May 2006

Northfield News

A Newsletter for the Friends and Clients of Northfield Information Services

Inside this Issue

Special points of Interest:

New Methods for Dealing with Estimation Error in Optimization

By Dan diBartolomeo

In recent years, the issue of estimation error in portfolio optimization has deservedly received a great deal of attention from the quantitative investment community. The vast majority of our portfolio construction methods revolve around basic concepts of portfolio theory put forward by Markowitz (JOF, 1952). Unfortunately, the mean-variance optimization process has many underlying assumptions that are not wholly realistic. It assumes that we know with exact certainty, the mean returns, volatilities and correlations that our assets will have in the future. It also assumes that the future is just a single long period, during which the aforementioned exact information about asset return behaviors will never change. By making these assumptions, we oversimplify the optimization problem to the point where the actual results of the optimization process are likely to be seriously flawed and produce poor actual investment results.

Northfield has always paid great attention to this issue. Brief papers on this subject appeared among our research in the early 1990s, for example "Estimation Error in Asset Allocation" from 1991, **http://www.northinfo.com/documents/42.pdf**. In 1994, we introduced a Bayesian statistical technique into our asset allocation optimization. Over the

(Continued on page 4)

The How and Why (Not) of "What If" (Part 1)

by Dan diBartolomeo

Recently, Northfield has gotten a number of inquiries from clients requesting assistance with the use of factor models within the context of "scenario analysis." Some European regulatory authorities such as the Financial Services Authority in the United Kingdom have issued discussion papers in the past year, advocating "stress testing" and "scenario testing" as part of risk management procedures for financial institutions. In our May 2005 newsletter we included an article on how to stress test the values of risk measures coming out of our risk models.

Part 1 of this article means to describe some important methodological considerations around how factor risk models should be used in scenario analysis. In our next issue, Part II of this article will consider the investment policy implications of scenario analysis. We'll see that while scenario analysis is an absolutely vital risk management tool for some financial institutions, we'll see that for most of the asset management community, formulating investment policy on the basis of scenario forecasts is likely to be harmful to investors in the long run. For the purposes of this article, we'll use the term "stress testing" to describe the process of examining the underlying assumptions and parameter estimation of our risk management process. The term scenario analysis is used to describe

(Continued on page 8)

Recent and Upcoming Events

Northfield Investment Seminar for Private Taxable Wealth Wrap-up

The Trapp Family Lodge • Stowe, VT • March 26-27, 2006

Northfield held a two day invitation only seminar at the Trapp Family Lodge in Stowe, VT to discuss the challenges in investing private wealth form both an investment and operational perspective. The contents of a recently released book written by Dan diBartolomeo, and his two co-authors, Jeffrey Horvitz, Vice Chairman of Moreland Management Company and Jarrod Wilcox, President of Wilcox Investment entitled, "Private Wealth and Professional Management" - a handbook prepared for the Research Foundation of the CFA Institute were presented.

The event began on Sunday afternoon with a two hour seminar session which was followed by a cocktail reception, wine tasting and dinner. The sessions continued on Monday.

The presentations included: "A Conceptual Framework for Helping Private Investors." "Private Wealth and Taxation," "Organizing the Management Firm for Private Clients," "Individual Retirement Plans and Location," "On Concentrated Risk," and "Assessment and Benchmarking for Private Wealth."

The complete seminar proceedings have been posted to **http://www.northinfo.com/papersearch.cfm**. The book can be purchased from the CFA Institute in either print or online versions. Visit **http://www.cfapubs.org/doi/abs/10.2470/rf.v2006.n1.3933** for more details.

Northfield London Seminar Wrap-up

The Radisson SAS Hotel and Resort • London • April 24, 2006

The Northfield Spring 2006 Investment Seminar was held in London at the Radisson SAS Hotel and Resort on April 24, 2006. The purpose of the seminar was to highlight recent advances in analytical techniques for the investment industry to our growing number of European clients and prospects. The seminar was sold out with over 75 attendees.

Northfield President Dan diBartolomeo and Professor Gregory Connor of the London School of Economics were the presenters. The topics included; "Hedge Fund Performance Measurement," "Investment Management for High Net Worth Individuals," "Integrated Global Risk Models," "Active Risk Budgeting with Optimization Tools," "Addressing Total Uncertainty in Optimization through an Augmented Objective Function," and "Time Series Variation in the Level of Security Market Risk."

