

Are You Modeling What You Intended?

Building Capital Market Assumptions for Monte Carlo Simulation

DAVID B. LOEPER, CIMA
CHAIRMAN/CEO

REVISION NOTICE - August 15, 2003

This revised version contains additional asset classes that were not included in the original version. Also, the correlation matrix has been revised by testing with a Cholesky Decomposition Matrix enabling all of these capital market assumptions to be used in a log normal Monte Carlo simulation AND for use in a mean variance optimizer WITHOUT THE NEED TO DEFINE CONSTRAINTS.

The financial services industry has been creating capital market assumptions for various purposes (mostly for mean variance optimization) for many years. The approaches utilized and the premises supporting the resulting assumptions, as well as the results that are generated, vary widely dependent on numerous factors. Are the assumptions based on the premise that one can forecast the future? Are they backward looking using only history as a guide? Are they built as a hybrid using history as a forecasting compass for a future assumption of mean reversion? Are they a hybrid that is modified downward to “be conservative” in the assumptions?

Regardless of the premises supporting the approach, it is absolutely critical that the capital market assumptions one creates for use in whatever system, be it mean variance optimization or Monte Carlo, have the effect of modeling what it is we are intending to model. While that statement is obvious, there is often a wide gap between how assumptions are built, and what effect those inputs have in our models.

In any mathematical model the assumptions are the instruction set telling the calculation engine what we want it to calculate. Math engines are not intelligent and they stupidly calculate (with the beauty of mathematical precision) exactly what you tell them to calculate. The precision of the results often lulls us into a false sense of security that we have done a thorough analysis, while completely forgetting that the math engine analyzed nothing and only did what it was told.

The math engines we use do not have the ability to warn you that your assumptions are erroneous, or even stupid. Whether your assumptions make sense and are modeling what you intended is a judgment call and the math engine exercises no judgment. It is your responsibility to evaluate whether the instructions you gave the engine model what you intended. If you instruct such math engines that money stuffed in a mattress will out-perform stocks and bonds for the next forty years, these engines naively comply and model exactly what you instructed them to model...usually to several decimal points.

When powerful new modeling tools become available we are often excited about the potential information they would enable us to discover. What they enable us to analyze was complex, but with the newfound power of the engine, it is easy. All we need to do is give them the right instructions, and these engines will produce this new information for us. Except, coming up with the right instructions is often far more difficult than actually creating the mathematics to crunch the instructions.

Regardless of the tool and your premises about how you are creating assumptions, it is absolutely critical that you evaluate whether the instructions you give the tool are modeling what you intended to model.

Monte Carlo is one of the greatest advancements in financial services because it enables us to potentially model uncertainty. It will have utility in many areas beyond financial and retirement planning as it is most commonly used today. No doubt, the power of this math engine will be applied in all sorts of new ways, from modeling the uncertainty of life spans, concentration risk, estate taxes and even whether one should exercise their stock options.

While Monte Carlo has all of this potential, for us to realize the benefits of this potential we must be sure the results we are getting from the engine are placed in the context of what we were attempting to model. There is no useful information provided by simulating 30,000 years of investing if the results of those simulations are completely unreal, either too aggressive or too conservative. Such results either put clients at too great of a risk of failing to achieve their financial goals, or if too conservative will cause them to make unnecessary sacrifices to the only life they have. Both of these violate key premises of Wealthcare, our revolutionary advisory process.

This balance of making sure assumptions produce results that are not either too aggressive or too conservative is extraordinarily delicate. The implications of being wrong (or even worse being negligent in whether we validated the reasonableness of the results) are far reaching. If our assumptions are too aggressive and are broadly applied, many investors will fail to meet their financial goals while if they are too conservative we will be needlessly sacrificing the financial lives of many people.

How can one tell whether the results the engine produced (based on our instructions) balance this delicate scale? That is completely dependent on what it is we are intending to model. If you believe, as Financeware does, that one of the greatest uses of Monte Carlo is the potential ability to model the fundamental nature of the capital markets in the context of history and stress test markets that have not yet but might occur, while measuring the confidence or uncertainty for an investor's lifetime package of financial goals, then it is critical that we evaluate whether our inputs produce results that model this.

Unfortunately for most investors, poorly designed inputs are NOT producing this result although it is generally the context of how the analysis that was done for them is described. This is misleading investors. Sometimes it is caused by using the wrong premises (like mixing an assumption of mean reversion or a market forecast for the next ten years into the inputs to Monte Carlo) and sometimes it is an attempt to be conservative to provide additional "safety" (at the price of one's only life).

Most of this is due to putting square pegs in round holes. Building square assumptions for a current asset allocation based on thorough analysis of market trends (a premise of certainty) does NOT fit well into the round hole of the uncertainty in historical context of what may occur. Reducing our return assumptions in the context of a tool that only models a flat return assumption and ignores uncertainty, has a devastating impact when applied in a model that exploits uncertainty.

Methods of Building Capital Market Assumptions

Perhaps the most common method of building capital market assumptions is simply using the average return and risk for the longest period for which we have data. On the surface, this sounds like a reasonable approach. After all, we can't do any better than using all the data we have...or can we?

When examined a little bit more thoroughly, this over-simplified approach can produce results contradicting the intent of our analysis. We are letting data drive our decision and if we don't validate if the results are in alignment with our intent, we can end up with completely misleading results.

