They had not sold it yet. So that was very uncomfortable.

I think it was in November that the Labour Chancellor, Denis Healey, changed the tax regime and between Christmas and New Year the Bank of England rescued *Burmah Oil*, which had run into difficulties, in a moderately reasonable way, without nationalising it. *Burmah Oil* at the time owned about a quarter of BP.

In October 1974 it was quite reasonable to say there was a binary possibility, that the Government would quite intentionally let a lot of companies go bust in order to nationalise them and that the shareholders would lose so the shares would be worth nothing. Alternatively, they would not do this and the shares would recover considerably. The second option took place and share prices more or less doubled between the end of December and the end of February 1975.

So at the moment I think we may have a binary position as well. Mr Reynolds mentioned the problem of credit not being available, and the sensible thing for a company that has too little capital to do is to write less business. Lend less in mortgages: house prices will come down, possibly quite a lot. Lend less to businesses: businesses will not be able to flourish. Then we are getting into a 1929 position when the market kept on going down for quite a long time. There was a big crash in 1929 and big drops in subsequent years.

Are we in a 1974 position when everything will recover or are we in the 1929 position when things will get a lot worse? There is a binary problem here and we do not know which it is.

The Chair: So are you suggesting that any trustees should look at both sides of that binary outcome and then ascribe whatever probability they think to either side?

Professor Wilkie: I am inclined to look at them both and take the worst rather than attribute probabilities to either of them. Imagine that you have half a chance of winning £1 million and half a chance of losing £1 million and those are the only two options. You can say that the expectation is zero but you know that will not happen. If you are wanting a cautious approach use the worst. You must think: is it reasonable that there is a possibility of that? Is there a non-negligible possibility of that sort of binary event happening?

Mr M. R. Kipling, F.I.A.: We have spent quite a lot of time talking about equities but there is also a section of the paper on fixed interest, Section 7. Perhaps because the gilts yield curve has been quite well-behaved in the last year or so, there has not been quite so much focus on that section, although admittedly the swap curve has been a little bit more difficult to explain in rational terms recently.

Nevertheless, I was particularly intrigued by Figure 37 where through a series of mathematical formulae, which I have not had the time or possibly even the ability to understand, the authors come up with what look like two quite practical stress tests to use against long-dated liabilities.

If I have understood what its purpose is, the dotted line, Stress 2, appears to be fairly close to a reasonably conventional twist stress of the type which, we heard earlier, definitely ought to be applied in addition to a level yield shift. Stress 1, on the other hand, seemed rather more interesting. It more or less had the very long end and the short end anchored with some sort of a bulge stress applied round about the middle term. I wondered if this was the intuitive interpretation meant to be taken away and whether, if one turns both stresses upside-down, one might have four useful tests in total to apply.

Mr Smith: As that chart was one of mine, it probably falls to me to reply. You are right, the testing up and down does make sense. We really struggled with the yield curves. The reason that

we struggled was because if you have 40 years of cash flows, you have potentially a 40-dimensional thing to model.

In this case we have a two factor stress model so you have to collapse those 40 dimensions into two dimensions. Inevitably, therefore, you are throwing away 38 dimensions, which is the majority of them. So, picking those 38 is key. An earlier speaker picked up that we had noticed this in looking at the inadequacy of simply checking parallel stress tests.

What we are showing here is the effect of calibrating a model in such a way that it hits the historic volatility of discount rates at each term but also the historic volatility of forward rates. That is what you are exposed to if, for example, you have an 11-year liability and you try and match it with a ten-year bond and you recognise there is a one-year mis-match.

If those are the sort of matching problems that you face, then this decomposition into two factors is probably about the best you can do if you are constrained to disregard the other 38. It is very dependent on the way that you set about the problem.

One of the things that we were quite surprised at was a technique called principal components analysis. It seems to be quite widely used in this context and it produces results that say things like the first three components are level and the level of twist in the convexity explains 99% of the variability.

What we found was that those numbers could be pretty misleading in that it is quite possible that your portfolio only explains 20% of the variability. The way percentages are quoted tends to be misleading. We are looking at a somewhat arcane list where measure of yield curve moves rather than something for your particular portfolio.

Although we are glad to hear that those curves are interesting, there is a layer of sophistication that we think people ought to be taking into account. In particular, if you use these two stresses to construct a portfolio that closely matches your liability, then you are unsurprisingly applying those stresses and you will come away with the conclusion that you do not need a lot of capital.

So you always need, for the purpose of understanding capital adequacy, to be testing several steps ahead of what you are using to construct the portfolio in the first place.

