Beyond Monte Carlo Analysis: An Algorithmic Replacement for a Misunderstood Practice

Originally published in Journal of Financial Planning

Shawn Brayman, BSc, MES
Chief Executive Officer
PlanPlus Inc.

December 2007

Beyond Monte Carlo Analysis: An Algorithmic Replacement for a Misunderstood Practice

by Shawn Brayman, BSc, MES

Shawn Brayman, BSc, MES, is the president of PlanPlus Inc, a financial planning software company from Lindsay, Ontario, Canada. He can be reached at shawn@planplus.com.

Executive Summary

- The application of Monte Carlo simulation (MCS) in financial planning is a largely misunderstood technology that does not deliver the results as advertised. This paper identifies those misunderstandings and outlines a simpler, more accurate algorithmic approach.
- The paper first demonstrates that MCS returns the same results as an algorithmic test when consistent assumptions are applied. When MCS is used to return an overall probability of success or failure of a plan, it has not tested the impact of higher or lower returns, nor the impact of sequence risk such as the order of good or bad years of return and the timing of the client's goals.
- When a consistent geometric mean assumption is used, MCS results are the same regardless of the variability of the portfolio. When the present value of client withdrawal strategies is consistent, MCS shows no variation, regardless of the nature and timing of the withdrawals. Randomizing life expectancy does not effectively test for longevity risk.
- An algorithmic approach (reliability forecast) allows for the calculation of the same partial probabilities as generated with MCS, but provides other benefits.
- A dozen iterations using a reliability forecast calculated the same probability distribution as MCS does using 10,000 or more simulations. This addresses concerns of error with MCS when used with a smaller number of simulations.
- Reliability forecast simplifies the ability to generate a matrix illustrating multiple success factors—not just a single success factor for a financial plan. It also separates the planning model from the probability distribution for randomized variables, easing the development effort.

Despite strong and persuasive academic evidence that Monte Carlo simulations (MCS) should be used as an option of last resort in forecasting financial results, financial advisors, software manufacturers, academics, and the media continue to promote this technology as the Holy Grail of financial planning.

Although MCS is an important statistical methodology that will continue to be used in research and other applications, this paper shows that its application in financial planning remains a largely misunderstood technology that does not deliver the results advertised by the industry. Advisors and their clients may mistakenly believe MCS has effectively tested variability of assumptions in the financial plan when this may not be the case.

This paper illuminates these misunderstandings with MCS and outlines a simpler, more accurate algorithmic approach that in many cases can replace the Monte Carlo randomization technique.

The purpose of the first section of this paper is to identify and dispel some popular myths surrounding Monte Carlo simulation, specifically that advisors believe

1. MCS "tests" the impact of higher or lower return assumptions, rather than just the "mean return" calculation used in an algorithmic solution.
2. MCS tests the impact of timing or sequence risk, specifically for when clients withdraw money from their portfolios, as well as the timing of poor returns and how those variables can affect the success of the investment plan.
3. Randomizing mortality tests for longevity risk.

Although used less frequently, some tools also apply MCS to inflation assumptions used in a financial plan. There is a short discussion on this use of MCS later in the sidebar "Impact of Monte Carlo on Inflation Assumptions."

Sample Case

The three misconceptions are addressed using a sample case: a retired couple each age 60, each with $400,000 in tax-sheltered retirement accounts ($800,000 total), and an assumed 15 percent average tax rate. They want to maintain an after-tax retirement income of $60,000 a year to age 85. After government benefits, it is estimated they require $52,160 a year from their retirement funds. They have a portfolio mix of 15 percent cash, 40 percent fixed income, 25 percent Canadian equity, and 20 percent U.S. equity.

The approach taken here was to solve when investment funds would be depleted using an algorithmic solution and then test what different results are generated using Monte Carlo simulations. The test was a simple Excel spreadsheet calculating the return on investments, receipt of revenue, and withdrawal of funds as required to fund lifestyle assuming an average tax rate. The key is not the specifics of this particular calculation, but rather that we will apply a simple geometric mean assumption and then compare it with the results of randomized returns or mortality generated by MCS and applied to the same calculation.

Historical data (Canadian CPI, Canadian 90-day T-bills, Scotia-McLeod Universe Bonds, Canadian equity represented by the TSX, and U.S. equity represented by the S&P 500) from 1950 to the end of 2005 were used to calculate both arithmetic and geometric returns, and the standard deviation for the portfolio, for the past 25, 35, and 56 years. For the "financial plan," the 56-year data assumptions were used to calculate the geometric mean of 8.96 percent and an inflation rate of 3.94 percent. The portfolio had a standard deviation of 8.07 percent. For this exercise, it does not matter what values we use as long as we test the same assumptions in both the algorithmic model and the Monte Carlo analysis.

Algorithmic Solution Versus Monte Carlo Simulation

Using the geometric mean of 8.96 percent, the drawdown scenario worked out almost exactly as planned, with the client's capital being depleted in 25 years. Each year, the client withdraws $52,159 from the retirement accounts, indexed at the rate of inflation.

In an algorithmic solution, using the geometric mean as the rate of return represents the point at which there is a 50 percent chance the return could be below this amount (failure) and a 50 percent chance it could be above this amount (success).