Exhibit 1 - Trailing Large Cap Returns Through 2001 (Source: Center for Research of Security Prices, top two deciles of market cap)

Trailing Period	Mean Return	Risk (Standard Deviation)
5 Years	12.70%	20.93%
20 Years	16.02%	17.78%
40 Years	11.72%	16.67%
60 Years	13.42%	15.84%
75 Years	11.92%	20.50%

In *Exhibit 1*, we can see how sensitive the risk and returns are based on trailing data. Are longer periods “better”? Is the 20-year data “better” than the 5-year data? Is the 60-year data better than the 40-year data?

Since we are trying to address building capital market assumptions for use in modeling the context of history in Monte Carlo simulations, maybe we should compare history to what happens when we use a trailing period as our instructions for simulation.

Exhibit 2- Comparison of all history to simulation results using trailing risk & return – (30-year periods using monthly data –30-year geometric mean returns - source: Center for Research of Security Prices, top two deciles of market cap)

	Result for simulation based on trailing return & risk					
	Historical	5 Years	20 Years	40 Years	60 Years	75 Years
0% (Best)	14.29%	21.89%	24.09%	19.32%	20.74%	20.94%
100% (Worst)	7.17%	-3.04%	2.69%	-0.77%	1.60%	-3.49%

In *Exhibit 2*, the first thing we notice is that the simulated range of returns is much wider than actual historical results. This is what we would expect since we are running 1,000 30-year trials for a total of 30,000 years of investing.

The basic idea of what we can do with Monte Carlo is simulate things that have yet to occur, but have some chance of occurring. In 1980, after several years of a flat market, we never would have expected the next twenty years would be better than any market seen up to that point. Likewise, at the peak of the tech bubble, it was difficult to see the potential for a three-year bear market, the likes of which we haven't seen in years. Although these sorts of markets were not “planned for,” Monte Carlo “stress-testing” told us there was a probability that they could occur. Monte Carlo could not forecast that these environments would occur, only that there was a chance that they might occur.

The best and worst results are more extreme, but one should understand there is only a 1 in 1000 chance that the best or worst simulated results would be realized. Our limited historical data falls in-between the extremes of the simulated ranges, constrained by the reality that we only have 76 years of data and what might happen over 30,000 years would obviously have greater extremes. Think of it as if you flipped a coin 76 times and then flipped it 30,000 times. The odds of getting 10 heads in a row by flipping the coin 76 times is remote, while you would almost certainly get 10 heads in a row at some point if you flipped the coin 30,000 times.

In looking at the extremes in *Exhibit 2* with the knowledge the extremes will be wider with Monte Carlo, it is difficult to ascertain which trailing period might more realistically model the potential future based in historical context. Which period models the extremes correctly? Is it the worst case, 1 in 1000 chance, lowest return of -3.0% a year as shown in the trailing 5-year data, -3.5% as shown in the trailing 75-year data, or +2.7% as in the trailing 20-year data? Or, might it be none of them?

How much more extreme should the range be? Where do the results fall within the extremes since the best and worst are only a 1 in 1000 chance, which no one should plan their life around? Here is where we start to see the impact of why we cannot blindly use “the last however long time period.”

Exhibit 3- Comparison of all history to simulation results using trailing risk & return – (30-year periods using monthly data –30-year geometric mean returns - source: Center for Research of Security Prices, top two deciles of market cap)

	Historical	Result for simulation based on trailing return & risk				
		5 Years	20 Years	40 Years	60 Years	75 Years
0% (Best)	14.29%	21.89%	24.09%	19.32%	20.74%	20.94%
5th %tile	12.94%	17.28%	20.19%	15.67%	17.28%	16.43%
25th %tile	11.69%	13.04%	16.59%	12.29%	14.07%	12.28%
50% (Median)	10.18%	10.63%	14.53%	10.36%	12.23%	9.92%
75th %tile	9.58%	8.18%	12.43%	8.38%	10.35%	7.52%
95th %tile	8.01%	4.85%	9.55%	5.68%	7.77%	4.25%
100% (Worst)	7.17%	-3.04%	2.69%	-0.77%	1.60%	-3.49%
Average	11.99%	12.43%	15.81%	11.52%	13.27%	11.66%

In *Exhibit 3* we examine where historical results fall relative to our simulated results and we can start to assess whether our assumptions were modeling the nature of history as we intended. Now, we know that we are modeling broader extremes, so where a particular historical result will fall should not match the simulated probability precisely. But, was it our intent to simulate over 50% of all market environments will do better than the best of actual historical periods? If that were the intent, then the trailing 20-years would be appropriate for our assumptions. Compare the “best” historical result to the results of the trailing 20-years. A simulation run based on the trailing 20-year data as inputs, is instructing the simulation engine to model a future with a probability that over half the time the market will do better than it ever had in the past. A little less extreme is the results from a simulation based on the trailing 60-years which projects that nearly 25% of all future markets will be better than the best the market has ever done. Using the trailing 20-years as inputs causes us to simulate approximately 70% of our trials being better than the top 5% of historical results. Was this what we intended to model?

In *Exhibit 4* we examine the average and lower range of historical results and where they fall relative to our simulations.

