NORTHFIELD | modeling financial markets worldwide

Closed Form Methods to Address Higher Moments without Monte Carlo Repricing

Dan diBartolomeo

Asia Seminar Series

October/November 2017

Introduction

- This presentation provides a new approach to risk assessment from numerical simulations. As risk-related regulation extends from commercial banking to other parts of the financial services industry, risk assessments arising from "stress tests" and "scenario analysis" have become more widely discussed and implemented.
- Unfortunately, banking methods for this kind of risk assessment are often counter-productive for long term investors who are not levered, as compared to somewhat different approaches used by actuaries.
- To resolve the shortcomings of numerical methods we have built a new process, extending the approach suggested in Meucci (2008) which combines Monte Carlos simulations with the flexibility to overlay complex explicit scenarios. The analytical output of the process is a robust representation of the distribution of possible outcomes, while being consistent with any mathematically feasible "stress scenario".

Is Risk Now?

- For financial intermediaries such as commercial banks that are generally highly levered, the conception of risk is about ***solvency.***
  - The liabilities of the entity are current and subject to immediate call.
  - The objective is to make as much money as you can each day while limiting the probability of going broke to some acceptable level so you likely to be in business tomorrow.
  - diBartolomeo (2010, JOI) found that the typical implied half-life of a financial firm is on the order of 20 years, but much shorter (e.g. 8 years) on a revenue weighted basis.
  - Risk is measured in value units (e.g. VaR, CVaR) and **material effort is spent to get accurate prices for assets**

- Geared hedge funds are in an even worse position as margin loans are at call, *and* prime brokers don't care about trading costs in a forced liquidation as long as they are whole.

Or Later?

- Sovereign wealth funds are the opposite end of the spectrum. You can't go broke if you don't owe anybody any money.
  - For long term unlevered investors, the key risk is ***the estimated variance of the future return stream*.**
  - $1 invested for 50 years at a fixed 8% annually produces $46.90
  - $1 invested for 50 years at an average 8% annually with a standard deviation of 20% produces only $17.42
  - If you want to get fancy you can adjust the volatility to account for skew and kurtosis. See Wilcox (2000).
  - *Precision in current asset pricing is largely irrelevant.*

- Pension funds and insurance companies have only actuarial liabilities, which are the present value of expected future liabilities.
  - The liabilities are not subject to immediate call.

Dealing with "Now" Risk

- Existing processes have worked in one of two ways. The first is Monte Carlo simulations of asset prices where there is random sampling from a parametric or empirical distribution to get a range of possible outcomes. Risk assessment are based on the lower tail of the portfolio value distribution.

- The second process is to forecast a single return value for a set or series of specific exogenous scenarios. For example "What will be the % change in the value of my portfolio *(notice it's a single point value)* if interest rates go up 2%? and oil prices go down 30%".
  - It is argued that if we look at enough different "stress scenarios" we can gain an intuition about "worst case" outcomes. Unfortunately, the way most stress scenarios are formulated, their actual probability of occurrence is very, very small. Investors predicating investment strategy on such low probability outcomes end up with portfolios that are materially sub-optimal in the vast preponderance of situations.

Issues with Stress Testing and Scenarios

- Any ***scenario should be mathematically coherent, which is often a non-trivial exercise in conditional probability.***
  - A partial equilibrium solution to a full equilibrium world
  - Alternatively, 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 50 factor risk 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.
  - For our 50 factor model, there will be 1225 relationships

- It's hard to simulate events that have never *yet* happened like the 1987 crash, "Flash Crash", or the August 2007 liquidity problem.
  - Chebyshev's inequality comes in very handy
  - 3, 800, 18

A Numerical Method We Like

- Since the future is unlikely to be exactly like the past, we should be interested in whether the sequence of past events we have lived through is typical or unusual, given available history.
- As described in our June 2013 newsletter, our preferred numerical simulation method for exploring the distribution of a set of outcomes is "bootstrap" resampling.
  - We can use bootstrap methods to answer the broader question of "what if things had been different" but drawn from a similar distribution. set of factor return experiences.
  - However, rather than using the actual sequence of events (e.g. factor returns) we will be using many sequences of randomized events drawn from an historic set of experiences.
  - In essence, we will assume that the future may follow *any one of an infinite number of paths that we might have experienced in the past.*