To perform a Monte Carlo simulation, the advisor must enter capital market assumptions (CMA), which are basically the advisor's rate of return and standard deviation assumptions. The MCS algorithm then generates as many series of random returns as requested. Overall, the average of the standard deviations and average of the returns of the simulations will fall within close proximity to the assumptions, given a large enough sample size.

If the rate of return assumption of 8.96 percent and the standard deviation of 8.07 percent from the portfolio are used as the CMA, the returns from the MCS will be distributed around this value. As a result, the arithmetic mean of the series will be 8.96 percent; however, the geometric mean of each series will be lower. How much lower is a function of the standard deviation. Because the MCS is using a lower rate of return, it will result in more failures than the 50 percent expected. Advisors will likely assume this is the Monte Carlo simulation illustrating the timing or sequence risk when, in fact, it is highlighting a flaw in the assumptions.

Even if financial advisors understand this issue, it is difficult for them to address it, as the geometric mean cannot be calculated from the arithmetic mean and standard deviation (although some estimation techniques exist). Without reverting to historical data, it may be impossible for advisors to generate matching assumptions.

This is not a criticism of how Monte Carlo simulation is applied by knowledgeable academics or mathematicians, but an observation of its pitfalls in an industry where many advisors struggle with the difference between arithmetic and geometric means.

Conclusion: Advisors may believe Monte Carlo is "testing their plan" when in fact it is calculating the outcome based on different assumptions.

Myth 1: MCS Tests Higher and Lower Returns

Advisors believe MCS "tests" the impact of higher or lower return assumptions and portfolio variability, rather than just the "mean return" calculation used in an algorithmic solution.

To validate the belief that MCS tests for return variability, this paper generated an MCS of representative returns using the arithmetic mean of 9.26 percent and 8.07 percent standard deviation for the portfolio, based on the 56-year data. This represents the same 8.96 percent geometric mean used in the plan.

With 10,000 simulations applied against the identical withdrawal structure as the algorithmic calculation, the results were 48.82 percent successful and 51.18 percent failed simulations—a result very close to the 50/50 expectation using the algorithmic solution.

Figure 1 shows the cumulative distribution of returns from the 10,000 simulations and the percentage of successes and failures in each band of returns. As an example, the column above 9.26 percent illustrates that we had 2,317 simulations of the 10,000 total (23.17 percent) with an average return between 8.26 percent and 9.26 percent, of which 771 succeeded (7.71 percent) and 1,546 failed (15.46 percent). As expected, MCS has provided a distribution of returns around the CMA. We can also see that failed simulations occurred even though the average return of the simulation was as high as 10.26 percent to 11.26 percent (sequence risk).

Figure 1: Success/Failure Distribution

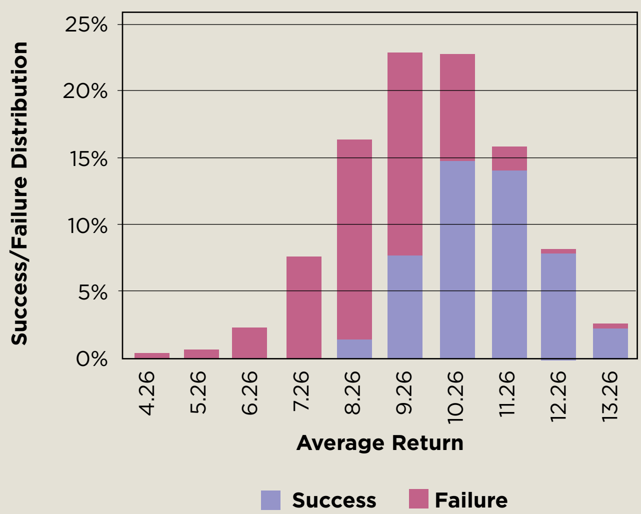

Table 1: Results with Same Estimated Geometric Mean

10,000 Simulations with Estimated Geometric Means of 8.96%			
Arithmetic Mean	**Standard Deviation**	**Success**	**Failure**
9.00%	2.84%	4921	5079
9.26%	8.07%	4882	5118
9.50%	10.45%	4841	5159
11.00%	20.63%	4727	5273

Because there are observably higher and lower returns, many advisors believe that this MCS represents a "test" of these different return ranges. This is not necessarily the case. As used, MCS totals the results of the 10,000 simulations and the overall average is still effectively 50/50. Although individual simulations may have been at higher or lower returns, they are all averaged back to a single probability of success, which is the same as the 8.96 percent geometric mean.

As an example, assume we randomly selected a sample from all the grade-five students in the country and applied a standard math test. In the sample class there are average students, some gifted learners, and some learning-disabled students. We report that the class average is 50 percent. Yes, the test was applied to the different types of students—but unless our analysis of the results included subset analysis, we cannot say we tested gifted learners, only that we tested grade-five students.

