

## Will the True Monte Carlo Number, Please Stand Up?

By: Moshe A. Milevsky and Anna Abaimova

Revised Date: June 12, 2006

The recent best-selling book by Lee Eisenberg titled *The Number* (published by Free Press, 2005) has focused national attention on a sum of money retirees need to live comfortably for the rest of their lives. According to Eisenberg, depending on one's desired life-style and income aspirations this number ranges from a few hundred thousand dollars to the millions. Many readers of the book have been surprised at how high this number can actually be. In fact, a PBS show that aired in mid-May entitled: "Can Americans Afford to Retire?" is just another indication of the national preoccupation with financial sustainability towards the end of the human lifecycle.

Of course, as most investment advisors have known for years, a retirement number – if it actually exists – is vague and imprecise since it depends on many economic unknowns, especially future equity market returns. After all this number" must be invested somewhere in order to produce income – and the portfolio return process is inherently random. This is precisely the reason so-called Retirement Income Monte Carlo (RIMC) simulations have become ubiquitous in financial planning. Numerous websites, software packages and vendors have devoted themselves to computing "the probabilistic number", namely the odds that a given dollar nest egg will last for the rest of a retiree's life.

Unfortunately – as we have found out the hard way – these estimates can be all over the map.

We set out to investigate the extent to which RIMC simulators – at least the ones that are widely available to the public – provide different answers, to what appears to be the same question which is: *Will my number last?*

To do this in a quasi-scientific manner, we hypothesized a 55-year old single male who is contemplating early retirement with a nest egg of \$500,000 from which he would like to withdraw \$25,000 adjusted for inflation each year. His entire nest egg was assumed to be invested and rebalanced in a portfolio of diversified equities which was projected to earn an arithmetic average 7% (after inflation) each year, which is equivalent to a geometric average of 5% with a standard deviation or volatility of 20%. Our test case retiree was assumed to be in average health and with no other source of retirement income. Obviously this situation might not represent the typical user of a RIMC simulator – since at the very least their desired \$25,000 per year should be supplemented by Social Security -- but our objective was to design a benchmark case that could be used across programs. We tried a variety of other cases and the results – which we will describe in a minute -- were generally the same.

Despite the inherent uncertainty, we embarked on this particular exercise with a pretty good idea of what the output or “the probabilistic number” should be. Most readers might not realize this, but there are unique situations in which it is possible to obtain the probability of running out of money before running out of life, with no simulations required! In fact, the above mentioned test case we constructed is an ideal candidate.

Let us explain how one can get analytic solutions to a randomness problem by examining a slightly different question. What is the probability the total return from the SP500 index will be greater than 5% in 2006? There are a number of philosophical approaches to dealing with such a question. One is to literally “dump” the one-year historical returns of the SP500 index “in a hat” and then sample from this hat a large number of times and then count the frequency with which the sample average exceeded 5%. This approach is at the heart of Monte Carlo. But, as many mathematicians know, another approach is to fit a curve (for example, the normal distribution) to long-term historical returns and then evaluate the curve – a.k.a. compute the tail probability -- at the 5% mark. Some readers will recognize this as the difference between parametric (the latter) and non-parametric (the former) approaches to forecasting. Interestingly,

when we adopt the parametric approach and assume a constant withdrawal rate over a random retirement horizon, one can also recover an analytic expression<sup>1</sup>. In plain English, in simple cases we know the odds of retirement ruin and sustainability without generating any random numbers – just like we know the odds of tossing 4 heads in a row on a fair coin is 1/16.

Thus, using this analytic methodology the *sustainability* – for a 55 year old who starts withdrawing a constant inflation adjusted \$25K from a \$500K retirement nest egg -- is 75%, which means that the *probability of ruin* is 25%. Remember that although Monte Carlo generators are based on randomized returns or spinning roulette wheels, the randomly generated results should only differ from each other by a statistical margin of error. In fact, that is the point of a properly designed Monte Carlo. The odds should eventually converge to the truth. Yet, our preliminary analysis of over ten different calculators – six of which are publicly available and listed in the accompanying table -- revealed a wide range of results for our 55 year-old retiree. The lowest sustainability number was 48% and the highest was 88%.

### **Why were the output results so different?**

Our first reaction was to blame the programmer or manufacturer (from here on: sponsor) of the RIMC for building a faulty product. Like all high-tech gadgets on the market some are better than others; but that is a simplistic knee-jerk reaction. 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, which leads us to the following insights.

(1) Every sponsor has their own view on the dynamics of equity markets and future stock prices. Some believe that equity prices obey random walks with no dependence (i.e. zero serial correlation) from year to year while others take a subjective view on the so-called equity premium and thus implicitly opine on the market's current valuation

---

<sup>1</sup> For readers who are interested in precise what this formula looks like and how this formula is derived, please see chapter 9 of the book: *The Calculus of Retirement Income: Financial Models for Pension Annuities and Life Insurance* by Moshe A. Milevsky, Cambridge University Press (2006)

levels. Along the same lines, different sponsors have their own correlation assumptions for asset classes. Some define equity/stocks as essentially being the SP500 index while others take a broader Russell 3000 view of portfolio allocation. All of this makes an obvious and important difference.