Basic Bootstrapping in Brief

- Mechanically, the process is easy and very, very fast.
  - We use any of our risk models to get the factor profile of the portfolio
  - Let's assume we want to make a period by period forecast of the return distribution for the next 12 months and that we have a 240 month history of factor returns.
  - To create our first sequence of synthetic history as our forecast, we draw random number N between 1 and 240. The factor returns for month N are now the first month of our first sequence of our forecast factor events. If we repeat the process 12 times, we will have one full sequence of potential future events.
  - Note that since the choice is random each time, not only is the order of events randomized but some observations may be omitted and some observations may be repeated more than once.
  - The probability of choosing each observation is 1/N at each moment
  - For each path we estimate the return on the portfolio for each month, assuming a random draw from the distribution of idiosyncratic risk.

Let's Add a Little More Realism

- Given the simple computational process, we can repeat this entire procedure many thousands of times in a few minutes to *produce a very robust estimate of the future distribution.*
  - At each point in each path, we can estimate calculate the estimated mean, volatility, cumulative returns, maximum drawdown, etc.
  - We can also analyze the cross-section of paths at each moment in time to describe the period by period distribution of the statistics.

- We can also account for serial properties
  - If we believe that asset returns are serially correlated randomizing the sequences will fail to represent this aspect of the data.
  - To address this we can follow the procedure above, but build our sequences of future events from blocks of multi-month periods so as to capture most of the dependence from one month to the next.
  - The length of the blocks would relate to the number of lags in an autoregressive process.

"Even God Cannot Change the Past": Agathon

- So far, we are just sampling from an empirical distribution.
  - Any of the paths we generate are plausible
  - All of the statistical relationships between factors would hold together
  - We can see how typical or atypical the actual sequence of history within the range of the paths we generate.
- The results are not a lot different than if we did Monte Carlo simulations that incorporated the higher moments and serial properties of the expected distribution.
  - But the use of an empirical distribution at least ensures that effective distribution is realistic (it did actually happen).
- But a lot of things have changed since 450 BC. Even if we can't change the past we can pretend that we can.

Let's Try Playing God

- In terms of our risk simulations what we really want is to combine the rich distributional information of a numerical simulation with the "intuitive" nature of a set of explicit scenarios. Such a combined process is described in Meucci (2008)
  - Attilio Meucci and Dan diBartolomeo were part of a session on this at the Society of Actuaries conference in March of 2005.
- It is possible to "stress test" the projections by filtering the set of past observations from which our projected sequences are built.
  - We could include only months from periods of economic recession, or had rising interest rates or include only months that were perceived as particularly volatile. *Meucci refers to this as "crisp conditioning".*
  - If we have a "seed sample" of N observations and we filter out P observations, the probability of any observation being drawn to fill a position in a particular path is *1/(N-P) or zero.*
  - Now we have dense simulated data in both time series and cross-section, *conditional on the stressful or benign filtering.*

Now Let's Get Flexible

"Who is to say that truth is in the crystal and not in the mist?"
Kahlil Gibran

- We can set up a more flexible process where the probability of any particular observation being drawn for inclusion in a bootstrap path is explicitly defined by the user.
  - Instead of the probability of inclusion being 1/(N-P) we can choose a vector of explicit values for each observation.
  - Each probability p_t must be between zero and one
  - The sum of all values of p_t must equal one
  - Meucci refers to this as *flexible conditioning*
- While obviously feasible, it is not immediately obvious how an investor would decide what values should populate the probability vector.

Scenario Based Flexible Conditioning

- We would like combine bootstrap simulations with explicit scenarios.
  - We can build the probability vector for inclusion of observations so as to fulfill the some explicit scenario within a confidence interval.
  - For example, we could say "Do a bootstrap simulation where ***on every path***, the 10 year interest rate rises between 297 and 303 basis points, and oil prices decline 11 to 13% over a 12 month interval".
  - Observations with increasing interest rates and declining oil prices get more weight and vice-versa.
  - We can specify any variable for which data exists for the seed sample. We are not limited to the factors of an underlying model.
  - We can generate several different scenarios and select the number of paths to be run for each to represent weights. We just do our cross-sectional statistics on the aggregated paths.
  - *The cross-sectional variation in the paths is an implicit measure of the likelihood of the scenario.* If all paths are similar we know that the only a small fraction of all feasible paths fulfill the scenario.