Exhibit 4- Comparison of all history to simulation results using trailing risk & return – (30-year periods using monthly data –30-year geometric mean returns and mean of all returns - source: Center for Research of Security Prices, top two deciles of market cap)

	Historical	Result for simulation based on trailing return & risk				
		5 Years	20 Years	40 Years	60 Years	75 Years
0% (Best)	14.29%	21.89%	24.09%	19.32%	20.74%	20.94%
5th %tile	12.94%	17.28%	20.19%	15.67%	17.28%	16.43%
25th %tile	11.69%	13.04%	16.59%	12.29%	14.07%	12.28%
50% (Median)	10.18%	10.63%	14.53%	10.36%	12.23%	9.92%
75th %tile	9.58%	8.18%	12.43%	8.38%	10.35%	7.52%
95th %tile	8.01%	4.85%	9.55%	5.68%	7.77%	4.25%
100% (Worst)	7.17%	-3.04%	2.69%	-0.77%	1.60%	-3.49%
Average	11.99%	12.43%	15.81%	11.52%	13.27%	11.66%

In *Exhibit 4* we examine the worst historical return of 7.17% and see that the simulation results based on trailing 75-year data, model a future where the likelihood of markets doing worse than they ever have is nearly 1 in 4 (the 75 th %tile). Also notice how much the simulated averages vary relative to history. With 30,000 years of data we might expect some small difference in simulations versus actual history. But, by blindly using trailing data, we can see that it may cause us to be off by more than 300 basis points!

The fact is that ANY “trailing period” can be skewed simply by the population of returns making up the trailing period. If there is an extra bull market as in the last 60-years of data, our numbers are skewed upwards and we end up modeling overly optimistic results, at least in the context of history. If this is our intent, that is fine, but if we are describing our analysis as evaluating the probability of what might happen based on the historical nature of the markets, we are doing nothing other than misleading the client.

In some cases the average return may be reasonable for a trailing period (as in the trailing 75-year data) but the risk may be higher or lower, which has similar unintended consequences exposed by evaluating the geometric mean which is a function of both mean return and standard deviation.

While we may feel as though long periods (like 30-years) should be “close enough and average out,” in reality we can easily see, simply by examining the data, that there is a wide and volatile range of results that can quickly change even with long periods.

Exhibit 5 - Rolling 30-year mean and geometric mean for large cap stocks - (source: Center for Research of Security Prices, top two deciles of market cap)

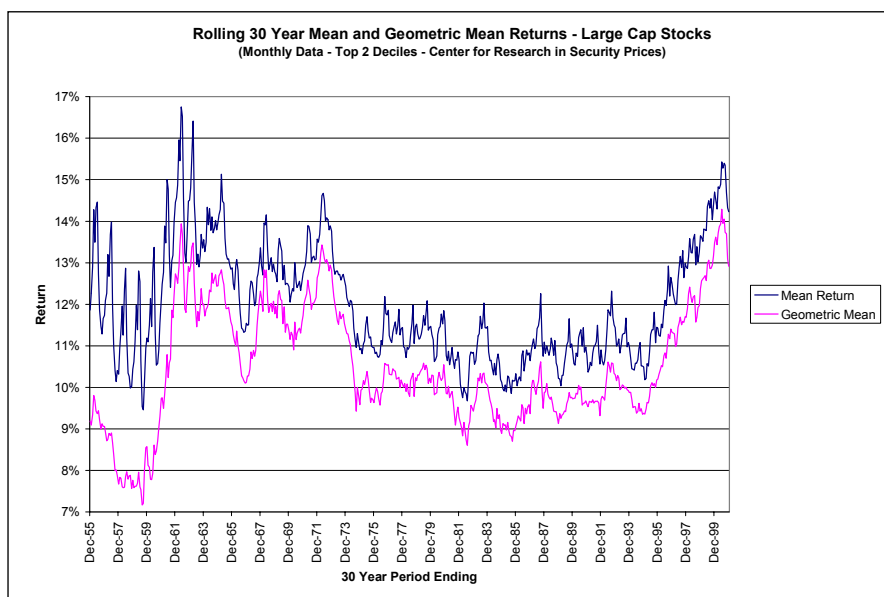


Exhibit 5 shows us that our assumption that “things average out” over the long term may be true relative to the extreme volatility of one-year returns, but still can be far more volatile than we presume, which is one more reason we should use Monte Carlo.

Changes In Capital Market Assumptions Should Be Rare

In the case of Monte Carlo simulation, if we did an exceptional job of creating our capital market assumptions, they should rarely if ever change unless there is a fundamental reason the markets have changed. Unfortunately, many have become so accustomed to using trailing returns (which invariably change each year with the addition of a single data point) that they actually expect capital market assumptions to change each year. Well-designed assumptions should model the nature of historical markets. Any method of building capital market assumptions that would cause them to materially change based on one additional year of data is critically flawed. A single data point is the random occurrence we model in Monte Carlo. If our method of building assumptions permits one data point to have this effect, we are randomly building our assumptions.

We see how volatile the 30-year mean returns are in *Exhibit 5*. Is it rational to assume that simply because a year has passed, that simply because we had a really good year, our assumption for the nature of markets going forward should be dramatically higher? Of course not. Yet if we change our assumption based on each new year of data, we are doing exactly that. If a return happened that falls outside of the range of what we expect, that might be time to reconsider. It is pretty hard to see any of the returns the market has produced to date as being outside of what one would statistically estimate.

It may be helpful to explain a few other points here. Some of you may be wondering why we used the top two deciles of the NYSE as provided by the Center for Research in Security Prices as opposed to the more widely used Ibbotson Large Cap/S&P500 data. There are a few basic reasons for this.