(2) Some sponsors consider retirement to have a random horizon based on various population or annuity mortality tables while others use Social Security Administration projections. Other sponsors or RIMC programs pick an arbitrary 25 or 35 years of retirement and ignore this important source of longevity randomness altogether. In some cases, a 90<sup>th</sup> percentile of the mortality table is used as a fixed horizon. Of course when these RIMC generators are used for couples as opposed to individuals, the results can be even more diverse.

(3) Some programs have sophisticated models for the term structure of interest rates (i.e. the yield curve) which drives bond returns while others assume a basic random walk process similar to the one used for equity prices. Indeed, it is debatable whether a sophisticated three factor model of the LIBOR swap curve is any better than a coin toss in predicting bond yield 25 years from now. But then again, one is left with an uncomfortable sense that something is missing when the RIMC simulation results don't care that the yield curve is on the verge of inverting or that long-term interest rates remain stubbornly low. In other words, everyone has their own answer to the question: how does a bond fund behave?

(4) The cost of living in retirement is often confused with the general level of macroeconomic inflation in the economy. Note that for many years that Bureau of Labor Statistics (BLS) has been computing a CPI-E, which is an experimental inflation index for retirees, which differs from the CPI-U (urban) or the CPI-W (wage earners). The projected increase in a retiree's annual expenditures might only be loosely correlated to the widely quoted CPI. In this vein, some programs explicitly (i.e. stochastically) model the evolution of inflation and then increase withdrawal rates by this amount each year

while others implicitly model the real after-inflation evolution of portfolio returns. While both methodologies have their merits, the end results will obviously differ.

(5) One factor that does affect the dispersion of results, but for which we can provide no justification is the use of a small number of scenarios or random numbers. While running the requisite 100,000 scenarios that would provide a minimal margin of error is computationally infeasible within most real-time engines, running as little as a few hundred scenarios can be woefully inadequate. Tails will be ignored and results will be biased by optimism.

### **What should be done about all this randomness?**

Some critics might view this as a critical flaw in the Monte Carlo methodology. We obviously don't think the baby should be thrown out with the bathwater. Other participants in the industry might not view this divergence of estimates as a problem or concern at all and will rightfully claim that every software developer or "quant" is entitled to their own opinion on how to model asset classes, mortality or serial dependence. In fact, it is not our objective to criticize any particular set of assumptions per se. Obviously, every RMC vendor or developer can justify their own assumptions and approximations. But, akin to widely reported economic forecasts, or even weather predictions themselves, the market will eventually reward those who get it right and punish those who are consistently wrong or provide unrealistic predictions. Indeed, we do have some sympathy with this free-market view.

Unfortunately, we do not believe the nascent retirement income planning industry can afford the luxury of ignoring the divergence of Monte Carlo results. First of all, many outsiders who are unfamiliar with the subtleties of simulations might erroneously interpret the dispersion of results as a flaw in the Monte Carlo methodology itself; instead of a powerful technique for measuring and explaining risk. More importantly, we suspect that one of the key players in this industry – namely security regulators such as the NASD, SEC or even state insurance departments – will remain deeply skeptical of

Monte Carlo forecasts and illustrations. Remember that these groups are frequently called upon to approve marketing and presentation material containing results of these simulations. If they continue to observe statistically divergent results on seemingly identical inputs, or even worse, results that are biased in the favor of the product that is implicitly being promoted, they will (justifiable) continue to frustrate the efforts of those who believe Monte Carlo is an ideal pedagogical tool.

In our opinion, one possible solution to this conundrum would be for the relevant stakeholders in the retirement income industry to develop a set of calibration tests (or certification procedures) for which all RIMC programs would generate the same results, within a statistically tolerable margin of error. This approach has been used quite successfully in the field of solvency and capital requirements, where the insurance or banking regulator allows a financial entity to use their own internal models to set capital requirements with the proviso that certain statistical calibration points must be met. We envision the same idea in the retirement income industry. The analytic formulas we described above – for some basic cases -- can be used to compute sustainability and ruin probabilities with full precision. This suggestion is akin to the age-old method of testing a weight scale by using dummy 1, 5 and 10 pound weights to calibrate accuracy. There is obviously no guarantee the scale will tell the truth when used with any other weights, but the odds are pretty good.

In sum, while attempting to impose an industry-wide standard or benchmark for RIMC results might be fraught with difficulty, we believe that developers and promoters of RIMC engines should calibrate their output to a number of baselines scenarios. Obviously, the industry as a whole stands to benefit from less randomness in Monte Carlo.