Shifting Time Scale

- It is also possible to shorten the apparent time length of a path to accommodate different forecast horizons irrespective of the time scale of the original factor return observations. Regulators often want scenarios to play out over short periods like one day or even instantaneously.
- Given an assumption of a particular fat tailed distribution, it is possible to functionally compress factor returns that have been observed over a given interval (e.g. month) into shorter intervals such as days, hours or minutes.
- For example, it is widely documented that high frequency financial return data has strong "fat tailed" characteristics.
- To convert an empirical distribution of monthly data to daily data, we assume the distribution has changed from normal to a T-5 distribution.
- For intraday horizons we assume no knowledge of the distribution and use the Chebyshev boundary probabilities. At ultra short horizons (seconds) we can invoke a stable Paretian distribution.
- See diBartolomeo "Fat Tails, Tall Tales, Puppy Dog Tails", Professional Investor (2007) for details

Calculating the Probability Vector

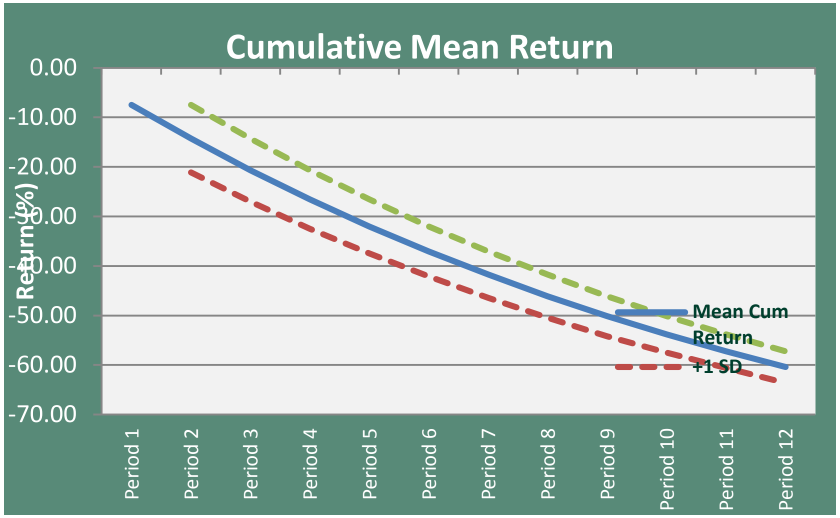
- Figuring out what probability vector best expresses a given scenario is an optimization problem. We want to find the vector of probabilities such that:
  - All values of p_t are equal to or greater than zero
  - All values of p_t are less than one
  - All values of p_t sum to one
  - The attributes described in the scenarios are fulfilled within the prescribed ranges.
  - We preserve maximum randomness by minimizing the sum of the differences (absolute or squared) between each value of p_t and 1/N

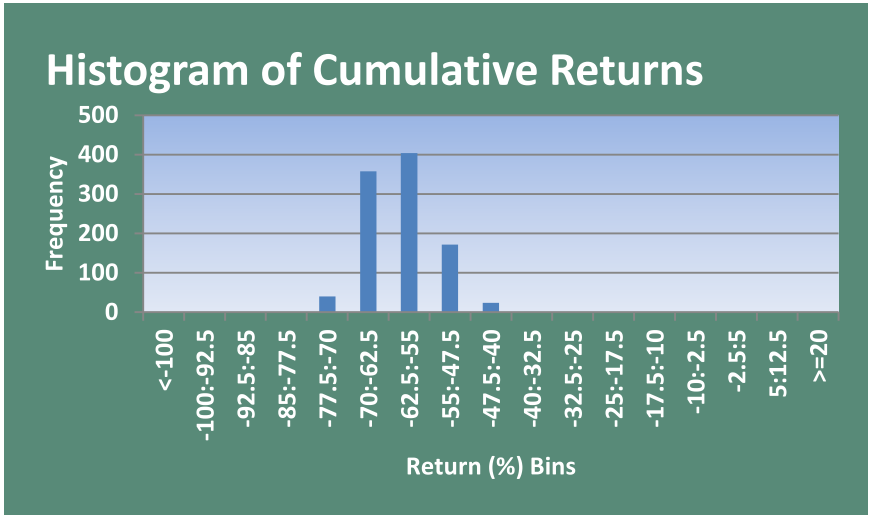
- If you use the Northfield Optimizer to solve the problem, it will also come up with the closest possible probability vector if the scenario is infeasible within the range of outcomes of the seed sample.