The seminar concluded with a well deserved post seminar reception. There was no cost to attend, however, donations to the Prince's Trust were strongly encouraged. The Prince's trust is a very worthwhile organization that makes a huge positive difference to the lives of many thousands of young people. Visit **http://www.princes-trust.org.uk** to learn more. The seminar proceedings will be posted soon to **http://www.northinfo.com/papersearch.cfm.**

Professor Gregory Connor Lecture Series

London School of Economics • April - May 2006

Through April and May, Northfield and UBS co-sponsored an evening lecture series on risk modeling by Professor Gregory Connor at the Financial Markets Group, London School of Economics. For five weeks on a Thursday evening, Greg presented chapters from his forthcoming book, with animated debate following. You can find out more at **http://fmg.lse.ac.uk/~connor/**. Northfield President Dan diBartolomeo presented at the final workshop. The topic was "Dynamic Volatilities and Correlations."

The lectures were well received and enjoyed by all. We highly recommend his forthcoming book, "Portfolio Risk Forecasting," to be published by Princeton University.

2006 Newport Annual Summer Seminar

Tennis Hall of Fame • Newport, Rhode Island • June 9, 2006

Tennis Hall of Fame

We are pleased to invite you and your colleagues to our annual summer seminar. The purpose of the seminar is to present recent research and technical advances to our clients and friends.

There will be five presentations; "Time Series Variation in the Level of Security Market Risk," "Addressing Total Uncertainty in Optimization: Parameter Uncertainty," "Life Cycle Funds and the Individual Investor," "Incorporating Higher Moments Into Financial Data Analysis," and "Active Returns from Passive Management: An Updated Look."

As always, our meeting date has been selected to coincide with the US Professional Championships of Court Tennis. Following the day's presentations, there will be a court tennis demonstration by Northfield President Dan diBartolomeo, and then a semi-final court tennis match. Court Tennis, or "real tennis" is the medieval sport that is the progenitor of all modern racquet sports.

After tennis on Friday evening, an oceanfront dinner party will be held at Johnnie's Atlantic Beach Club and Pavilion in nearby Middletown RI (Purgatory Road).

There is no charge for participation in any aspect of this event, however, we will be accepting donations on behalf of the Pine Street Inn, a Boston homeless shelter. The full seminar agenda and the online registration form have been posted to **http://www.northinfo.com/events.cfm**.

2006 Northfield Annual Research Conference

The Greenbrier • White Sulphur Springs, West Virginia • October 23-25, 2006

The Greenbrier

We are pleased to announce our 19th annual research conference at the Greenbrier, in White Sulphur Springs, West Virginia. The conference will officially begin on Monday, October 23rd and end on Wednesday, October 25th.

As is customary at Northfield events, a complete recreational and social calendar will accompany the working sessions. The Greenbrier is an award winning resort and national historic landmark offering over 50 recreational activities including 3 championship golf courses, indoor and outdoor tennis courts and a 38,000 square foot spa.

Further details and the complete conference agenda will be posted on the Northfield Website at **http://www.northinfo.com/events.cfm** as they become available. Contact Kathy Prasad at 617.208.2020, **kathy@northinfo.com** for more information.

(New Methods, Continued from page 1)

past few years, we had made the patented NFA parametric resampling methodology available in our Open Optimizer application.

We began a new round of research into these issues when we decided to create something of our own to replace the parametric resampling approach in the Optimizer. It was hoped that we would introduce the new technology in April of 2006. In the academic literature, there has been considerable attention paid to the issue of errors in estimation of the return distribution parameters, very little research of practical use in dealing with the distortions caused by the "single period" assumption. As we dug deeper into the issue, it became clear that much additional research was needed on this aspect of the problem. The theoretical aspects of this work have been largely completed, and we expect to introduce a set of new features into the Optimizer in the near future that will address estimation error in a more complete way than other current methods.

From a conceptual perspective, we would argue there are four (maybe even five) possible sources of uncertainty in an optimization. The largest source of uncertainty is the risk of financial markets and securities that we routinely address with our existing factor risk models. The second source is estimation error in our return forecasts. The third aspect is estimation error in our forecasts of portfolio risk. Finally, we have the distortions created by the assumption that the future is one long period. If we wanted to be extremely thorough we might also consider that transaction costs are not known exactly in advance, particularly for large block trades.

