Commentary

Beverly Hirtle

I am very pleased to speak here today and to comment on these three very interesting and constructive papers dealing with value-at-risk modeling issues. In my view, each paper is an excellent example of what academic research has to tell practitioners and supervisors about the practical problems of constructing value-at-risk models. Each paper examines a particular aspect of value-at-risk modeling or validation, and offers important insights into the very real issues that can arise when specifying these models and when considering their use for supervisory purposes. In that sense, the papers make important contributions to our understanding of how these models are likely to work in practice.

Danielsson, de Vries, and Jørgensen

The Danielsson, de Vries, and Jørgensen paper examines some key issues surrounding the question of how well current state-of-the-art, value-at-risk models capture the behavior of the tails of the distribution of profit and loss, that is, those rare but important instances in which large losses are realized. As the paper points out, this question is a fundamental one in the world of value-at-risk modeling, since both risk managers and supervisors are presumably quite concerned about such events. In fact, one of the key motivations for the development of value-at-risk models was to be able to answer the question, If something goes really wrong, how much money am I likely to lose? Put more technically, risk managers and the senior management of financial institutions wanted to be able to assess both the *probability* that large losses would occur and the *extent* of losses in the event of unfortunate movements in markets. When supervisors began considering the use of these models for risk-based capital purposes, the fundamental questions were much the same. Thus, for all these reasons, the ability to model the tails of the distribution accurately is an important concern.

As the Danielsson et al. paper shows, this ability is especially key when there is suspicion that the distribution might feature "fat tails." As you know, the phrase fat tails refers to the situation in which the *actual* probability of experiencing a loss of a given size—generally, a large loss that would be considered to have a low probability of occurring—is greater than the probability predicted by the distribution *assumed* in the value-at-risk model. Obviously, this disparity would be a matter of concern for risk managers and for supervisors who would like to use value-at-risk models for risk-based capital purposes.

Beverly Hirtle is a vice president at the Federal Reserve Bank of New York.

The paper suggests a method for addressing this situation. I will not go into the details of the analysis, but the paper proposes a method of estimating the overall distribution of potential profits and losses that essentially combines fairly standard methods for specifying the middle of the distribution with an alternative approach for estimating the tails. The paper then tests this modeling approach using random portfolios composed of U.S. equities and concludes that, at least for these portfolios, the "tail estimator" approach outperforms value-at-risk models based on a normal distribution and historical simulation.

When thinking about the practical implications of the proposed tail estimator technique, at least one significant question occurs to me. The empirical experiments reported in the paper are based on a fairly large data sample of 1,500 trading-day observations, or about six years of historical data. While this long data history may be available for certain instruments, it strikes me that these are more data than are likely to be available for at least some of the key risk factors that could influence the behavior of many financial institutions' portfolios, particularly when regime shifts and major market breaks are taken into account. Thus, the question that arises is, How well would the proposed tail estimator approach perform relative to more standard value-at-risk techniques when used on an historical data set more typical of the size used by financial institutions in their value-at-risk models, say, one to three years of data? At its heart, the question I am asking is whether the tail estimator approach would continue to perform significantly better than other value-at-risk methods under the more typical conditions facing financial institutions, both in terms of data availability and in terms of more complex portfolios. This is a question on which future research in this area might focus.

Christoffersen, Diebold, and Schuermann

The Christoffersen, Diebold, and Schuermann paper addresses another key practical issue in value-at-risk modeling, namely, whether the volatility of important financial market variables such as stock price indices and exchange rates is *forecastable*. By asking whether volatility is forecastable, the paper essentially asks whether there is value to using recently developed econometric techniques—such as some form of GARCH estimation—to try to improve the forecast of the next period's volatility, or whether it makes more sense to view volatility as being fairly constant over the long run. In technical terms, the question concerns whether *conditional* volatility estimates, which place more weight on recent financial market data, outperform *unconditional* volatility estimates, which are based on information from a fairly long historical observation period.

The answer, as the paper makes clear, is that *it depends.* Specifically, it depends on the horizon—or holding period—being examined. The results in the paper indicate that for holding periods of about ten days or more, there is little evidence that volatility is forecastable and, therefore, that more complex estimation techniques are warranted. For shorter horizons, in contrast, the paper concludes that volatility dynamics play an important role in our understanding of financial market behavior.

The basic message of the paper—that the appropriate estimation technique depends on the holding period used in the value-at-risk estimate—implies that there is no simple response to the question, What is the best way to construct value-at-risk models? The answer will clearly vary with the value-at-risk estimates' purpose.

As valuable as the contribution of the Christoffersen et al. paper is, there are some extensions that would link the work even more closely to the real world issues that supervisors and risk managers are likely to face. In particular, the analysis is based on examinations of the behavior of individual financial time series, such as equity price indices, exchange rates, and U.S. Treasury bond returns. Essentially, the analysis considers each individual financial variable as a very simple portfolio consisting of just one instrument. An interesting extension would be to see how or whether the conclusions of the analysis would change if more complex portfolios were considered. That is, would the conclusions be altered if the volatility of portfolios of multiple instruments were considered?

