



Does the “Bucket Approach” Destroy Wealth?

January 7, 2019

by Larry Swedroe

The “bucket approach” to retirement planning has been routinely adopted by financial planners, ever since it was popularized by Harold Evensky. Clients keep several years of assets in safe, liquid investments, while investing the rest of their portfolio more aggressively. But new research shows that this approach actually destroys a portion of clients’ wealth.

This research comes from [Javier Estrada](#), a professor of financial management at the IESE Business School in Barcelona, Spain. Before we get to Estrada’s research, let’s review how the success and failure of a financial plan is measured using Monte Carlo simulations.

The challenge of measuring failure rates of financial plans

Retiring without sufficient assets to maintain a minimally acceptable lifestyle (which each person defines in their unique way) is an unthinkable outcome. That’s why, when investors are planning for retirement, their most important question is: How much can I plan on withdrawing from my portfolio without having a significant chance of outliving my savings?

The answer generally is expressed in terms of a safe withdrawal rate (SWR) – the percentage of the portfolio you can withdraw the first year with future annual withdrawals adjusted for inflation. While historical returns can provide insight, it’s critical that investors not simply project the past into the future. Current valuation metrics should be used instead.

Another problem that must be considered is that investment returns are not constant, and systematic withdrawals during bear markets cause portfolio values to fall to levels from which they may never recover. Over the 26-year period from 1973 through 1999, the S&P 500 Index provided a nominal return of 13.9% a year and a real return of 8.2% a year. With hindsight, you might think that, given the 8.2% real return, it would have been safe to withdraw 7% a year in real terms (that is, take 7% of the portfolio’s starting value and increase the withdrawal by the inflation rate each year). Unfortunately, had you retired at the end of 1972 and followed that strategy, you would have run out of funds within 10 years, by the end of 1982. This was because of “sequence of returns risk.” Adverse returns in the early stages of retirement are particularly harmful; the S&P 500 Index lost almost 40% in the 1973-1974 bear market.

Additionally, investors must address our limited ability to estimate future returns and the fact that the sequence of returns matters a great deal. The way to solve for these problems is to use a Monte Carlo simulator. Monte Carlo simulations require a set of assumptions regarding time horizon, initial

investment levels, asset allocation, withdrawals, rate of inflation and, importantly, the distribution of and correlations between annual returns for various asset classes.

Two numbers determine the expected final wealth distributions in Monte Carlo simulation programs: the average annual return (again, derived from current valuations/yields rather than historical ones) and the standard deviation of the average annual return. The Monte Carlo simulator then randomly selects a return for each year and calculates wealth values over the expected retirement period. This process repeats thousands of times to calculate the likelihood of possible outcomes.

Monte Carlo simulation outputs typically are presented as odds of success. For example, the simulation's result might show a 90% chance of you not outliving your assets. Said another way, the failure rate, in this case, is an estimated 10%.

Addressing the shortcomings of Monte Carlo analysis

However, while the failure rate is an essential tool when evaluating SWRs, as Estrada, author of the October 2016 paper "Refining the Failure Rate," pointed out: "This variable is silent about how long into the retirement period a strategy failed." He continues: "Two strategies that sustained withdrawals for 10 and 25 years of a 30-year retirement period have both failed, but a retiree would be far from indifferent between them."

Estrada's study, which covered 21 countries over the 115-year period from 1900 through 2014, showed that two strategies could have the same failure rate but fail at very different points along the retirement horizon, with one supporting a retiree's withdrawals for a longer time.

Estrada showed that over the 86 30-year retirement periods he considered, a 4% withdrawal strategy from a global 60/40 portfolio would have failed 20 times, or in 23% of the periods. However, those 20 failures looked very different. In some cases, the plan failed with only two years remaining; in others, it failed with 14 years remaining. Those represent two very different outcomes, with very different consequences. Yet, they both count the same way as failures.

To overcome this issue, Estrada introduced "shortfall years" – the average number of years a strategy fails to support withdrawals over all periods in which the strategy failed – as a complement to failure rate. While adding shortfall years is an improvement upon relying solely on failure rate, it uses two distinct metrics, which won't necessarily always agree with each other.

For instance, strategy A could have a lower failure rate than strategy B but a higher shortfall years metric. This suggests that strategy A will fail less often, but when it does fail, it will fail earlier on average into a retirement period than strategy B. Thus, using both failure rate and shortfall years will require retirees to make trade-offs as opposed to a decision based off a single variable that defines a clear choice between the two strategies.

To combat this joint-optimization issue, Estrada introduced a new metric – risk-adjusted success (RAS) – in his April 2017 paper, "From Failure to Success: Replacing the Failure Rate." Given its important implications for evaluating retirement strategies, my colleague, Tim Jost, an institutional

services advisor with my firm, Buckingham Strategic Wealth, supplied the following analysis.