If we confine ourselves to just four sources of optimization uncertainty we must also realize that these different sources can interact. We can think of a four by four matrix of uncertainty sources. To the extent that this matrix is symmetric about the diagonal, we have potentially ten separate contributions to the overall uncertainty of the problem. So far, no one has proposed methods that attempt to deal with all ten.

There have been three broad methods proposed for dealing with particular subsets of the overall uncertainty problem. The first approach of "Bayesian methods" retains the normal optimization process, but makes adjustments to the input numbers that are meant to compensate for the uncertainty in the estimation in return distribution parameter estimation. One example of Bayesian approaches being applied solely to expected returns for assets include the Jorion (JFQA, 1986) method that is used in our ART asset allocation software. Another is the Black-Litterman model (JFI, 1991) that is widely used in asset allocation. At the security level, the "alpha scaling rule" from Grinold (JPM, 1994) is very popular, although it has some very restrictive assumptions that practitioners often incorrectly ignore. On the risk side, Ledoit and Wolf (JPM, 2004) provide a method for adjusting asset covariance estimates. The advantage of Bayesian method is that they are very numerically efficient in that the traditional optimization process is used. Only the input data values are adjusted.

The second approach used to deal with estimation error has been to add additional terms to the optimization objective function to explicitly address the new sources of uncertainty. Ceria and Stubbs (Axioma, 2004) dealing with expected return uncertainty only by assuming a confidence interval on security alpha estimates while assuming the risk estimates are exactly correct. A similar paper by Goldfarb and Inyegar (MOR, 2003) assumes min/max bounds on the alpha estimates but without distributional assumptions. They also assume an exactly known factor covariance matrix but with possible errors in factor exposures. This method also assumes that risk errors and return estimate errors are orthogonal and hence there is no need to worry about their interaction. This is probably a very unrealistic assumption for most equity strategies. Finally, Halldorsson and Tutuncu (JOTA, 2003) assume min/max bounds on both the alpha and covariance estimates, again without distribution assumptions. These methods are sometimes referred to as "robust optimization" as they can often be used to solve for the optimal portfolio under a "worst case scenario."

One potential concern about "worst case" optimization is that dealing with the worst case makes a great deal of sense in a situation where even a single bad return result can drive the investor bankrupt such as a highly leveraged hedge fund. On the other hand, the vast majority of investors who are not leveraged will survive an infrequent very negative return outcome to invest through other periods of more favorable returns. By always optimizing for the worst case, the investor creates a pessimistically biased investment policy on an ongoing basis. Clearly, this issue might be viewed as one aspect of the distortions caused by the "single period" assumption.

A principle drawback of both Bayesian and "robust" methods is that we have to have information about the nature and magnitude of our likely future forecasting errors. Just to deal with potential errors in forecast returns we have to estimate the covariance matrix of our forecasting errors. It is not easy to distinguish between forecasting errors and random events from a distribution of returns that we have identified, as described by Huber (Northfield Annual Conference, 2001).

(Continued on page 5)

(New Methods, Continued from page 4)

In addition, for securities problems the number of securities for which we have forecasts is very likely to exceed the number of observations over which we can measure the effectiveness of our forecasts. Such a situation means that the covariance matrix of forecasting errors may not be positive definite, leading to the need to build a factor model of forecasting errors. In that most practitioners purchase factor models of the unconditional security returns from specialized vendors such as Northfield, we think it unlikely that asset managers will be willing to undertake to build factor models of their forecasting errors. To deal with this dilemma, the implementation of most Bayesian and "robust" methods involve a variety of simplifying assumptions as described in Idzorek (Zephyr Associates, 2003). Unfortunately, the simplifications are sometimes gross enough to render the benefit of the method questionable or mathematically equivalent to a trivial adjustment of a regular optimization (e.g. multiply the regular risk terms by 1.5).