It is little known, but is disclosed, that the Ibbotson Large Cap data does not represent the S&P500 prior to 1957 since it did not exist. All of the Ibbotson Large Cap data, prior to 1957, represents the S&P90. Also, while the index is generally accepted as a “passive” representation of large cap stocks, in addition to the change from 90 to 500 stocks in 1957, the index had a fixed composition of 400 industrial companies, 20 transportation, 40 financial and 40 utility companies, which was changed in the mid 1980s. It is managed by committee and as much as 10% or more of the stocks can change in any one year. Sometimes this is due to mergers and acquisitions (a new stock has to be added to keep the count at 500 stocks) and sometimes this is due to performance.

Finally, while the low turnover, fundamental index construction and basic committee decisions may be a statistical “nit,” we feel the data is not completely objective for these reasons and may be why the numbers are consistently higher both in terms of return and risk. Whether evaluated from a trailing geometric mean, rolling 10-year periods or simple average, the returns and risk are higher than a completely unbiased index. Even the maker of this index acknowledges the “list effect” as positively impacting the performance of the index. (See the paper by David M. Blitzler, PhD entitled “The S&P500” presented on March 26 th , 2001).

For these reasons we have decided that a passive, unbiased and objective approach may be more representative of the fundamental nature of the markets as opposed to markets influenced by human effect. We are changing our official capital market assumptions from the Ibbotson data originally used as our basis so that we may correct for this influence. Our reputation should permit us to expand beyond what was popular when we first started our company. Although the statistical difference is small, it none-the-less appears to exist.

What Asset Classes Should We Simulate?

On this topic we have written many other papers that cover in greater detail what makes an “asset class” and what may be an attempt at false precision that ends up increasing estimation error rather than reducing it. (Visit <http://www.financeware.com/advisors/whitepapers.asp> for more information.) When it comes to segmenting portfolios by market capitalization, one should realize that stocks are stocks and the market weights them on their relative value. Any specific allocation to small, mid or micro cap stocks that is not based on the relative market weight should be understood as a bet against the nature of the markets.

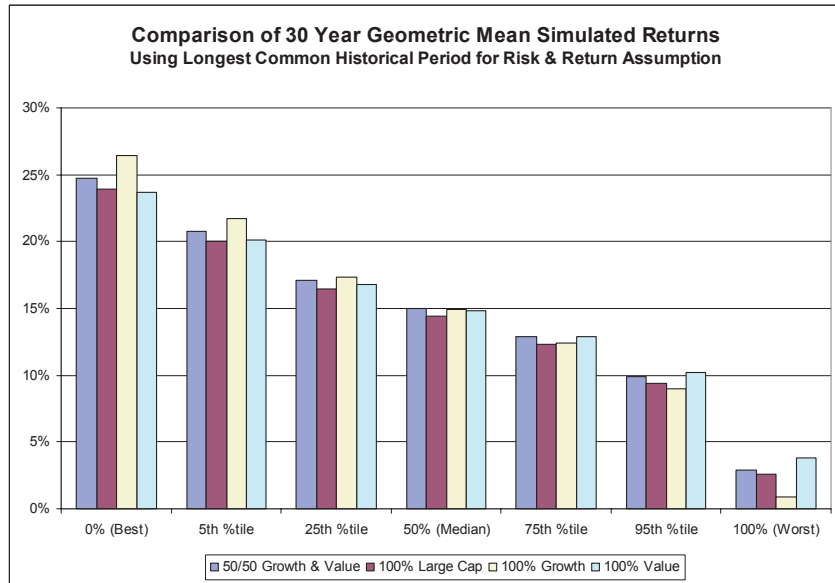
In the case of foreign or domestic stocks, there is no fundamental reason why stocks with headquarters in Germany or Japan should be expected to perform better or worse than any stock with headquarters in the USA. Not that they don't... they do just as each stock in the USA performs differently than another. The element that makes foreign stocks different from domestic stocks is the introduction of additional

risk, due to currency and potentially political risk, with some amount of compensation for this risk in the form of lower correlation and potentially more localized economic trends. From the perspective of applying Monte Carlo, we believe it would be unwise to make specific assumptions about returns, and instead focus on the fundamental impact of additional risk and diversification. Together, we suspect the impact is not statistically significant relative to the other assumptions being made and if one attempts to model foreign stocks, they should not produce a material difference as a bet against the nature of stocks in general. The data to prove this does not yet exist, but the rationale is sound as it is based on not betting against the nature of the market in general.

In the case of “style” analysis, we really have a case of the tail wagging the dog. A little common sense here goes a long way toward making better decisions. For example, what makes up the universe of large cap stocks? Most style analysis starts with a passive index of a diversified market as a whole, and splits it in two (normally based on price-to-book ratio). Did I create anything new by doing this? I just took the whole, and split it in half. When combined, what do they equal? The whole. If I have the whole, what do I have? I have the sum the two halves, or...the whole.