A further improvement using “risk-adjusted success”

RAS, which uses a single metric for choosing an appropriate retirement strategy, is defined as the ratio between the mean-expected value of outcomes and the standard deviation of outcomes, making it similar to the Sharpe ratio. Like the Sharpe ratio, and all else equal, an investor should pick the retirement strategy with the highest RAS. A higher RAS means either a strategy can sustain more years of expected withdrawals or there is less uncertainty about the years of withdrawals it can sustain before it fails.

While RAS provides an improvement over using the joint variables of failure rate and shortfall years, it still poses its own issues, primarily due to its use of volatility as a measure of risk (volatility doesn't distinguish between upside and downside fluctuations). This means that achieving positive outcomes way above a desired result would actually reduce the RAS even though investors should desire these outcomes.

To improve upon this limitation of RAS, Estrada introduced a downside risk-adjusted success (D-RAS) metric in his October 2017 paper, “Replacing the Failure Rate: A Downside Risk Perspective.” Instead of defining risk using the standard deviation of years where retirement is sustained, like the RAS metric does, D-RAS measures risk as the “semideviation” of years sustained.

Putting downside risk in context

Semideviation measures only downside volatility with respect to a benchmark, thus not penalizing fluctuations above a benchmark. As a result, D-RAS measures the dispersion of only failed outcomes (where the years that a retirement strategy sustains withdrawals is shorter than the length of the retirement period). A higher D-RAS value indicates a strategy can sustain more years of expected withdrawals or there is less uncertainty about the years of withdrawals it can sustain when it fails.

Using the same data from his prior work, which you'll recall covers 21 countries over the 115-year period from 1900 through 2014, Estrada compared the optimal allocation suggested by both RAS and D-RAS. He looked at 11 different stock/bond allocations ranging between 100/0 (all stocks) and 0/100 (all bonds) for both metrics. Some highlights from Estrada's findings included:

- On average, D-RAS selects more aggressive allocations than RAS. This is largely due to D-RAS not penalizing aggressive strategies that leave large bequests at the end of a retirement period.
- D-RAS selects strategies with an average stock allocation of 85%; RAS selects strategies with stock allocations that average only 61%.
- Failure rate is more negatively correlated with D-RAS (-0.79) than with RAS (-0.52). This suggests that failure rate and D-RAS rank competing strategies more similarly than failure rate and RAS. Remember, selected strategies have the highest RAS and D-RAS values, so one would expect these ratios to be negatively correlated with failure rate, which investors want to be lower rather than higher.

Given that D-RAS tends to select strategies with relatively low failure rates that tend to leave large bequests, Estrada suggested that D-RAS should be used as a single comprehensive metric in place of RAS and the commonly used failure rate/shortfall years framework. Time will tell if retirement planning software providers incorporate single-variable metrics like RAS or D-RAS into their Monte Carlo simulation software.

However, to address the problems Estrada highlighted in his earlier research, Monte Carlo simulation software can report not only the success/failure rate, but also the median age at which a plan fails and the median amount by which it fails. Investors would be able not only to determine the estimated odds of failure, but also how long into retirement their portfolios could maintain their desired lifestyles.

They could also determine how large an adjustment would be required to prevent failure. This enables them to design effective “plan Bs” – a contingency plan that lists the actions to take if financial assets drop below a predetermined level. Actions might include remaining in or returning to the workforce, reducing current spending, reducing the financial goal, selling a home and/or moving to a location with a lower cost of living.

Estrada made another important contribution with his October 2018 study, “The Bucket Approach for Retirement: A Suboptimal Behavioral Trick?” He examined whether the “bucket approach” is an optimal strategy.

The bucket approach

The bucket approach was developed to deal with the problem of “mental accounting” – the tendency to categorize and evaluate economic outcomes by grouping assets into a number of non-interchangeable “mental accounts.” The approach, which has many variations, essentially calls for parking safely in cash, or short-term, high-quality bonds, a few years of withdrawals, and then investing the rest of the portfolio more aggressively. If the aggressive portion of the portfolio suffers a sharp loss, the retiree can withdraw from the cash reserve, avoiding liquidating assets at a depressed valuation.

By avoiding selling funds from a portfolio that has just suffered a sharp loss, the bucket approach addresses the need for safe near-term liquidity and the goal of long-term growth of wealth. In his own practice, Evensky uses a two-bucket approach that he can effectively implement and monitor. He maintains a cash reserve for clients that is sufficient to handle liquidity needs over a five-year period and invests the remainder of client assets with a longer-term horizon.

Estrada noted that the bucket approach is appealing for several reasons:

- There is no need to understand sequence-of-returns risk or to realize that selling an asset just after its price has fallen sharply is usually not a good strategy.
- It is comforting, enabling a retiree to stop worrying about the possibility of having to liquidate assets at a bad time, as his or her near-term withdrawals are covered.
- It is consistent with the well-known behavioral bias of mental accounting; a retiree is likely to find the separation between the withdrawal account and the investment account appealing.

- It is easy to implement; a retiree only needs to determine his or her annual withdrawals for the next few years, protect those funds by parking them in safe and liquid assets, and invest the rest in more aggressive assets.