The third method of dealing with parameter estimation error is resampling. Northfield's early research in this area involved bootstrap resampling, as first proposed in Bey, Burgess and Cook (University of Tulsa, 1990). Resampling involves combining Monte-Carlo simulations of potential errors in the input parameter estimates into the optimization process in order to produce an entire "family" of optimal portfolios. Once you have the entire family of portfolios, various schemes have been proposed to pick the "best" from the family of optimal portfolios, or to average the family of optimal portfolios in a particular way, such as Michaud (Harvard University Press, 1998). One drawback of resampling is that it is very numerically intensive, requiring long computation times to derive a stable solution for large problems. Other technical criticisms of resampling that involve potential biases in asset weights appear in Scherer (FAJ, 2002). A major advantage of resampling is that it provides a very rich set of information as to the statistical significance of the differences between portfolios. We will see in our discussion of the single period assumption that such information is an important component to efficient management of transaction costs.

All three approaches to parameter estimation error are different routes to the same destination. Basically, we are trying to find the most efficient way to "over-diversify" our portfolio, in order to compensate for the fact that the information on which we base our decisions is frequently wrong. All three have strengths and weaknesses, both theoretical and practical.

In the months ahead, our Open Optimizer will be equipped with three functions that can be used separately or jointly to address estimation error. To deal with estimation error in expected returns (alphas), we will provide a function that will automatically rescale the user's alpha inputs in a way that we believe avoids restrictive underlying assumptions of the Grinold scaling procedure. The user's input alphas will first be mapped into the user's choice of a normal or fat-tailed distribution. Some investment professionals believe that the information content of most investment strategies is concentrated only at the extremes of the distribution so mapping to a fat-tailed distribution is more appropriate. A statistical procedure is then run to relate the user's rescaled alphas to the factor loadings of the risk model that is being used. Portions of the user's alpha that seem related to a particular factor in the risk model are converted to a factor alpha that is scaled in a way that is consistent with expected volatility of returns to that risk factor. The portion of the user's alpha that seems uncorrelated with the risk model is scaled by the idiosyncratic risk of the security involved. The theoretical basis for this approach comes from Bulsing, Sefton and Scowcroft (UBS, 2003).

The second new function which will address uncertainty in the risk estimates will be Bayesian in spirit, but will actually involve a change in the optimization process. During an optimization, the confidence interval on the risk measure (i.e tracking error) will be repeatedly estimated from the closed form calculations available in the statistics literature for the standard error of a standard deviation and the standard error of a correlation coefficient. The additional uncertainty arising from estimation error in the risks will then be directly incorporated into the optimization objective as a new term. The result of this process will be to increase the perceived overall risk of the portfolio, make the risks of individual securities more similar to the average of all securities, and make the pair-wise correlations of securities closer to the average of all pair-wise correlations, as is consistent with Ledoit and Wolf. To the extent that estimation uncertainty on portfolio risk arising from time series trends in risk levels rather than sampling error, we are considering a minor adjustment to our model estimation methods.

The final new function has been the result of recent original research at Northfield. We are attempting to address the distortion of optimization caused by the assumption in mean-variance optimization that the future is one long period during which our expectations about the distributions of securities will not change. Essentially, we assume the initial portfolio has a particular distribution of return, and

(Continued on page 6)

(New Methods, Continued from page 5)

that the optimized portfolio has a different distribution of return that we unambiguously like better (i.e. higher returns or lower risk). This assumption does no harm if transaction costs are free, since we can revise our portfolio at no cost as our beliefs change. However, if transaction costs are non-zero this is clearly a problem.

Many approaches have been proposed for full multi-period optimization starting with Mossin (1968). Cargill and Meyer (AFPF, 1987) focused on the risk side of the multi-period problem, while Merton (Blackwell, 1990) introduces continuous time analog to MVO. Pliska (Blackwell, 1997) introduces a time analog to MVO and Li and Ng (MF, 2000) provide a framework for multi-period MVO using dynamic programming. Sneddon (Northfield Client Conference, 2005) introduces a closed form solution to the multi-period optimization including optimal turnover, under certain simplifying assumptions. However, the sort of parameter estimation error problems discussed above have severely limited the practical implementation. For most investment practitioners, it is hard enough to make reasonable forecasts of average expected returns. It would be vastly more problematic to have to accurately predict expected returns, time period by time period.

Other authors have proposed methods to try to build some intelligence into the portfolio rebalancing process to avoid unnecessary trading costs. Rubenstein (JPM, 1991) examines the efficiency of continuous rebalancing and proposes a rule for avoiding spurious turnover. Kroner and Sultan (JFQA, 1993) proposed a “hurdle” rule for rebalancing currency hedges when return distributions are time varying. Engle, Mezrich and Yu (Smith Barney, 1998) proposed a hurdle on alpha improvement as the trigger for rebalancing.