We can test if the MCS results are strictly a function of the geometric mean by generating other CMAs that have the same geometric mean. The paper uses the estimation technique (Bernstein and Wilkinson 1997),

$$\text{Geometric Mean} \approx \text{Arithmetic Mean} - \text{Standard Deviation}/(2(1 + \text{Standard Deviation})),$$

to calculate three additional arithmetic mean and standard deviation combinations, all of which result in a geometric mean of about 8.96 percent.

In Table 1, we can see four combinations of average return and standard deviation that result in a common 8.96 percent geometric mean. In other words, an arithmetic return of 11 percent +/– 20.63 percent has an estimated 8.96 percent geometric mean.

A criticism of an algorithmic solution would be that because the geometric mean is the same in all four scenarios, it would generate identical results. An expectation would be that Monte Carlo simulation should behave differently because of the significant differences in volatility. In fact, with 10,000 simulations used for each, we can see from Table 1 that there was only a very small variance in the success or failures, even when the volatility is increased several times in magnitude (2.84 percent to 20.63 percent). The impact on the analysis is minimal as long as the geometric mean remains the same.

If these four scenarios were presented to the client using MCS, they would all indicate the client has a 47–49 percent chance of success, even though common sense tells us these are dramatically different proposals. Although larger standard deviations may stretch out both "tails" of our distributions, when the cumulative success/failure impact is calculated, there is no significant difference.

Conclusion: MCS results are strictly a function of the geometric mean and not sensitive to the level of risk when used to report a single "probability of success." MCS will average out results so no "test" of the range of returns is occurring.

Recommendation: The geometric mean of each simulation could be calculated and the success/failure result subtotaled by this variable to allow reporting based on higher or lower returns.

Myth 2: MCS Tests Impact of Client Goals—Sequence Risk

Advisors believe that MCS tests and illustrates the impact of a client's specific withdrawal strategy and the further impact of the order of good and bad years of return.

A white paper titled *Understanding Monte Carlo Simulation,* on the Wealthcare Capital Management Web site, marshaled a convincing argument on the value of MCS, stating, "…the effect of the uncertainty in timing of various returns, contributions and goals can have far more impact than the average return or even the risk of the portfolio" (Loeper 2007).

Because this appears to be a widely held belief in the industry, the next objective was to test whether MCS will illustrate the impact of the timing of withdrawals, based on different client goals, and the impact on the order of returns or sequence risk. To do this, a variety of withdrawal strategies were tested. The key is to ensure that the present value of the distributions based on the 8.96 percent rate of return remains constant. This ensures an apples-to-apples comparison.

The same 20,000 simulations were applied to each of the following seven withdrawal strategies so we could ensure that any variance was a function of the withdrawal strategy, not the simulations.

- Even withdrawal of $52,159 indexed
- A $57,423 withdrawal for the first five years (about a 10 percent increase), then reduced withdrawal of $49,932 for the duration of the plan—withdrawing an additional $30,000 of the $800,000 capital up front
- Withdraw $65,000 for the first five years (25 percent more) and a reduced amount thereafter
- Withdraw $78,000 for the first five years (50 percent more) and a reduced amount thereafter
- Withdraw $104,300 for the first five years (100 percent more) and a reduced amount thereafter
- Withdraw an additional lump sum of $100,000 in the 12th year of the plan
- Withdraw an additional lump sum of $100,000 in the 20th year of the plan

From Table 2 we can see that regardless of the wide variety of withdrawal strategies, the failed simulations only varied from 51.7 percent to 52.3 percent. The same seven withdrawal strategies were then applied against an MCS with 1,000 simulations using the 11 percent return +/– 20.63 percent, which is an 8.96 percent geometric mean. The results ranged from 51.8 percent to 52.1 percent failures—exactly in line with the results from the base 9.26 percent, +/– 8.07.

This appears to fly in the face of popular belief. Why? If we look back at Figure 1, we can see that although the randomized returns of MCS generate some "ruin" scenarios where the returns ranged from 9.26 percent to 10.26 percent, it equally generated "success" scenarios based on better returns early on where the returns ranged from 8.26 percent to 9.26 percent. Sequence risk is not a two-headed coin and can equally result in simulations to the client's advantage or disadvantage. A more telling question is, why would it ever be expected that two statistical approaches (geometric mean and MCS), using the same assumptions, could have ever returned a different result? It would be frightening if they did!

Conclusion: Although the client's goals may influence an individual simulation, they have little or no impact on the success or failure of a full Monte Carlo simulation summing multiple simulations.

Observation: Recognizing that we could subgroup MCS results and report on failed simulations within return bands, the question becomes, of what practical use is this to the advisor and client? Does it help if we tell the client, "Your plan could succeed with a lower return or fail with a higher return"? Unless strategies exist that can reduce the risk of short-term poor returns without equally reducing the ability to take advantage of short-term great years, the only option is to reduce portfolio volatility overall, which will reduce the client's likelihood of achieving their goals.

Impact of Monte Carlo on Inflation Assumptions

The inflation assumption is not the variable that is "sensitive" in the financial plan, other than where you have revenue streams that are not similarly indexed. By randomizing the return, the inflation, or both, we are randomizing the real return, which generates the need for more or less capital.