The results already suggest that the ability to forecast volatility is somewhat dependent on the financial variable in question—for instance, Treasury bond returns appear to have forecastable volatility for holding periods as long as twenty days, compared with about ten days for some of the other variables tested. It would be interesting, then, to build on this observation by constructing portfolios comprised of a mixture of instruments that more closely mirror the portfolio compositions that financial institutions are likely to have in practice. Such an experiment presumes, of course, that the risk manager is interested in knowing whether the volatility of the *portfolio* can be forecast, as opposed to the volatility of individual financial variables. In practice, risk managers and supervisors may be interested in knowing the answer to both questions.

LOPEZ

Finally, the paper by my colleague Jose Lopez addresses another important area in the world of value at risk: model validation. The paper explores the question, How can we assess the accuracy and performance of a value-at-risk model? To answer this question, it is first necessary to define what we mean by "accuracy." As the paper points out, there are several potential definitions. First, by accuracy, we could mean, how well does the model measure a particular percentile of the profit-and-loss distribution? This is the definition that has been incorporated into the market risk capital requirements through the so-called backtesting process. As the paper points out, approaches to assessing model accuracy along this dimension have received considerable attention from both practitioners and researchers, and the properties of the associated statistical tests have been explored in several studies.

However, the main contribution of the Lopez paper is its suggestion that alternative approaches to evaluating the performance of value-at-risk models are possible. For instance, another potential approach involves specifying a characteristic of value-at-risk models that a risk manager or a supervisor may be particularly concerned about—say, the model's ability to forecast the size of very large losses—and designing a method of evaluating the model's performance according to this criterion. Such approaches are not formal hypothesis tests, but instead involve specifying what is known as a "loss function," which captures the particular concerns of a risk manager, supervisor, or other interested party. In essence, a loss function is a shorthand method of calculating a numerical score for the performance of a value-at-risk model.

The results in the Lopez paper indicate that this loss function approach can be a useful complement to more traditional hypothesis-testing approaches. I will not go over the detail of his analysis, but the loss function approach appears to be able to provide additional information that could allow observers to separate accurate and inaccurate value-at-risk models. The important conclusion here is *not* that the loss function approach is superior to more traditional hypothesis-testing methods or that it should be used in place of these methods. Instead, the appropriate conclusion, which is spelled out in the paper, is that the loss function approach is a potentially useful supplement to these more formal statistical methods.

A further implication of the analysis is that the assessment of model performance can vary depending on who is doing the assessing and what issues or characteristics are of particular concern to the assessor. Each interested party could assess model performance using a different loss function, and the judgments made by these different parties could vary accordingly.

Before moving on to my concluding remarks, I would like to discuss briefly the material in the last section of the Lopez paper. This last section proposes a method for implementing the loss function approach under somewhat more realistic conditions than those assumed in the first section of the paper. Specifically, the last section proposes a method for calibrating the loss function in the entirely realistic case in which the "true" underlying distribution of profits and losses is unknown. Using a simulation technique, the paper demonstrates how such an approach could be used in practice, and offers some illustrations of the type of information about model accuracy that the approach could provide.

The material in this last section is a promising beginning, but before the actual usefulness of this application of the loss function approach can be assessed, it seems necessary to go beyond the relatively stylized simulation framework presented in the paper. The ideal case would be to use actual profit-and-loss data from a real financial institution's portfolio to rerun the experiments presented in the paper. Admittedly, such data are unlikely to be readily available outside financial institutions, which makes such testing difficult. However, the issue of whether the proposed loss function approach actually provides useful additional information about model performance is probably best assessed using real examples of the type of portfolio data that would be encountered if the method was actually implemented.

Concluding Remarks

In making a few brief concluding remarks about the lessons that can be drawn from these three papers, I would like to point out two themes that I see running through the papers' results. First, as discussed above, the papers highlight the point that in the world of value-at-risk modeling, there is no single correct way of doing things. The papers illustrate that the "right approach" often depends on the question that is being asked and the circumstances influencing the concerns of the questioner. The most important contribution of these papers is their helping us to understand what the "right answer" might be in certain situations, whether that situation is the presence of a fat-tailed distribution or different holding period horizons. Furthermore, the papers illustrate that in some situations, multiple approaches may be required to get a full picture of the behavior of a given portfolio or the performance of a particular model. In both senses, the three papers in this session have helped to provide concrete guidance on how to make such choices as circumstances vary.

The second theme that I see emerging from these papers is a little less direct than the issues I have just discussed. In my view, the papers reinforce the point that value-at-risk modeling—indeed probably most types of risk modeling—is a dynamic process, with important innovations and insights occurring along the way. It has been several years since I myself first started working on value-at-risk issues, as part of the original team that developed the internal models approach to market risk capital charges. Even at that stage, many financial institutions had already devoted considerable time and resources—over periods spanning several years—to the development of the models they were using for internal risk management. Despite this long history, these papers clearly indicate that serious thinking about value at risk is still very much a live issue, with innovations and new insights continuing to come about.

For that reason, no value-at-risk model can probably ever be considered complete or final; it is always a matter of keeping an eye on the most recent developments and incorporating them where appropriate. This is probably a pretty obvious observation to those of you who are involved in risk modeling on a hands-on basis. Nonetheless, it is an important observation to keep in mind as new studies emerge illustrating new shortcomings of old approaches and new approaches to old problems. These studies—such as the three presented here today—do not reflect the failure of past modeling efforts, but instead demonstrate the importance of independent academic research into the practical questions facing risk managers, supervisors, and others interested in risk modeling.