Markowitz and Van Dijk (FAJ, 2003) propose a rebalancing rule designed to approximate multi-period optimization, but argue it is mathematically intractable (at least in closed form). However, numerical methods have also been applied to this purpose. Bey, Burgess, Cook (1990) use bootstrap resampling to identify “indifference” regions, along a fuzzy efficient frontier. Michaud (1998) uses resampling to measure the confidence interval on portfolio return and risk to form a “when to trade rule”. Elaborated upon in Michaud and Michaud (New Frontier Associates, 2002) and patented.

Our approach to this issue is to consider a “probability of realization” with respect to improvements in expected utility. The certainty of our preference for the optimized portfolio is based on looking at the population statistics of the two distributions. However, if our time horizon is finite, we really should be looking at the sample statistics of the distributions, not the population statistics. Over any finite interval, the likelihood that the higher utility of the optimized portfolio will actually be realized is somewhere between 50% and 100%. When we trade off improvements in utility versus transaction costs our amortization constant should reflect the probability of realization for gain in expected utility as less than one.

The probability of realization is an approximate function of the tracking error between the initial portfolio and the optimal portfolio. The third new feature of the Optimizer will compute this tracking error value and calculate the probability of realization at each step of the optimization. The user’s input of the amortization constant will then be dynamically adjusted to ensure the most efficient trade-offs between expected changes in utility and transaction costs.

Overall, we expect these three new functions to provide significant improvement in the stability and intuitiveness of optimization results, while making the trade-offs between returns, risks and trading costs in a more efficient fashion.

Northfield Staff Speaking Engagements

On March 14, Dan diBartolomeo presented "Estimation Error in Optmization Arising from the Single Period Assumption." at the QWAFAFEW meeting in Boston. On April 11, Dan did a related presentation at the Alpha Strategies UBS Conference at Duke University on "Dealing with Total Uncertainty in Optimization through an Augmented Objective Function."

On May 2, Dan spoke at the Separately Managed Accounts Conference in Boston where he presented "Realistic Tax Optimization."

On May 23, Dan presented "Optimal Algorithmic Trading" at a CHARISMA conference in London sponsored by Brunel University.

Technical Support Tip: How to Use The Excel Run0 Add-In — NER0

By Christine Milne

NER0 allows a user to run portfolio risk analysis and "what if" scenario analysis quickly and easily without leaving the familiar Excel environment. All set-up and output reports are completed within Excel. The add-in is not a replacement but a useful tool on top of the Northfield Optimizer running in the background.

After adding the add-in to Excel, you can see these six icons relating to NER0. The problem set-up is run in the "Analysis Wizard" – the first icon on the toolbar.

Information on the location of the risk model data files and the Optimizer software need to be filled in, along with references to your Excel workbook where the portfolio and benchmark information is listed. These can be in percent, shares, market cap or equally weighted.

The Analyzer allows you to use either the Northfield supplied industries and sectors, or your own industry classification, or a combination of the two. You can also store the classification in files outside the workbook and reference those, or keep the classification tables in the workbook.

Up-to-date prices can be fed via a link into the Excel workbook and can be updated automatically for use in the analysis.

The Composite Assets utility allows the user to create a fund-of-funds set-up. Each sub-fund may be represented in the portfolio as a single "line item" by creating a new composite asset sheet in your Excel workbook – the fourth icon on the NER0 toolbar.

Quick "what-if" analysis can be run by changing portfolio holdings in the Excel sheet and re-running the analysis. The output reports are in Excel format. You can save them as new tabs in your existing workbook with the portfolio and benchmark data, or as a new Excel file.

The Output Reports

- Summary – Tracking Error, Stock Specific and Factor Risk, Total Risk, Portfolio beta. The report also includes some VaR numbers, calculated simply from total risk, and a summary of any exceptions found.
- Exceptions – Any missing stocks, with their name and weight, found in the portfolio or benchmark.
- Risk – This is similar to the risk decomposition report in the Northfield Optimizer stand-alone system. It also lists the top 5 and bottom 5 risk factors by variance contribution. A small sample is shown below.
- Stock – shows the stock specific risk information for each security in the portfolio, as in the Stock Marginal Contribution report in the main system.
- Penalties – the weight in each of the industries and sectors used – Northfield or user-defined.
- Main Table – this contains all the underlying data about each security represented in both the portfolio and the benchmark – exactly as seen in the stand-alone.