A Monte Carlo with 10,000 simulations was run using the historical arithmetic mean for inflation of 3.99 percent, +/–3.3 percent. In each case, the withdrawal rates on the investments, government benefits and client expenditures were all indexed using the MCS derived inflation values. The results averaged 49.1 percent successful simulations and 50.9 percent failed simulations. In other words, there was no significant difference applying MCS.

Myth 3: Random Mortality Values Test for Longevity Risk

Proponents feel that planning must include a test for longevity risk—that is, the chance the client will outlive his or her money. A form of randomized mortality can be applied in a financial plan to act as a longevity test.

I can understand why advisors believe this is an important test by looking at the sensitivity of the plan to the mortality assumption. The original assumption was a mortality of age 85. If the mortality assumption is delayed by five years, a smaller amount

Table 2: Results with Different Withdrawal Strategies

Withdrawal Scenario	Success	Failure
Even withdrawal	48.2%	51.8%
10% increase initial years	48.1%	51.9%
25% increase initial years	48.1%	51.9%
50% increase initial years	47.9%	52.1%
100% increase initial years	47.7%	52.3%
$100,000 lump sum in 12th year	48.3%	51.7%
$100,000 lump sum in 20th year	48.2%	51.8%

Table 3: Impact of Longevity on Capital Required or Supportable Lifestyle		
Sensitivity to Mortality, Income		
Mortality	Capital Required	Goal Level
85	3,975	60,000
90	69,060	56,003
95	122,491	52,465
99	158,225	50,407

of the tax-sheltered funds is withdrawn each year and additional open capital must fund the difference.

In Table 3, we see that a five-year extension of our sample case creates a present value capital gap of $69,060. An additional 5 years, to age 95, adds an additional $53,431 on this amount to total $122,491. The table also records how this would affect lifestyle. As an example, ten extra years of longevity translates into a $7,535 reduction in annual, after-tax income.

As we can see from the mortality tables (Table 4), in a plan for a 60-year-old couple, planning to age 85 would be sufficient for 68.8 percent of males (100% – 31.2%) and 49.4 percent of females (100% – 50.6%).

Even though it is referred to as applying a "Monte Carlo" to mortality, the approach used was to randomly generate values between 1 and 87,813—the number of males alive at age 60 from the original 100,000. Each random value would then correspond to one of these 87,813 people getting a pink slip from the Grim Reaper. I checked where they stood on the list for all ages and determined the mortality. (Example: random life 796 or 795 would die at age 99.)

By using a randomized mortality based on tables, which appears to be the obvious way to illustrate longevity risk, one arrives back at the same dilemma as with MCS—the average washes out any testing of the extremes around it. The mortality randomization spreadsheet generated a series of 2,000 simulations, or mortality assumptions, with an average mortality age of 80.43, +/–0.7. In other words, the average mortality was almost at exactly the same place in the table where one-half of the 87,813 males remained alive, as should be expected.

Consider the sample case where we know that if the client lives beyond age 85, it would be a failed simulation, and to die sooner would result in surplus capital. By randomizing mortality, we should end up with 31.2 percent failed scenarios and 68.8 percent success—exactly the same as the mortality assumption we made.

The trend we experienced with returns continues with mortality. Although randomization may generate failures based on the client living to age 101 and running out of money, as many of the simulations could be that the client dies at age 61 and capital remains—a "successful" scenario (although the client may not see it that way!).

Conclusion: By doing an apparent randomization to represent longevity, advisors are, on average, basing their planning on a lower mortality assumption than would be used in traditional plans. Further, when analyzed in isolation, the randomization should return the identical probability as simply referencing the mortality table for the age at which the plan fails in an algorithmic solution.

Table 4: Sample Mortality Based on Canadian Tables (1997)				
Mortality Tables, Canada				
Age	Male	% Gender	Female	% Gender
0	100,000		100,000	
60	87,813	100.0%	92,757	100.0%
77	44,873	51.1%	71,498	77.0%
85	27,402	31.2%	46,934	50.6%
90	12,409	14.1%	27,303	29.4%
99	796	0.9%	3,402	3.7%

Based on Statistics Canada results from 1997, assuming 100,000 lives at birth, the table illustrates the number of individuals still alive at different ages.

Development of an Algorithmic Replacement for MCS

In this first section, I documented that Monte Carlo simulation returns the same results as the algorithmic solution if the same assumptions are used and the total successful/failed results across all simulations are summed. Regardless, many advisors, media, and academicians are still attracted to MCS. One remaining feature of MCS that may account for this popularity is the ability of MCS to illustrate a partial probability of success.

For example, using a geometric mean Table 5 shows the distribution in a financial plan for a 60-year-old male with assumed mortality at age 85. In simple terms, we know that if the client lives longer than expected, with worse returns than we forecast, the plan will fail. The challenge is, if the client lives longer than expected but also has a higher return, we cannot determine success or failure.

A traditional algorithmic approach using the geometric mean identifies the 50/50 solution. If it is not a geometric mean and just a point rate of return, the algorithmic approach becomes a "Succeed/Fail," "Surplus/Deficit" result. What MCS does that traditional algorithmic solutions do not is provide a breakdown of the 50 percent "unclear outcome" so the advisor is provided a probability of success that is unambiguous.