Mr M. H. D. Kemp, F.I.A. (closing the discussion): May I first of all thank the authors for what I found to be a very interesting paper. The discussion today suggests to me that those who have spoken have also found it most interesting. Like the Chair, I too would encourage everyone to contribute to discussions at future sessional meetings.

Quite a few themes were raised in the discussion. As one might have guessed, there was a strong focus on fat tails, Mandelbrot and how the world is not Normally distributed. There was also some discussion, which I felt was potentially overly negative, about financial economics more generally, so I will first comment on this.

I would question whether it is true that financial economics always assumes that the world is Normally distributed, as some of those contributing to the discussion seemed to imply. For example, within the credit world it is fully accepted that organisations do go bankrupt from time to time. The standard way of modelling credit risk does therefore involve just such an assumption. Only in equity-land is it perhaps more common to see the assumption adopted that the world 'smoothly' evolves through time in a way that does not involve extreme jumps from time to time. So perhaps the criticisms of financial economics that some people have raised during the discussion are more relevant to how people think financial economics is applied in equity-land than to how it is actually applied in credit-land.

Quite a few people commented on the subjectivity of models. The message is that models inevitably involve a fair degree of subjectivity but that this does not detract from the merits of the paper itself, which takes this fact of life into account. For example, Kathryn Morgan thought that the paper was sufficiently useful for her to want to send it to her EU colleagues working on Solvency II.

Another topic aired in the discussion (but less well covered in the paper) was that of cyclicity. This led into a wider discussion about whether or not capital adequacy should be toned down after a strong market decline. Cyclicity is clearly a very, very important issue at the

present time. As Ian Reynolds pointed out, banks currently have two choices if they are strapped for cash. They can raise more capital or they can reduce risk. We are seeing at the moment a very severe deleveraging, a reduction in risk appetite across the board. This topic is likely to remain very high on the agenda for quite a while, particularly for banks but, one suspects, also for insurance companies.

Another area of discussion initiated by David Wilkie focused on some of the weaknesses of published indices, and some of the challenges that this presents. He came up with what I thought was an interesting comment which was that most observed outliers are actually data errors. Would that this were always the case. However, clearly, when we are looking at extreme events, it is the 'true' outliers (rather than the data 'errors') that are key.

Richard Kelsey suggested that one way of developing practical approaches might be to focus not principally on the very extreme events but on somewhat less extreme events, maybe the one in ten, one in 15 year events. But as Andrew said, this was not the mandate of this paper, which was to look at the one in 200 year event.

If I can add a few thoughts of my own to the paper and to the discussion, the thing that I would focus on most is that I have noticed that whenever we talk about these kinds of problems we tend to focus on the 'probability' of outcome. I would ask people also to think about the 'cost' of the outcomes when they think about risk management.

One of the potential benefits of the thinking underlying financial economics is that it draws out the distinction between our own views about how the world might evolve and the cost of actually going out and hedging or protecting ourselves against the risks involved. These include the risk that our views about how the future might evolve might be wrong. There is something out there called 'the market', and it rather often shows up how wrong our own views can be!

Several times during the discussion (indeed it was brought out by Ralph Frankland at outset) people commented on the strong equity market focus of the paper. I would again echo the request for more work on credit risk and, dare I say it, also on liquidity risk. These are the types of risk that have been most challenging for those parts of the financial services industry which have suffered the most distress of late.

The other main area that I suggest the working party explore further, if they wish to take up the mantle and continue operating next year and in the years ahead, would be the issue of cyclicalities. Of course the authors have pointed out that they do not want to do the Regulator's job, and this is a very fair comment. They are not paid to do that. However, I do think that it behoves us all (particularly in a professional context) to think about the wider impact of the actions that we might take.

My final comment thus refers to the discussion between David Wilkie and Sally Bridgeland regarding whether always to chose the worst case outcome in a binary situation when setting capital bases. One of the challenges of doing this is that you would then potentially become part of the pro-cyclicalities issue, i.e. part of the problem rather than part of the solution. Pro-cyclicalities is not just about banks, or even also insurers. It can also be about pension funds. Indeed, if a problem is sufficiently large to have become systemic, then it probably actually affects us all.

The Chair: I want to express my own thanks for a very topical, thought-provoking paper. Thanks from all of us to the authors, to the closer, Mr Kemp, and to those who participated in this evening's discussion.

REFERENCE

MANDELBROT, B.B. & HUDSON, R.L. (2005). *The (mis)behaviour of markets: a fractal view of risk, ruin and reward*. Profile Books, ISBN 1861977905.