Northfield INFORMATION SERVICES, UK Ltd.

Portfolio: Global : Benchmark: FTSE All world : Report Created 09-Ja

Summary	
Factor Tracking Variance	4.6401
Stock Specific Tracking Variance	6.0204
Total Tracking Variance	10.6605
Tracking Error	3.265
Total Risk of Portfolio	13.0265
Total Risk of Benchmark	14.1011
R-Squared	0.9489
Portfolio Beta	0.8999

Greatest Effect on VarContr

Factor	#	ActiveExp	FactorVar	%VarContr
INDUSTRIAL SECTOR	6	-0.0478	206.044	20.65
SIZE	15	-0.0493	61.9756	19.71
DEVELOPING MARKET	14	-0.0862	73.7876	14.85
BLIND FACTOR 2	18	-0.1918	12.9848	12.94
EURO	90	-0.0108	77.4917	5.32

The Excel Run0 process can also be run as a batch. The Northfield Excel Batch Application – NEBA – retains the flexibility of NER0 but allows for regular risk analysis of multiple portfolios, all within Excel, with less manual intervention.

The last icon in the NER0 toolbar shows the "Function Crib Sheet." This sheet can guide VB developers in automating the execution of the add-in from their own code, for full integration within Excel.

- For further inquiries contact Technical Support in Boston: **support@northinfo.com** or call +1 (617) 208-2080.
- European clients can contact Christine Milne in London: **christine@northinfo-europe.com** or call +44 (0)20 7801 6250.
- Clients in Asia can contact Nick Wade in Tokyo: **nick@northinfo.com** or call +81 3 5403 4655.

(How and Why, Continued from page 1)

the process of considering the "what if" consequences of one or more specific scenarios such as "oil prices rise 50% over the next year, or government bond interest rates rise 300 basis points in the Euro-zone."

To do scenario forecasting based on a factor risk model, we must begin with having a risk model that is composed of factors that are both relevant and intuitive in this context. For example, if we're going to consider what would happen if oil prices rise 50% in the next year, we need to have a model that has oil prices as one of its factors. This means that endogenous models of security behavior that rely on security characteristics (e.g. accounting ratios such as P/E of a stock) are going to be much more limited in their usefulness than models that are based around the macroeconomic influences on security returns. It should be noted by the proponents of endogenous models (of which Northfield certainly is in other contexts) that the original rationale for endogenous models was to provide a way to link security behavior to macroeconomic factors while allowing for time varying-elasticity. This issue is made clear in a paper by Barr Rosenberg and James Guy (Financial Analyst Journal, 1976). Similarly, purely statistical models that describe risk only as a set of orthogonal latent factors that are without immediate economic interpretation are equally unattractive.

All linear factor models express expected asset returns as the product of a factor exposure (i.e. beta coefficient or elasticity) and a factor return. In the scenario analysis context, our hypothetical events are single realizations of the factor returns. If we want to know what will happen to a portfolio if oil prices rise 50% over the next year we simply need to determine the portfolio exposure coefficient to changes in oil prices (hopefully it's a factor in our model) and multiply by 50%. This piece of the arithmetic is very simple indeed, and can be done automatically in the Northfield optimizer by entering the factor return forecast as a "factor alpha" in the .MDL file.

The probability of any "point forecast" scenario is zero. The likelihood that oil prices will rise exactly 50.000000000% over the next year is essentially zero. For scenario analysis to be sensible, the expected outcomes must be expressed in the form of a range, such as that oil prices will rise between 45% and 55% over the next year. Given a probability distribution of oil price changes, we can calculate probability that a rise in oil prices will be greater than 55% and the probability that a rise in prices will be more than 45%. By subtracting the former from the latter, we have the finite probability associated with the range of outcomes in our scenario.

For a "what if" scenario to be useful, it has to be plausible. Often investors try to use scenarios that are implausible because they represent partial equilibrium solutions to a full equilibrium world. They assume a large impact in one factor while concurrently no effects in other factors of the model. If we consider what would happen if oil prices rose 50% over the next year, we would also realize that such an economic event would mean that any other factors in our model that are correlated with oil prices would likely experience their own factor returns. For example, firms in the energy sector would have different profitability than energy consuming firms, currency exchange rates between oil exporting and oil importing countries would be likely to change. Even the return relationship between large and small capitalization stocks would likely be impacted, as the capital intensive nature of oil exploration means that almost all oil is produced by large firms. As such, a plausible scenario forecast usually implies calculating an expected range of outcomes for every factor in the model.