I applied MCS with both returns and randomized mortality assumptions to our sample case. The results of 10,000 simulations were an average of 72.9 percent successful and 27.1 percent failed scenarios, with a standard deviation of 1.57 percent.

Of the 72.9 percent successful results, many of these simulations could represent scenarios such as "the client lives to 100 with higher than expected returns" or "the client dies at age 62 with lots of money still in the bank." The client indicated a life expectancy of 85 and the MCS results provide no information how likely they are to achieve their goals using that assumption.

Generating an Algorithmic Model

Milevsky and Abaimova (2006) state, "While running the requisite 100,000 scenarios that would provide a minimal margin of error is computationally not feasible within most real-time engines, running as little as a few hundred scenarios can be woefully inadequate."

Although large numbers of simulations may be feasible with simpler tests (such as a pretax depletion of capital), as the complexity and sophistication of the planning model increases (for example, tax calculations, integrated government benefits, multiple goals, and so on), the computational burden increases dramatically. Each simulation requires a complete iteration of the planning algorithm.

Daryanani (2002) states, "In Monte Carlo simulation, the accuracy of the results is proportional to the square root of the number of iterations (trials)…the number of iterations with sensitivity simulation is fixed at twice the number of input variables."

Milevsky and Abaimova analyzed ten Monte Carlo simulation engines by applying a common case with as close to the same assumptions as possible in each tool. Their results: "The lowest sustainability number was 48 percent and the highest was 88 percent." They address the wide range of results and state, "The true reason for the divergence of results is more complicated and subtle. In fact, we traced the range of results to a number of factors,…" (Milevsky and Abaimova 2006). They then go on to list several reasons for the variation. It is often difficult to audit algorithmic solutions—let alone systems that merge the complexity of multiple MCS components into a single answer.

In pursuing an algorithmic solution, I had four objectives:

1. To generate the same partial probabilities as returned by MCS
2. To generate more accurate and faster results than provided by MCS
3. To generate more useful (granular) results for the client and advisor
4. To create a more supportable model

Table 5: Success/Failure Likelihood Based on Algorithmic Solution

	Pre-decease Age 85 (68.8%)	Outlive Age 85 (31.2%)
Returns Above 8.96% (50%)	A 34.4% chance the client will die before age 85 and have higher than projected returns (Surplus – Success)	A 15.6% chance the client will outlive age 85 with higher than projected returns (Unclear outcome)
Returns Below 8.96% (50%)	A 34.4% chance the client will die before age 85 but have lower than projected returns (Unclear outcome)	A 15.6% chance the client will outlive age 85 with lower than projected returns (Capital shortfall – Failure)

In developing an algorithmic approach, the first step is to realize that *the probability distribution of returns and client mortality are independent of the plan itself.* In other words, client mortality or the possible distribution of returns are not influenced by when the client retires or when the pension starts to be paid. Earlier, I demonstrated that the MCS results do not vary from the algorithmic solution when properly applied. As a result, we can extract the randomized variables and test for success algorithmically.

I applied a methodology that I have called a "reliability forecast." This name is taken from the approach used by 19th century maritime insurance companies trying to predict the failure rate of their ships coming in. In our case, the ship is our life and the route is our financial plan. The shoals are the uncertainty of some of our underlying assumptions like mortality and rates of return.

The first step was to create a probability matrix of the likelihood of attaining a specific age and a specific rate of return (see a partial matrix in the Appendix).

$$\text{Probability of Death}[40] \times \text{Probability of Return}[17] = \text{Probability Matrix}[40,17]$$

The probability of death for any given year is calculated by referencing mortality tables. For a 60-year-old male there were 87,813 of 100,000 alive at the beginning of the year. By age 61 there were 86,842. This means that for 60-year-old males, 971 of 87,813, or 1.105 percent, will die before age 61.

To calculate the probability of the returns, I used the lognormal deviate for the returns:

$$[\ln(1 + \mathrm{T}) * N - \ln(1 + R)*N]/[\ln(1 + \mathrm{Std.Dev}) * \mathrm{SQRT}(N)]$$

where T is the target return and R is the geometric mean (Kritzman 1995). This was then used to calculate the probability of achieving that return using the NORMSDIST function in Excel, which is the standard normal density function. For the initial trial, I used 0.5 percent increments on the return and assumed 25 years for N. Taking the difference between each return increment then gave the probability of achieving that particular return. By adding greater granularity on the rate-of-return axis, one can increase the sensitivity of the reliability forecast.

Looking at the appendix, if I have calculated a 0.7 percent chance of having greater than a 13.46 percent return and a 1.7 percent chance of the client dying at 93 years of age, the likelihood of the client dying at age 93 with that return is 0.7 percent × 1.7 percent, or 0.012 percent. The extremes in the grid are inclusive of any "tails"—that is, any return greater than 13.46 percent or mortality at any age 100 or greater, and so on.