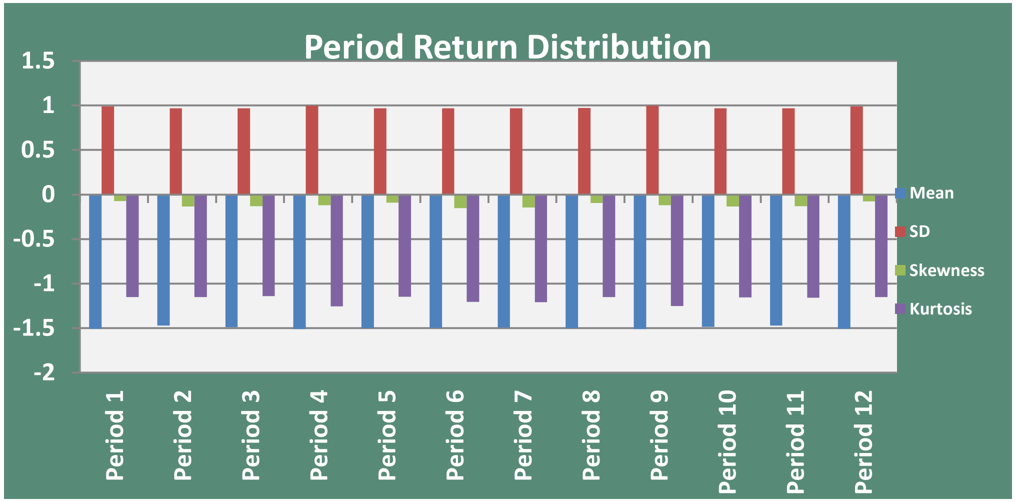
A Quick Review

- We use our regular risk models to get a time invariant representation of the portfolio and/or liabilities.
  - Our SIENS process can be used incorporate complex derivatives
  - Unlike normal risk model usage, we represent not only variation around the mean, but uncertainty of the mean return.
- We use bootstrap resampling to compile a wide range of alternative simulations of history drawn from a seed sample of historical data.
  - The probability of any observation being included in a simulated path can be conditioned by filtering (crisp) or by a probability vector (flexible).
  - The path driven simulations provide a rich set of statistical metrics in both time series and cross-section.
  - For any feasible explicit scenario there exists a corresponding best probability vector. Multiple scenarios may be easily combined
  - Finding the best probability vector is a tractable optimization problem.