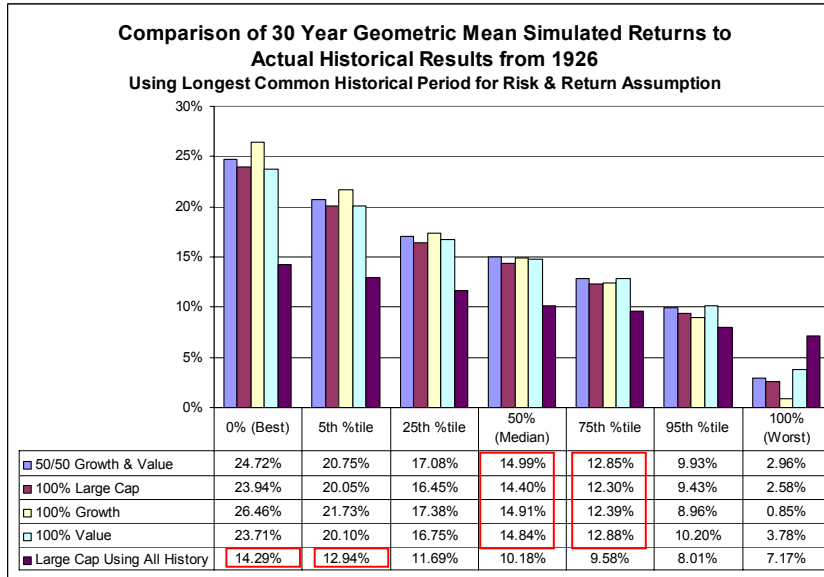
Style segmentation is a fairly recent fad generated by excitement about mean variance optimization. Of course, the two halves will not independently perform the same as the whole and there appears to be some fundamental differences, but only because we created them. By intentionally splitting the universe, we find that value will yield more, growth will yield less, etc. Attempting to model the nature of growth and value stocks is fraught with problems about not having enough data, similar to speculation about returns in foreign stocks. *Exhibit 6* shows how small the impact is using all available data.

Exhibit 6 – Comparison of simulation results for large cap and large cap growth & value (23-years ending 2001).



Maybe you perceive some material differences in this data, but to me these results are about the same. What is worse though is in our attempt to fine-tune our accuracy by splitting the whole into halves, we are limiting the data we have available. Perhaps before we leap to the decision that the small differences shown in *Exhibit 6* are enhancing our precision, we should evaluate whether limiting our view of data to the last 23 years (remember the trailing return problems we exposed?) is fundamentally causing us to model something we had no intention of modeling.

Exhibit 7 – Comparison of simulation results for large cap and large cap growth & value (23-years ending 2001) to all 30-year periods based on monthly data to 1926.



Are we modeling what we intended? We know the “best” simulated results should be better than the best of history and the “worst” should be worse, based on the number of years we are simulating. In *Exhibit 7* though, notice how our simulated results had nearly 75% of our trials doing better than the top 5% of all historical results. Is this additional precision? Did we really want half of the trials to be better than THE BEST history had ever produced?

While we do not object to modeling growth and value per se, and we create capital market assumptions for them (which happen to be similar to their large cap index of which they each make up half), we must make sure that when we make such attempts we don’t blindly let the data cause us to model something we had no intention of modeling.

Conservatism, Mean Reversion and Forecasted Capital Market Assumptions

So far we have analyzed how we need to be cautious about using trailing returns, changing our assumptions each year because of one year of data, and focusing on a tree (growth or value) while ignoring the forest. Do these if you wish, but in so doing make sure you validate whether the results are modeling what you intended.

The recent bear market has made it popular for advisors to complain about “return assumptions being too high,” It is rather ironic that many of those currently suggesting our assumptions are too high, just four or five years ago complained the assumptions were too low. Well-designed capital market assumptions should not change if they are designed to model the nature of the capital markets. While it may be easier to sell lower return assumptions when the markets are declining, and...it appeared as if we were overly conservative in our assumptions when the markets were roaring, the return assumptions should fall near the middle of the range of results, and should stay there. Our simulation will take care of modeling both higher and lower results, along with their probability, if our assumptions model the nature of capital markets.

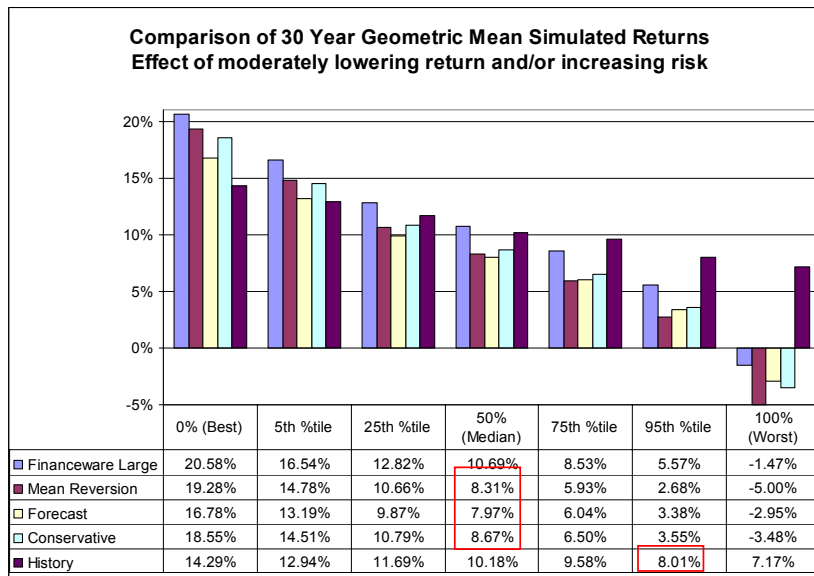
Monte Carlo is about modeling uncertainty, and just as oil and water do not mix, taking certainty (mean reversion or forecasts) and mixing that into a model designed to measure uncertainty can have severe, unintended consequences. While there may be some brilliant forecasters out there, their forecasts serve little purpose when what we are trying to understand is the odds of them being wrong. Mean reversion, by the way, is just another form of forecasting. Monte Carlo will model some periods of mean reversion, albeit randomly as they really occur. Trying to influence the results of Monte Carlo by inputting a mean reversion assumption is in essence double counting the impact and the likelihood of the effect.