This matrix is displayed in a histogram (Figure 2) to allow a visual representation. Each cell represents the likelihood of dying at that specific age with that specific realized return. As illustrated, the highest single likely outcome is that the client dies around age 80 to 84 with the expected rate of return of 8.96 percent. If we summed all

Figure 2: Probability Distribution Matrix with Success/Failure Boundary Line

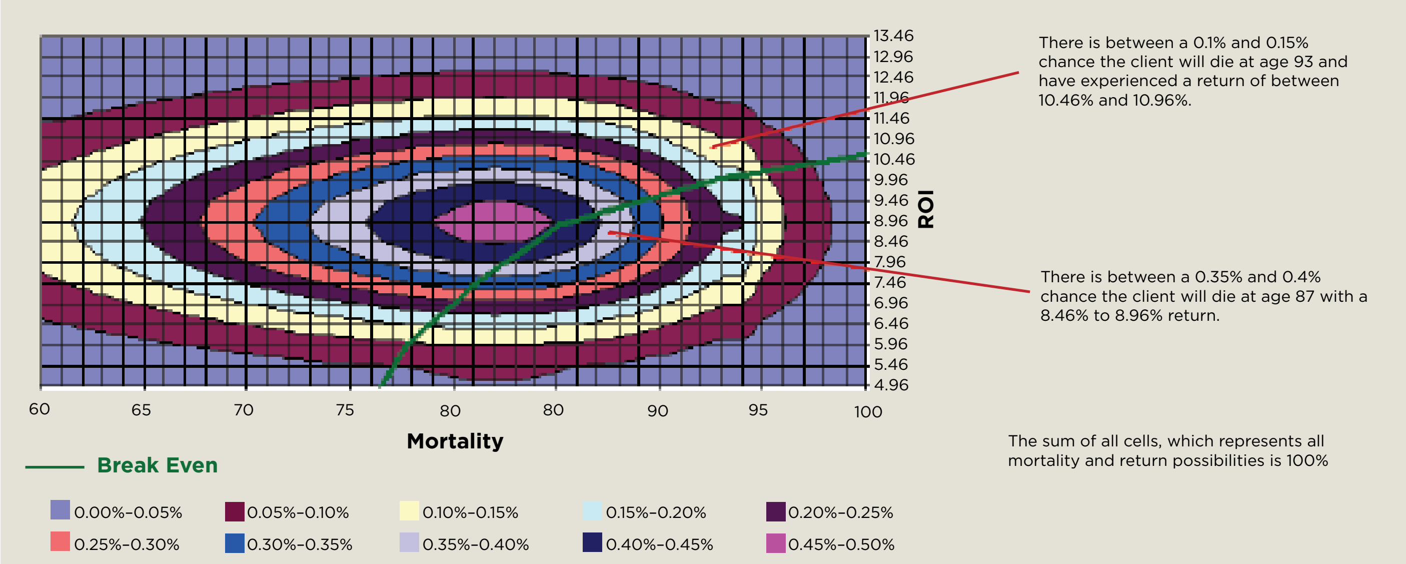

Table 6: Age at Which Funds Exhausted Based on Various Return Assumptions

Return	4.96	5.46	5.96	6.46	6.96	7.46	7.96	8.46	8.96	9.46	9.96	10.46	10.96
Age Funds Depleted	76.5	77	78	79	80	81	82.5	84	85	89	93	99	100+

cells, it would add to 100 percent, since there is a 100 percent chance the client will die with some rate of return.

With clients of the same age and similar portfolios, the grid would be identical. The probability matrix is strictly a function of mortality tables, based on the client's age, and the probability of achieving specific returns, based on the capital market assumptions.

If we ran 100,000 Monte Carlo simulations of the returns, and randomized the mortality, then did a density scatter-gram of the age and return used in each simulation, we should end up with exactly the same graphic!

The next step was to use the same calculation we used for the sample case to solve for the age at which funds are depleted based on each of the rates of return in the grid.

Table 6 shows that if the client achieved a 5.96 percent rate of return, funds would be depleted at age 78. With a 10.46 percent return, funds would be depleted at age 99.

For the 13 rates of return between 4.96 percent and 10.96 percent, I calculated the age that funds ran out, after which the returns always exceeded any life expectancy. This is the boundary line between success and failure using an algorithmic calculation.

The green line in Figure 2 represents the series of points where mortality age and return are balanced with no shortfall or surplus. A cell to the lower right of the green line indicates an outcome where the client has outlived his or her money, or a failure. All of the squares left or above the green line represent successful outcomes. If I total all the squares in either quadrant demarcated by the green line, it represents the probability of success or failure.

Applying this methodology resulted in 74.5 percent success and 25.5 percent failure, which compares with 72.9 percent, +/– 1.57 percent with MCS using 10,000 simulations.

The biggest advantage of the reliability forecast is precision. The matrix is composed of 40 ´ 17 cells. Each time you run a single Monte Carlo simulation, it allows you to "paint" one dot, or pixel, of the picture. If you ran 100 simulations, it would be insufficient to even put one dot in every cell. It might take 4,000 to 5,000 Monte Carlo simulations to get the equivalent to one decimal precision and 100,000 before the picture actually starts to take similar shape and clarity. If the age range is larger or the standard deviation of the portfolio greater, there would be more cells requiring more simulations to arrive at a similar picture.