**About the authors:** *Moshe A. Milevsky, Ph.D. is the Executive Director of the IFID Centre and Associate Professor of Finance at the Schulich School of Business at York University in Toronto, Canada. Anna Abaimova is a Research Associate at The IFID Centre and a recent graduate of the Schulich School of Business.*

**Table #1:**

This table is part of a preliminary study of retirement income Monte Carlo (RIMC) calculators. We tried to adhere as closely as possible to the following input assumptions when comparing the various programs. Specifically, we assumed a generic single 55 years old male who withdraws 5% of his initial retirement nest egg that is invested in an all-equity portfolio expecting to earn a real (after-inflation) 7% per annum (geometric mean of 5%) with a volatility of 20% after all investment expenses and management fees. If required, we assumed an inflation rate of 2% and the RP2000 mortality table. The analysis ignores income taxes or required minimum distributions (RMD) and we did not assume any additional pension or annuity income. In this table we only include results from RMC generators that are publicly available free of charge and have listed the URL. The description of the program/host is in our own words. An updated version of this study will be available on [www.ifid.ca](http://www.ifid.ca).

Host and Calculator	URL	Description/notes	Test case result (probability plan is sustainable)
Gummy-Stuff: Gummy's Monte Carlo Safe Withdrawal Calculator (By Peter Ponzo)	<a href="http://www.gummy-stuff.org/sensible_withdrawals.htm">http://www.gummy-stuff.org/sensible_withdrawals.htm</a>	Post-retirement Spreadsheet Monte Carlo simulations (number of iterations chosen by user); allows for asset allocation, annuities; uses historical returns; retirement period is specified by user; inflation can be specified or is randomized (using historical successive rates); output: % of simulations where the money did not run out before specified end of the period	88% chance investment will last 25 years, 85% chance portfolio will last 30 years
Moneychimp: Monte Carlo Retirement Calculator	<a href="http://www.moneychimp.com/articles/volatility/montecarlo.htm">http://www.moneychimp.com/articles/volatility/montecarlo.htm</a>	1000 iterations; accounts for pre-retirement saving and post-retirement spending phases; no inflation parameter; output: probability investment will last for a user specified retirement period, given a user specified return and volatility	73% chance portfolio will last 25 years, 67% chance investment will last 30 years
Retire Early: Standard or Advanced FIRECalc	<a href="http://capn-bill.com/fire/">http://capn-bill.com/fire/</a>	Post-retirement; allows for asset allocation and social security income; uses, but does not simulate (up to roughly 250) historical return paths; inflation – linked to PPI or CP or user specified; output: probability of not running out of money before the end of specified period; allows for change in	Standard FIRECalc: 81% chance investment will last 25 years, 75% chance portfolio will

		withdrawals and sale of major assets; provides link to a longevity calculator to estimate probability of surviving to a given age. Standard FIRECalc is the simplified version.	last 30 years Advanced FIRECalc: Results vary. e.g. 50-85% chance investment will last for 25 years, 30-70% chance investment will last for 30 years
Society of Actuaries: Retirement Income calculator -Beta Version (By William Leslie)	<a href="http://www.soa.org/ccm/content/areas-of-practice/finance/research/retire-income-calc-beta-version/">http://www.soa.org/ccm/content/areas-of-practice/finance/research/retire-income-calc-beta-version/</a>	Allows for pre-retirement period (savings and asset allocation), variable and fixed (COLA adjusted), deferred and immediate annuities, social security, pensions; user specified asset class return parameters; models inflation; output: table of forecasted scenario outcomes – real wealth levels - based on 200 different sequences of returns for each age group or possible end-of-retirement points (based on mortality tables).	80%
Society of Actuaries: Retirement Probability Analyzer (by Moshe Milevsky and The IFID Centre)	<a href="http://www.soa.org/ccm/content/areas-of-practice/special-interest-sections/pension/retirement-probability-analyzer-software/">http://www.soa.org/ccm/content/areas-of-practice/special-interest-sections/pension/retirement-probability-analyzer-software/</a>	Uses numerical partial differential equation approach; allows for changing asset allocation and variable and fixed (COLA adjusted), annuities, pensions; user specified yield curve (real and nominal) and asset class parameters; output: probabilities of ruin and of various wealth levels within lifetime and 10,20,30 year points; projected income in retirement, survival probability statistics.	74%
T. Rowe Price: Retirement Income Calculator	<a href="http://www3.troweprice.com/ric/RIC/">http://www3.troweprice.com/ric/RIC/</a>	-500 simulations -post retirement; gives an assessment of whether monthly income goal will be met, for a desired simulation success rate; suggests strategy changes; user specified retirement length, allows for some asset allocation, marital status; assumes 3% inflation for withdrawals	80%-90% chance the investment will last 25 years, 70%-80% chance the investment will last 30 years