Absolute Risk Scenario Output

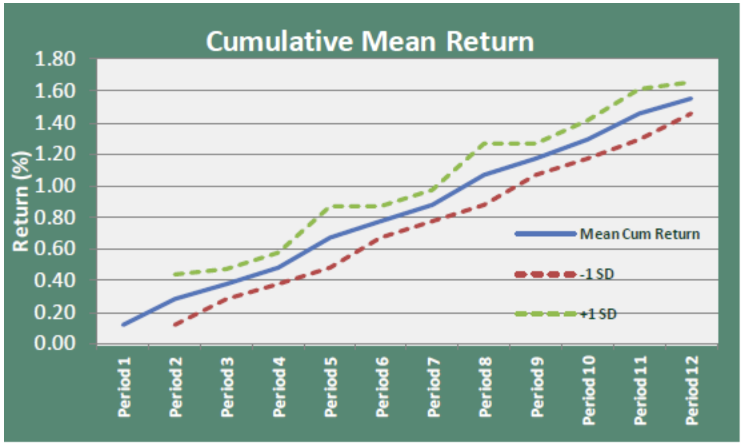


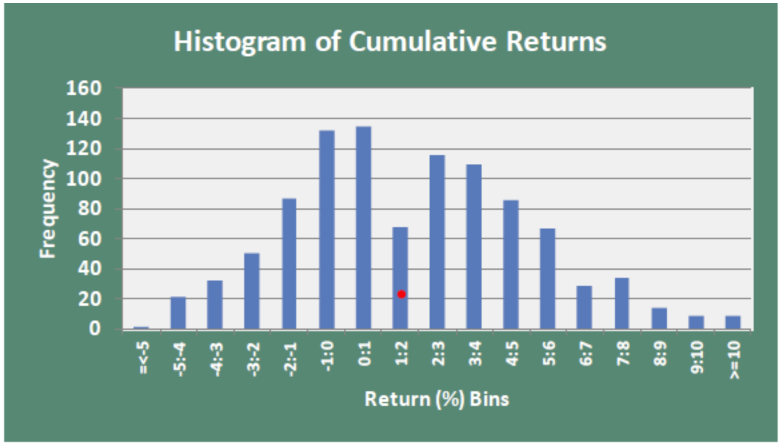


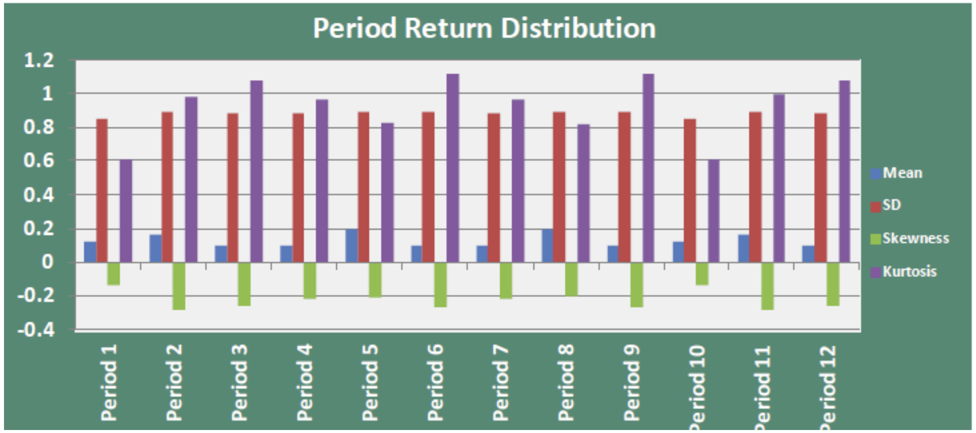


Metric	Mean Value
Cum Ret. (%)	-60.41
Std. Dev (%)	6.24
Skewness	0.45
Kurtosis	2.89
Excess Kurtosis	-0.12
Periods with positive active returns	519
Periods with negative active returns	11481