The scenario must be mathematically coherent, which is a non-trivial exercise in conditional probability. The expected outcome for each factor must be coherent in terms of the expected outcome of every other factor, not just the factor or factors for which we intend to explicitly forecast outcomes. Let's assume that we have a 20 factor model, of which oil prices listed as factor #1. If we hypothesize a 45% to 55% rise in oil prices, we must ensure that our expected range outcome for factor #2 is consistent with the correlation between factor #1 and factor #2. Our expected range outcome for factor #3 must be consistent with the correlation between both factor #3 and factor #1, and the correlation between factor #3 and factor #2. If we have a 20 factor model, we have 190 cross-relationships to consider. For a 50 factor model, the number is 1225.

One way to construct coherent forecasts is to use a "chained Bayesian estimation" for the expected range outcome of each factor. Here is an example using a three factor model estimated over 60 data periods.

	Volatility	Correlation		
Factor 1	30	1		
Factor 2	20	-.10	1	
Factor 3	10	-.39	-.25	1

(Continued on page 9)

(How and Why, Continued from page 8)

Now let us assume we hypothesize a 60% to 90% return to factor 1 over one time period which ranges between a two and three standard deviation event (60/30 and 90/30). As such, the most likely range of outcomes for factor 2 is from -4 to -6 (2 * -.1 * 20 to 3 * -.1 * 20).

We now have to consider factor 3. Given our expectation for factor 1, we get expected range of -7.8 to -11.7 (2 * -.39 * 10 to 3 * -.39 * 10) for factor 3. Given our expectations for factor 2, we get a range of .5 to .75 (-4/20 * -.25 * 10 to -6/20 * -.25 * 10) for factor 3. We now can combine the two forecasts as:

Lower Value for Factor 3 = w_1 (-7.8) + w_2 (.5)
Upper Value for Factor 3 = w_1 (-11.7) + w_2 (.75)

One way to determine what values we should use for w_1 and w_2 is to invoke Bayes Theorem that states the best weighting of two forecasts is proportional to their inverse variances. This method is relatively easy to calculate in a spreadsheet even for a large number of factors. However, it should be noted that this particular procedure provides an order dependant solution. It is presented as only one of many procedures that can provide coherence in the scenario vector.

w_1 = 1/Var(1) / (1/Var(1) + 1/Var(2))
w_2 = 1/Var(2) / (1/Var(1) + 1/Var(2))

The reason these two forecasts don't match is the observed correlation structure between the factors. The variance of a correlation coefficient (ρ) is given by:

$$Var(\rho) = (1 - \rho^2)^2 / (n-2)$$

Variance 1 = .0123
Variance 2 = .0152

w_1= (1/.0123) / (1/.0123 + 1/.0152) = 81.3 / 147.1 = 55.25%

w_2 = (1/.0152) / (1/.0123 + 1/.0152) = 65.8 / 147.1 = 44.75%

Lower value for Factor 3 = .5525 (-7.8) + .4475 (.5) = -4.09
Upper value for Factor 3 = .5525 (-11.7) + .4475 (.75) = -5.26

Unfortunately, our typical estimates of correlation coefficients are likely to be suspect from the start. Even if we are able to mathematically construct a coherent scenario forecast vector, it is unlikely to be reliable during the sort of extreme events that are of interest. Financial institutions are using scenario analysis to consider the possible impact of rare, extreme events on their portfolios. Being rare events, the sample of past events from which correlation estimates may be obtained is likely to be very small. In addition, there is a well recognized phenomenon wherein the expected correlation of market factors (but not other factors) may be expected to rise during times of economic upheaval. There is a simple but not widely known rationale for this effect. Let us assume the conventional view that the value of every investment asset is simply the presented value of future expected cash flows discounted at some rate that reflects the risk (uncertainty) of the cash flows. We presume that investors have set the discount rate in consideration of their own level of knowledge of the economic circumstances of the asset. If some "good news" (unanticipated information) arises that suggests that future cash flows will be larger or that future discount rates will be lower, investors adjust their forecasts accordingly. However, the investor's level of confidence may be reduced because they were "surprised" by the new information (their perception of uncertainty increased). This causes the discount rate to rise, muting the increase in the value of the asset. On the other hand, if the news is "bad," the expectation of future cash flows may decline or the discount rate may rise. However, the decline in confidence based on the arrival of news drives the discount rate up separately. These two effects reinforce each other causing a stronger decline in value of the asset. The asymmetric nature of this effect causes the correlation between assets to be higher during periods of price decline. For more information see Brown, Harlow and Tinic (Journal of Financial Economics, 1988) and Vuolteenaho (Journal of Finance, 2002). The correlations between factors in Northfield models can be edited in the .COR file, however care must be taken that the revised correlation structure for the factor covariance matrix does not result in a factor matrix that is not positive semi-definite, as irrational negative risk forecasts may result.