Finally, the same holds true for something we normally view as a positive...conservatism. If we do all this analysis, and we test the results to make sure they fall properly in the context of history as Finaceware does, wouldn't it be a good idea if we just lowered the returns somewhat to "be safe?"

If you think about it, this is really just the opposite side of the same coin as we discovered by letting data drive our decisions causing us to simulate 75% of the trials as being better than the top 5% of history.

If our intuition is right on this, being conservative should produce results that are far more pessimistic than we intend. The problem with this is opposite of the overstating of results, which causes clients to have false expectations about how confident they could be, thereby encouraging them to spend more than they should. Conservatism in the inputs instead has the consequence of needlessly sacrificing the only life our client has. To us, both should be considered criminal.

Exhibit 8 – Comparison of simulation results for large cap based on Finaceware recommended assumptions, actual history and modeling lower returns, and/or higher and lower risk.



As in our other analysis, the effect of testing 30,000 years of investing makes the simulated best and worst case more extreme than for history. Likewise, as one would anticipate, our pessimism in trying to be "smarter" or "conservative" may have an unintended impact on what we were trying to model.

In *Exhibit 8*, look at the 95th percentile historical return of 8.01%. If we lower our return and increase our risk by 2% as modeled in the "mean reversion," we are simulating that nearly HALF of all future

markets will produce results WORSE than the bottom 5% of historical results. In the “forecast” we assumed some clairvoyant analyst “knew” that returns would be 3% lower than historical norms and risk would be lower by 2% with similar, but even more pessimistic results. The “conservative” used the Financeware assumptions for risk and return, but lowered the return by 2% to “be safe.” It is unlikely that the typical use of Monte Carlo would be to model such pessimism. The result of ignoring this (or not even bothering to check) is that clients of advisors that make these sorts of mistakes will be needlessly sacrificing the only life they have.

The Financeware Assumptions

Financeware has been analyzing the capital markets, and how Monte Carlo modeling can be used, for several years. Some consider us industry leaders in this area. Our common sense approach to building assumptions is based on the notion that while we know we cannot be perfect, we should make sure we do not unintentionally skew results one way or another.

Most of our assumptions are built by averaging rolling historical periods, which has the tendency to compensate somewhat for extremes. Generally, we look at the 768 ten-year periods we have data for, and average them to compensate for skewness and frequency. For certain data types we have less data, but a similar methodology is still used because we feel we still have sufficient data. For asset classes that we feel we do not have enough data to use or that would cause the trailing period effect covered earlier, we rely on the nature of the macro asset class where we have good data and then slightly modify those results (i.e. add a risk premium to foreign stocks, but assume correlation of less than 1.0 and use stock returns for the return assumption).

Correlations are generally based on extremely long term monthly data for most of the classes, but recent data is used (only for correlations and to estimate a relative risk factor), for sub-classes that have only short-term data.

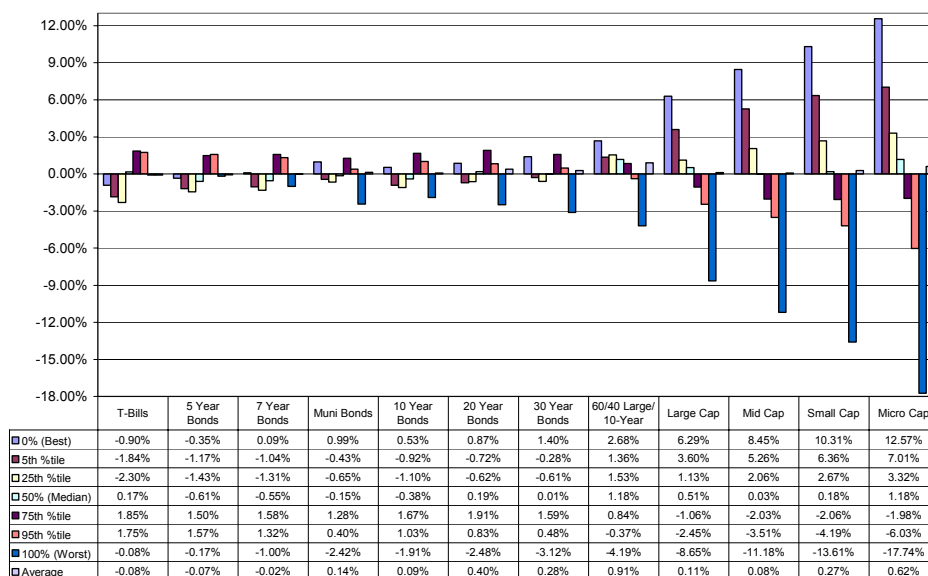
All of the results are validated to make sure we are modeling what we intended. That is, generally the extremes are wider than actual history, but the averages and results within the distribution set do not cross our intent of what we are trying to model (i.e. 5 th percentile results do not fall in the 50 th or 75 th percentile).

It is important to note that without significant manual adjustment, the extremes simulated for short-term (5 years or less) fixed income investments will not produce “best” results greater than the best of history even though the worst simulated results are still less than history. These differences are extremely small and are likely due to our assumption that returns are log-normally distributed, which is generally accepted as a reasonable approximation of the nature of equity returns. However, for short-term fixed income investments in particular, returns are serially correlated and are not log-normally distributed. This effect does impact the entire fixed income distribution and is likely due to the kurtosis in the nature of those returns (tails actually increase in frequency relative to returns between the middle and end of the distribution set). Users should be very careful though on how they describe the simulation results to their clients if most (>75%) of the investor’s assets are in short-term fixed income and should probably target a higher than normal confidence level (>90%) if this is the case.