Historical Benchmark Relative Output

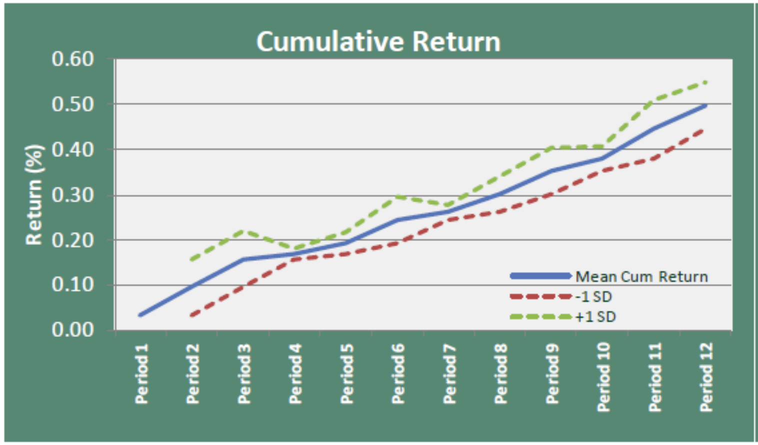


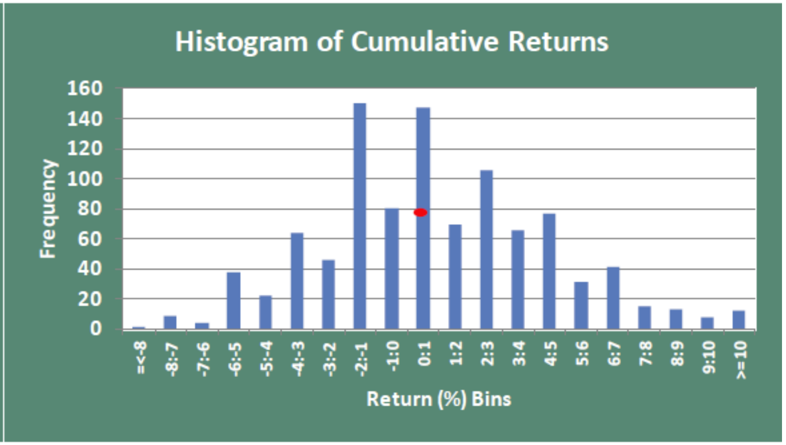


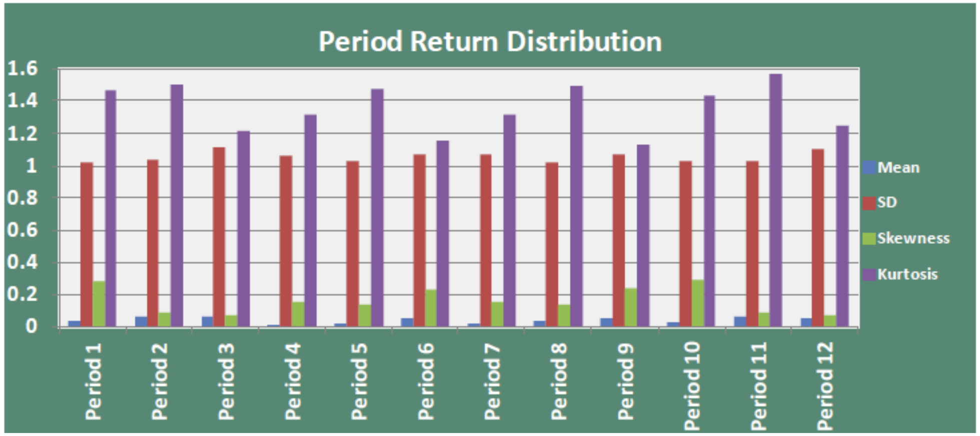


Metric	Mean Value
Cum Ret. (%)	1.55
Std. Dev (%)	3.13
Skewness	0.34
Kurtosis	-0.08
Periods with positive active returns	472
Periods with negative active returns	528

Scenario Benchmark Relative Output







Metric	Mean Value
Cum Ret. (%)	0.50
Std. Dev (%)	3.56
Skewness	0.30
Kurtosis	0.09
Periods with positive active returns	364
Periods with negative active returns	636

Higher Moments and Portfolio Formation

- At each moment in time investors have to form their portfolio. When must they be concerned if their portfolio contains securities where higher moments are present in the return distribution?
  - Markowitz (1952) says he would have preferred to use semi-variance as the risk measure instead of variance. Too computationally difficult for the 1950s. He later argued that just mean and variance were sufficient for practical cases.
  - Samuelson (1970) argues that investors should define their choices to maximize utility over all moments.
  - Hlawitscka and Stern (1995) show the simulated performance of mean-variance portfolios is nearly indistinguishable from the utility maximizing portfolio.

Should Investors Care About Higher Moments?

- Wilcox (2000) shows that the importance of higher moments is an increasing function of investor gearing.

- Cremers, Kritzman and Paige (2003) Use extensive simulations to measure the loss of utility associated with ignoring higher moments in portfolio construction
  - They find that the loss of utility is negligible except for the special cases of concentrated portfolios or "kinked" utility functions (i.e. when there is risk of non-survival though bankruptcy).
  - For an agent manager being fired may be the equivalent of non-survival. Ethical issue: Whose risk are we worried about?

- Satchell (2004) Describes the diversification of skew and kurtosis
  - Illustrates that plausible utility functions will favor positive skew and dislike kurtosis.

A Key Advancement

- "Repricing" based risk engines do some form of simulation (historical, Monte Carlo) to estimate the distribution of possible dollar values for a portfolio of K assets over N observations.
  - Computational burden is N*K, which is material for large N and K

- Even for highly non-linear instruments the computational burden is reduced to what is necessary to estimate the skew and kurtosis of the each asset.
  - If we are interested in a particular confidence interval (e.g. 95% Value at Risk) we reduce the computational burden to 2 * K

Extension to Factor Models

- For any portfolio viewed in a factor model, there exists a numerically equivalent full covariance matrix. See diBartolomeo (1999)
- In a factor model context, any asset portfolio can be replicated as combination of factor mimicking portfolios plus some amount of idiosyncratic risk.
- This feature allows us to easily capture a second source of higher moments: the skew and kurtosis observable in the factor returns.
  - If the returns to a factor (FMPs) have skew and kurtosis, we can now readily include those effects in the closed form computation of portfolio level skew and kurtosis.
- We can also capture this second source of higher moments via bootstrap simulations of factor returns with fixed factor exposures
  - This analysis is *already available* in the **Optimized Scenario** function within the Northfield PRISM application.