Testing the Algorithmic Solution

Although the first results comparing the reliability forecast and MCS were astoundingly close, I tested further and applied two of the withdrawal strategies used in Myth #2 (one with greater upfront withdrawals and one with a lump-sum withdrawal):

- Withdraw $78,000 for the first five years (50 percent more) and a reduced amount thereafter.
- Withdraw an additional lump sum of $100,000 in the 12th year of the plan with reduced withdrawals throughout.

Table 7: Sample Improved Sensitivity Analysis

Sensitivity Analysis			
Age	**Probability**	**Success**	**Failure**
70	80.7%	100.0%	0.0%
75	65.7%	97.8%	2.2%
80	47.1%	83.4%	16.6%
85	27.4%	68.3%	31.7%
90	11.5%	58.9%	41.1%
95	2.6%	54.6%	45.4%
Return	**Probability**	**Success**	**Failure**
11.5	22.4%	89.2%	10.8%
10.5	37.2%	77.8%	22.2%
9.5	54.4%	68.7%	31.3%
8.5	70.9%	61.9%	38.1%
7.5	84.0%	57.9%	42.1%

Probability of Success					
Age/Return	**7.5**	**8.5**	**9.5**	**10.5**	**11.5**
70	100.0%	100.0%	100.0%	100.0%	100.0%
75	100.0%	100.0%	100.0%	100.0%	100.0%
80	88.6%	92.3%	97.6%	100.0%	100.0%
85	73.6%	78.0%	85.2%	93.4%	100.0%
90	63.8%	68.0%	75.1%	84.2%	94.7%
95	59.2%	63.2%	70.1%	79.3%	90.6%

When the rates of return were used to test for when funds would be depleted (our boundary line for success), it resulted in identical ages for each scenario. This should not come as a surprise as the present value of these withdrawal strategies was identical to the initial withdrawal strategy. So, the algorithmic solution still predicts exactly the same result of 74.5 percent.

Running 1,000 simulations with both mortality and returns and the $78,000 withdrawal strategy resulted in 73.7 percent success and 26.3 percent failure. The MCS with the lump-sum withdrawal in the 12th year resulted in 73.5 percent successful and 26.5 percent failed simulations.

The next test was to alter the algorithmic solution to assume an accelerated withdrawal of $58,633, indexed to deplete all funds by the end of the 80th year. This resulted in a new boundary set to be applied to the reliability forecast and a projected likelihood of 59.5 percent success and 40.5 percent failure. The same withdrawal strategy using MCS and 1,000 simulations resulted in a 60.3 percent chance of successful and 39.7 percent failed simulations.

The final test was to randomly select three factors. I arbitrarily set the withdrawal at $65,000, the arithmetic return at 10 percent, and a standard deviation of 11 percent. I used an estimation technique to arrive at a 9.5 percent geometric mean. After adjusting the probability distribution for the returns (a new mean and standard deviation), the planning tool was used to calculate the boundary set where funds were depleted. The reliability forecast resulted in a 51.6 percent success and 48.4 percent failure rate. The MCS analysis was applied using the new capital market assumptions and 10,000 simulations. This resulted in a 51.2 percent success +/–1.42 percent.

Conclusion: The use of the reliability forecast provided the same probability of success as MCS with a dozen iterations compared with tens of thousands. (This achieved my first two objectives in developing an algorithmic approach.)

The Next Level: Proper Sensitivity Analysis

A financial plan assists the planner and the client to make decisions on trade-offs. If the planning horizon is extended from age 85 to age 90, given that the other variables remain fixed, the client might need to reduce the withdrawal amount from $52,159 to $48,684—a noticeable drop in after-tax income.

The question then becomes, is the client concerned enough about longevity that they are prepared to reduce their after-tax income immediately by $3,500 a year to ensure they have enough capital to support the same lifestyle until they are 90 years old? There is no correct answer to this question—just a properly informed decision by the client.

The MCS approach introduced a different type of ambiguity because the nature of the trade-offs is not evident to the planner or the client. Giving the client a simple 70 percent success and 30 percent failure provides no context in which to make a decision, an observation made in numerous papers critical of Monte Carlo simulation. What percentage of success is sufficient for the client?

Using the reliability forecast, it is a simple matter to subtotal the probability distribution in whatever manner desired to illustrate various likelihoods (as illustrated in Table 7). This could illustrate the probability of success or failure at any target planning horizon, any rate of return, or a matrix that displays the sensitivity of success based on age or return. The histogram itself provides a visual depiction of all of these factors.

Conclusion: Rather than a single probability of success, the reliability forecast can easily be subtotaled in different ways, providing a true understanding of the sensitivity of specific variables like mortality or returns. (This achieved our third objective in developing an algorithmic approach.)