Some extreme market events may seem plausible (or have even been observed) but are so unlikely that we must consider the potential for non-parametric outcomes. For example, the one-day stock market crash in October, 1987 reduced US stock prices an average of about 25% in a single day. Given the observed volatility of the US stock market averages to that point (and since) this was a more than twenty standard deviation event. Under the normal distribution assumption the likelihood of this event is zero to more than 30 decimal places, and yet it happened. However, if we assume that we don't know the distribution of returns, we can use Chebyshev's inequality to consider the maximum likelihood of this event:

$$Prob(<Z) = 1 - Z^{-2}$$

This suggests that the likelihood of an event being more

(Continued on page 10)

(How and Why, Continued from page 9)

than 20 standard deviations if we know nothing about the distribution is about one in four hundred. Given that there are about 250 trading days each year, and hundreds of financial markets around the world, we should be seeing these huge kinds of "crash events" on a very frequent basis if the distribution of asset returns is unstable or unknowable. As this high frequency is not observed, we must conclude that markets rarely break down into chaotic states.

One useful exercise in this regard is to consider the level of volatility that would correspond to a particular parametric probability. For example, we know that under the normal distribution assumption, a three standard deviation event covers roughly 99% of a distribution. Under Chebyshev, a ten standard deviation event is required to cover the same amount of an undefined distribution. So if we want to consider risk under the hypothetical likelihood of a collapse into chaos as 1 in a 100, we can just multiply the variances of our factors by about eleven ($10^2/3^2$). This is easily done for Northfield models by editing the factor variances in the .MDL file.

It should be noted that for every factor model of asset behaviors there exists a mathematically equivalent full covariance matrix. Changing factor correlations or factor variances can lead to failure of positive semi-definiteness in the equivalent full covariance matrix, leading to the previously noted irrational results.

Mike Knezevich Joins Northfield

Mike Knezevich will be joining Northfield effective June 1, 2006. He will be taking over the management of Northfield's Technical Support Department and will be responsible for organizing on-going training programs for staff, clients and partners.

Mike previously spent seven years with Barra, with roles in both technical support and client consulting. He also presented at Barra's Equity Portfolio Management Conferences, conducted numerous workshops and served on the steering committee of the San Francisco chapter of QWAFAFEW.

In addition to his experience in the field, Mike holds a Masters degree in Applied Economics from the University of Michigan.

Starting in June, Mike can be reached at **mike@northinfo.com**.

If you have any suggestions of what you would like to see covered in upcoming issues, please e-mail your ideas to staff@northinfo.com

Boston Office
184 High Street, 5th Floor
Boston, MA 02110
Phone: 617.451.2222
Fax: 617.451.2122
Sales: 617.208.2050
Tech Support: 617.208.2080

London Office
Shakespeare House
168 Lavender Hill
London, SW11 5TF
Phone: +44-(0)-20-7801-6260
Fax: +44-(0)-20-7801-6261

Tokyo Office
Shiroyama JT Trust Tower
4-3-1 Toranomon
Minato-ku
Tokyo 105-6016
Phone: +81 (0)3 5403 4655
Fax: +81 (0)3 5403 4646

Northfield INFORMATION SERVICES, INC.

Northfield News is a publication of Northfield Information Services, Inc., 184 High Street, 5th fl., Boston, MA 02110. If you have any questions or comments regarding the content of this newsletter, please call us, or e-mail us at staff@northinfo.com, or visit our home page at http://www.northinfo.com