Building the Intuition

- The key to understanding the incorporation of higher moments that we cannot say that a particular security has more or less risk.
  - We need to know the sign of the position weight, or the active weight against a benchmark

- Consider a simple "out of the money" call option
  - The distribution of return has a large positive skew with a large positive return (at low probability) and maximum loss truncated at -100% (at higher probability).
  - Compared to a symmetric distribution with the *same standard deviation,* I have less risk if I have a long position, but more risk if I have a short position.

Combining Two Ideas into A Short Cut

- Satchell (2004) Illustrates that plausible utility functions will favor positive skew and dislike kurtosis.
  - It is therefore very natural to consider skew an increment to return and kurtosis an increment to volatility.
- For an asset with positive skew, compared to a symmetric distribution with the *same standard deviation,* I have less risk if I have a long position, but more risk if I have a short position.
  - For positive skew, a long position has a smaller *effective volatility* and a short position has a larger *effective volatility.* For positive excess kurtosis, effective volatility is larger and vice versa.
- For any confidence interval, the magnitudes of the adjustments to both return and volatility are given by Cornish and Fisher (1937)
  - Practitioners have long used this kind of adjustment such as "convexity adjusted yield", a standard bond metric.
  - We can convert the four moment problem to mean-variance

A Bit More "Lateral Thinking" Combinations

- For investors not accustomed to thinking in terms of a specific utility function for their portfolio, even knowing the four moments of their portfolio and related metric (e.g. VaR) may be unintuitive.
- If we know (or can infer) mean-variance risk aversion for the investor we can convert the Cornish Fisher adjustment to expected return into another "utility equivalent" adjustment to volatility. Expected volatility is adjusted three ways (skew, kurtosis, and the incremental impact on return).
- Conversely, we could adjust post CF expected returns for the change in volatility expectations. This difference may be an intuitive way to describe the "multi-period, long term cost of risk".
- Wilcox (2003) provides an approach to inferring optimal risk aversion (or risk tolerance) for an investor.
  - In Northfield terminology, a simple rule of thumb is RAP = 6 times volatility (after adjustment for higher moments)

Intuition for the Factor Conscious

- To the extent we aggregate the impact of higher moments as adjustments to the expected volatility of a security or portfolio, we could also represent these adjustments as different magnitudes of factor exposure.
- For example, it is common for practitioners to think of "market timing alpha" as having a low "beta" (market factor exposure) in down markets and a high "beta" in up markets.
  - This concept goes back to Treynor and Mauzy (1966) where beta was subdivided into sensitivity to market returns, and sensitivity to market returns squared (implicit tail events).
- Volatility adjustments arising from kurtosis represent changes in the factor exposures and idiosyncratic risk. Volatility adjustments arising from skew represent changes in factor exposures conditional on the sign of the position weight.

Dealing With Portfolios

- Even if we know how to adjust expected volatility or expected return at the security level, we have to consider how these adjustments would play out in a portfolio.
- If we have estimates of the four moments of the return distribution for each individual asset, we can *calculate the estimated four moments of the portfolio return distribution in closed form.*
  - Satchell and Hall (2013)
- Alternatively a basic observation allows us to have an immediate intuition.
  - Portfolio variance is a function of squared portfolio weights
  - Skew is a function of cubed portfolio weights
  - Kurtosis is a function of fourth power portfolio weights
- As a portfolio diversifies, skew and kurtosis diversify away faster than variance.

A Simple Measure of Diversification

- Assume a hypothetical world where all portfolios are equal weighted, all security returns are uncorrelated and all securities have the same volatility.
  - In this special case, the variance of a portfolio declines linearly with the number of positions
  - Skew declines with the number of positions to the 3/2 power and kurtosis declines with the number of positions squared.

- For any real portfolio we can find the "effective number of positions" as:

$$(\text{Average of Individual Security Volatility/Portfolio Volatility})^2$$

Conclusions

- We have long held reservations as to the usefulness of "stress tests" and "scenario analysis" for financial institutions where day to day solvency is not the primary goal of risk management.
  - Strategies based on low probability scenarios are sub-optimal for the vast preponderance of circumstances.

- Irrespective of our view, regulatory reforms in many countries are forcing more financial organizations to at least consider these concepts in their risk management process.

- Combining our normal factor risk models, bootstrap resampling and scenario driven conditioning can provide a rich set of information about the potential distribution of future periodic or cumulative return outcomes over a short or long time horizon, in a way that can be more intuitive for fundamental investors. **The process is also very computationally efficient**