In *Exhibit 9*, we see the comparison of the new Financeware capital market assumptions relative to history and measured by the difference in geometric mean returns. It is important to use geometric mean returns for testing the results because geometric mean considers both the risk and the average return and can therefore potentially expose an error caused by either. One can actually calculate the geometric mean (often referred to as compound return) if they have the mean return and standard deviation. The

difference between mean and geometric mean is directly related to the standard deviation. We believe this shows that our assumptions do model what we intended...the general nature of the markets and the probabilities of various returns relative to history.

Exhibit 9 – Difference in geometric mean and average return for simulation results using the Financeware capital market assumptions compared to actual history for 30-year periods.



As we would expect through simulation and “stress testing,” the worst result is consistently less than history and the extent of this is dependent on how volatile the asset class is. As we move up the risk curve, the extremes relative to history get more extreme. We can also see the impact of serial correlation in the “best” results for short term fixed income and how relatively minor the difference is.

The amount of data in this chart, and the “noise” from the nature of short-term fixed income investments shades some of the beauty in simulated lognormal, well-built capital market assumptions. *Exhibit 10* “zooms in” and shows the same information for only asset classes dominated by equities.

Exhibit 10 – Equity Classes - Difference in geometric mean and average return for simulation results using the Finaceware capital market assumptions compared to actual history for 30-year periods.

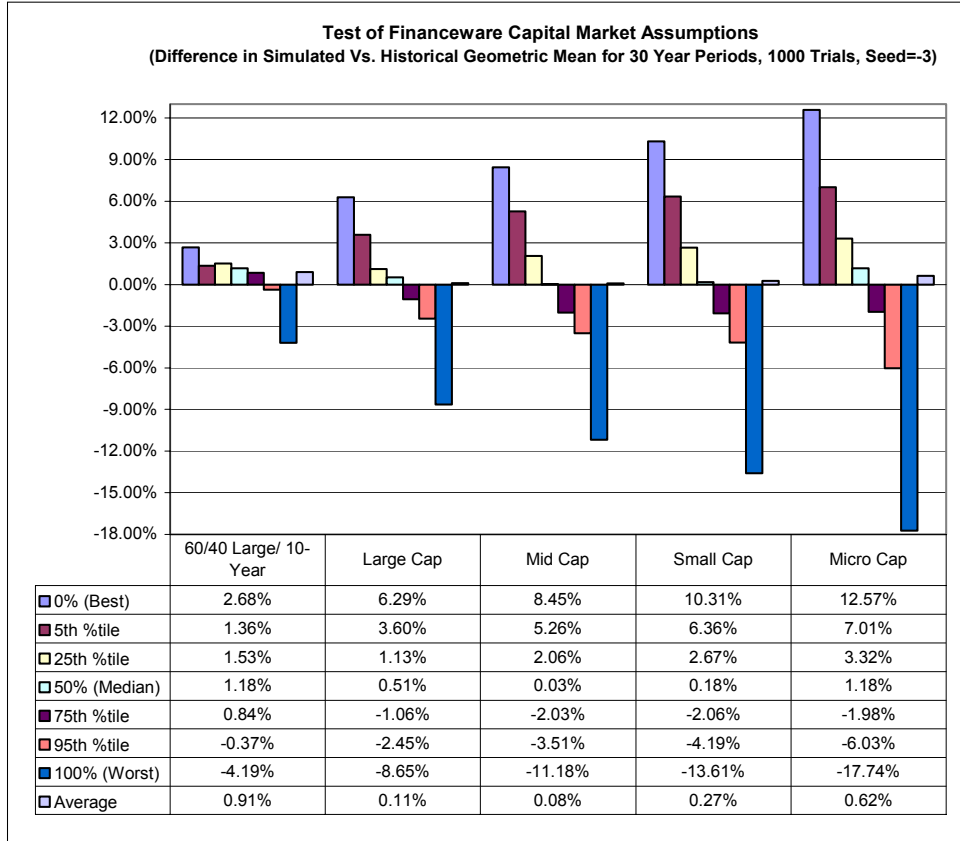


Exhibit 10 clearly demonstrates results that at the tails are more extreme, increasing as asset class volatility increases, with differences relative to history being minor as we move toward the center of the distribution. To us, this represents what we are trying to model. These results could be conceptually described as “stress testing the historical nature of markets” to assess the relative likelihood of markets we have already seen, repeating, and evaluating the likelihood of new, more extreme markets, emerging.

Exhibit 11 – Detailed analysis of simulated geometric mean and average return for 1000 trials (seed -3) versus actual history.