While the bucket approach may have psychological benefits (which may justify its use), that doesn't mean it's optimal (i.e., provides the highest likelihood of success). To explore this issue, Estrada examined how bucket strategies performed relative to static strategies (e.g., a typical 60/40 portfolio that is rebalanced annually) both in the United States and globally. His database again covered 21 countries over the 115-year period between 1900 and 2014, as well as 11 static strategies – with fixed allocations to stocks and Treasury bills ranging from 100/0 (all stocks) and 0/100 (no stocks) with nine allocations in between – and three variations of the bucket approach.

The three bucket strategies each consisted of two buckets, one with funds parked in Treasury bills (bucket 1) and the other with funds invested in stocks (bucket 2). All three strategies set aside two years of inflation-adjusted withdrawals and parked them in Treasury bills, investing the rest in stocks. Each of the three bucket rules determine whether to make the annual withdrawal from bucket 1 or bucket 2, and whether to refill bucket 1 when it has less than two years of real withdrawals, depending on the performance of the stock market.

The three bucket rules were designed to avoid withdrawing funds from bucket 2 after stocks have performed badly, but differed in how they assess the performance of stocks and particularly in how they define bad performance. Bucket rule 1 focused on the return of stocks over the previous year. It sought to avoid withdrawing funds from bucket 2 after it has gone down the previous year. It is easy to implement and, according to Estrada, it is among the most popular bucket rules recommended by financial planners. Bucket rule 2 compared the return of stocks over the previous year to the long-term geometric mean return, withdrawing funds from bucket 2 only when the return of stocks over the previous year was higher than the long-term mean return. Bucket rule 3 compared the geometric mean return of stocks over the previous five years to the long-term geometric mean return. Its goal was to avoid withdrawing funds from bucket 2 after a run of low stock returns (relative to long-term performance) expected to mean revert in the near future. Conversely, it aims to withdraw funds from bucket 2 after a strong five-year performance (relative to historical performance) is also expected to mean revert in the near future. In all cases, returns were real returns adjusted for inflation.

Estrada's analysis was based on a \$1,000 portfolio at the beginning of retirement, a 4% initial withdrawal rate with subsequent annual withdrawals (adjusted for inflation) made at the beginning of each year, and a 30-year retirement period. The annual withdrawals were taken proportionally from stocks and Treasury bills with static strategies, and from one of the two buckets (depending on each specific rule) with the bucket strategies. After the withdrawal is made, the portfolio is rebalanced to the target asset allocation in the static strategy cases and no rebalancing takes place in the bucket strategy cases.

The results

Estrada found that a bucket approach *underperforms* static strategies based on all four ways of assessing performance – the failure rate, the number of shortfall years, Estrada's RAS metric, and his D-RAS metric. Although the results obviously varied across the 21 countries he examined, those for

the average country largely confirmed and strengthened those already discussed for the United States. Those non-U.S. findings could be considered “out-of-sample” relative to the baseline U.S. results. Estrada also tested setting aside various years of withdrawals, from one year through five years. He found that, in general, as he extended the number of years, the worse the performance of the bucket strategy became.

Estrada explained the poor performance of the bucket approach as follows: “Most implementations of the bucket approach, and clearly the most popular versions that involve parking in bills a fixed number of annual withdrawals, distribute funds from more aggressive buckets into more conservative buckets, but not the other way around. Put differently, although bucket strategies avoid selling low by withdrawing from bucket 1 after stocks performed badly, they do not take advantage of also buying low as static strategies do through rebalancing.” Static strategies, Estrada observed, sell assets that have become relatively more expensive and buy assets that have become relatively cheaper, thus enhancing the performance of the overall portfolio.

The results led Estrada to conclude: “However plausible, comforting, consisting with mental accounting, and easy to implement the bucket approach may be, simple static strategies, which call for periodic rebalancing and are just as easy to implement, would make retirees better off.”

Estrada’s results lead one to question why the bucket approach is so popular. The most likely answer is that it’s due to lack of knowledge. Estrada’s research makes that excuse less defensible. Another answer is that the bucket approach allows investors to surrender to the bias of mental accounting, allowing them to stop worrying (even though Estrada showed that the approach increases their odds of failure). One of a financial advisor’s important roles is to help investors overcome their personal biases, allowing them to make economically optimal decisions. In this respect, advisors are undermining their core value by allowing investors to succumb to a wealth-destroying bias.

For those still clinging to the bucket approach, Estrada also found that bucket rule 1, the simplest and perhaps most popular rule, was outperformed by bucket strategies that take a longer perspective of stock market performance. Bucket rule 3, which takes the longest approach, outperformed bucket rule 2.

While increasing the number of years of savings you hold in bucket 1 might allow you to sleep better, it also increases the odds of running out of money.

Larry Swedroe is the director of research for The BAM Alliance, a community of more than 140 independent registered investment advisors throughout the country.