Observation: As outlined throughout this paper, the same results could be achieved using Monte Carlo simulation by altering current practice, calculating the mean returns from each simulation, and subtotaling successes/failures in appropriate bands, albeit at an even greater compu-

Table 8: Summary of Pros/Cons Monte Carlo Versus Reliability Forecast

Issue	Monte Carlo	Reliability Forecast
Return assumptions	• Based on arithmetic mean • Estimation technique can convert • Advisors do not understand	• Based on geometric mean, same as financial plan
Test higher and lower returns	• Not as utilized • Requires additional computation and subtotaling of results in return bands • Misunderstood	• Easily subtotaled
Test impact of client withdrawals and order of returns	• Misunderstanding • Does not occur	• Not calculated
Test longevity using randomized mortality	• Not as utilized • Could add additional layer of subtotaling	• Easily subtotaled
Precision	• Poor due to computational load • Require 100,000 simulations	• High • 12 calculations are necessary to generate the boundary line
Ability to generate overall probabilities of success	• Yes	• Yes
Ability to provide improved granularity on sensitivity analysis	• Requires additional computation and subtotaling logic	• Yes, easily subtotaled
Ability to maintain the planning tool	• Complex with integration of randomized variables to plan algorithm • May restrict certain capabilities	• Simpler • The algorithmic solution is easier to maintain and audit

Appendix: Sample Probability Distribution Matrix

Mortality	Age/Return	13.5	13.0	12.5	12	11.5	11.0
1.1%	60	0.008%	0.009%	0.016%	0.028%	0.044%	0.065%
1.2%	61	0.008%	0.009%	0.18%	0.030%	0.048%	0.071%
1.3%	62	0.009%	0.010%	0.019%	0.033%	0.053%	0.078%
1.5%	63	0.010%	0.011%	0.021%	0.036%	0.058%	0.085%
1.7%	93	0.012%	0.014%	0.025%	0.044%	0.069%	0.102%
1.7%	94	0.012%	0.013%	0.025%	0.043%	0.068%	0.100%
1.1%	95	0.008%	0.009%	0.016%	0.028%	0.045%	0.066%
0.9%	96	0.006%	0.007%	0.013%	0.022%	0.035%	0.051%
0.7%	97	0.005%	0.005%	0.010%	0.017%	0.026%	0.039%
0.5%	98	0.003%	0.004%	0.007%	0.012%	0.019%	0.028%
0.3%	99	0.002%	0.003%	0.005%	0.008%	0.013%	0.019%
0.2%	100	0.002%	0.002%	0.004%	0.006%	0.010%	0.015%
100%		13.46	12.96	12.46	11.96	11.46	10.96
Distribution of Returns		0.7%	0.8%	1.5%	2.5%	4.0%	5.8%

tational cost than today.

Refining the Planning Model

The importance of the reliability forecast approach in isolating the probability distribution of the assumptions from the intricacies of the planning model is critical. By separating these two components, *any algorithmic solution can be used to calculate the boundary line.* The better the algorithmic solution, the more reliable the results.

The industry is continuing to evolve with new insights and observations that allow us to understand the client's behavior and model it to help make decisions. As an example, if our planning model uses MCS-derived inflation assumptions, it becomes more difficult to be able to reflect insights like those of Tacchino and Saltzman (1999) on non-inflationary withdrawal rates in retirement, unless we create layers of complexity that will be impossible for any advisor to understand. A reliability forecast approach allows firms to focus on ensuring that the base model of their planning tools reflects the best understanding of investments, taxation, retirement spending, and more.

Having opened the door on the use of a reliability forecast to replace Monte Carlo simulation, there are a variety of extensions to this original research that are warranted, such as optimal cell size in the matrix and the assumed number of periods in the calculation of return distribution.

Summary of Findings

The reliability forecast should be used instead of Monte Carlo simulations in many cases. Because every client with the same age/sex/mortality and risk/return/portfolio will in turn have the same probability distribution, the algorithmic model allows us to extract these variables and then apply client-specific planning requirements to determine the boundary between success and failure. As a result, the reliability forecast can be faster, more precise, less subject to error, and less subject to misunderstanding. Any tool or test can be used to calculate the algorithmic boundary between success and failure. Table 8 summarizes the pros and cons of MCS compared with the reliability forecast.

Firms should focus on ensuring that the underlying planning model they use is the best the industry can provide. It should accurately model a planning scenario for a client that incorporates tax sensitivity, the best understanding in income distributions during retirement, portfolio theory, and whatever other innovations research supports.

JFP

References

Bernstein, William J. and David Wilkinson. 1997. "Diversification, Rebalancing, and the Geometric Mean Frontier." November 24: http://www.effisols.com/basics/rebal.pdf.

Daryanani, Gobind. 2002. "Sensitivity Simulations: A Faster Alternative to Monte Carlo." *Journal of Financial Planning* 15, 9 (September): 104–116.

Kritzman, Mark. 1995. *The Portable Financial Analyst.* The Association for Investment Management and Research.

Loeper, David. 2007. "Understanding Monte Carlo Simulation," Wealthcare Capital Management.

Milevsky, Moshe A. and Anna Abaimova. 2006 "Will the True Monte Carlo Number Please Stand Up?" *Journal of Financial Planning: Between the Issues* July: http://www.fpanet.org/journal/BetweenTheIssues/Essays/071506.cfm.

Tacchino, Kenn and Cynthia Saltzman. 1999. "Do Accumulation Models Over-State What's Needed to Retire?" *Journal of Financial Planning* February: http://www.fpanet.org/journal/articles/1999_Issues/jfp0299-art9.cfm.