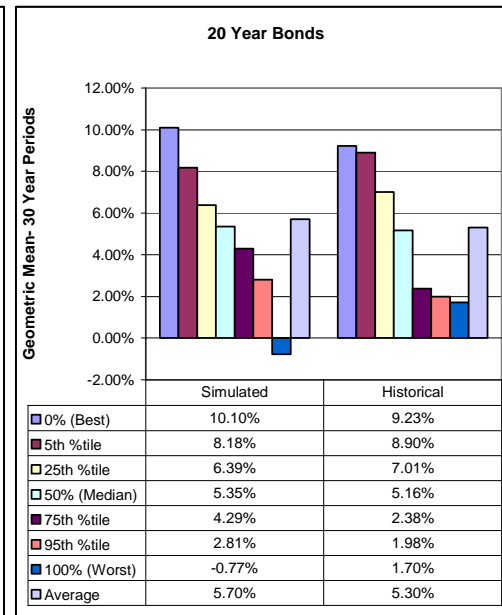
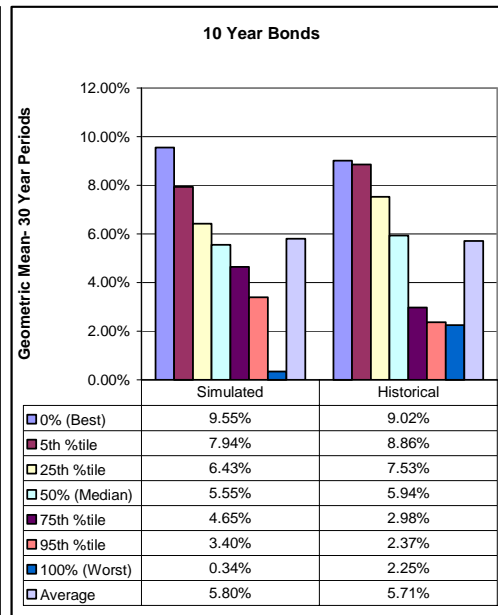
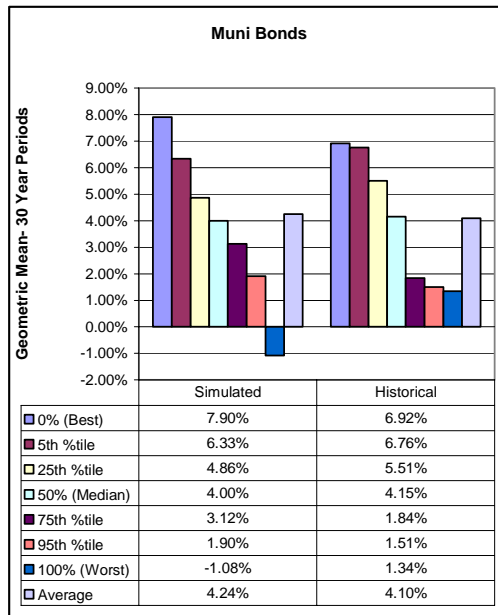
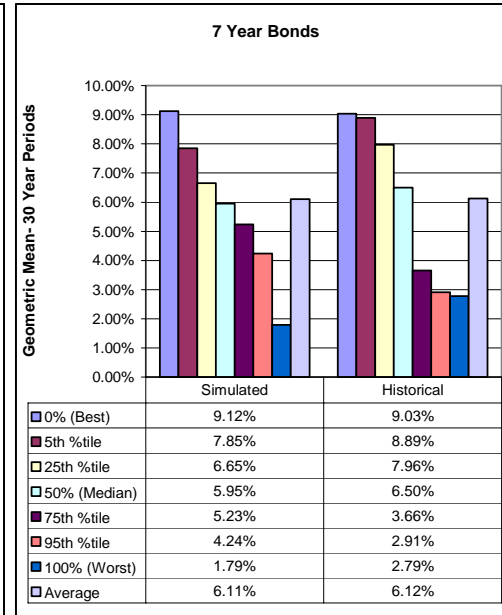
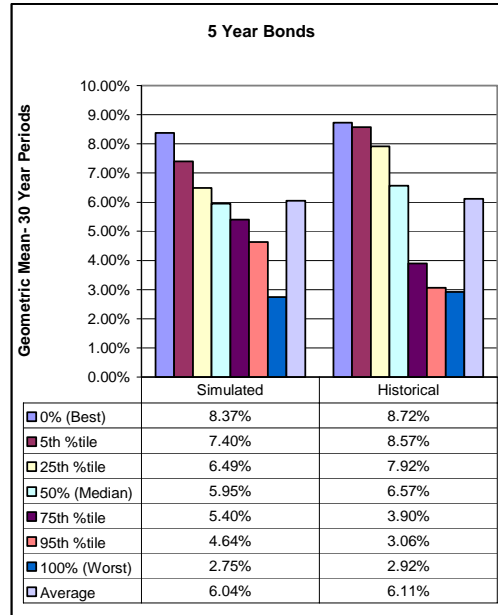
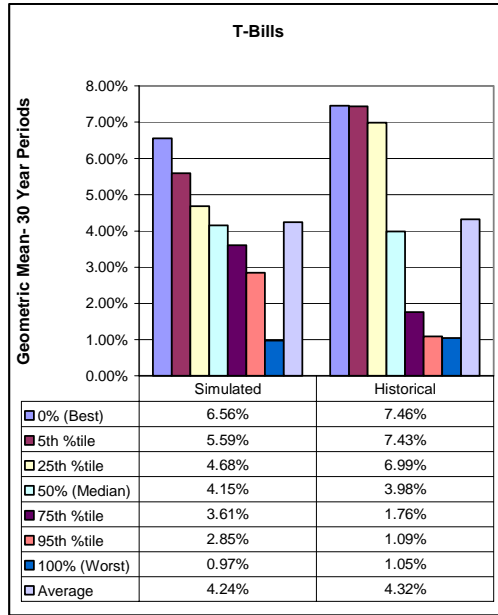


Exhibit 12 – Detailed analysis of simulated geometric mean and average return for 1000 trials (seed -3) versus actual history.

