The wrong-side-of-maybe fallacy

Michael Kitces & Derek Tharp | Kitces.com | 28 February 2017

EXECUTIVE SUMMARY

Monte Carlo analysis has become a common standard for evaluating the health of a prospective retiree's spending plan in retirement. And while the framework itself may be a rigorous way to look at a wide range of potential outcomes, there's one big problem: it's difficult to know how to interpret the results. What do you do with a scenario like a "95% probability of success"? Who wants to risk being part of the other 5%?

In this guest post, Derek Tharp (our new Research Associate and a Ph.D. candidate in the financial planning program at Kansas State University) delves into behavioral biases and challenges that influence our ability to interpret the results of a Monte Carlo analysis and, in particular, the so-called "wrong-side-of-maybe" fallacy.

The wrong-side-of-maybe fallacy refers to our tendency to interpret a projection as "wrong" if the outcome is inconsistent with the most likely outcome forecasted. It can quickly interfere with the proper interpretation and understanding of probabilistic forecasts because, unfortunately, we have a tendency to try and evaluate the accuracy of a forecast based on a single outcome - for example, if there's a 95% probability of success, and it turns out to be a failure, the forecast must have been "wrong" even though the accuracy of probabilistic forecasts can only be evaluated over a series of multiple outcomes (i.e. to see if a 95% probability of success is really success in 19 out of 20 attempts, on average).

Of course, when working with retirees, the reality is that they only get one shot at retirement. It's unlikely that we'll encounter retirees who are sympathetic to the nuances of proper probabilistic interpretation if their retirement is turning out poorly!

Furthermore, the wrong side of maybe fallacy isn't the only psychological barrier to properly interpreting Monte Carlo analysis. Behavioral research suggests that when clients are asked to express their comfort level with a particular Monte Carlo probability of success, rather than giving a response that answers the question at hand, they may be substituting in an expression of their intensity of feelings regarding failure instead. In other words, a retiree's indication of their comfort with a 90% probability of success may actually be an expression of how much they dislike the idea of the 10% outcome (running out of money in retirement) rather than their actual willingness to tolerate the 10% risk of it occurring.

Fortunately, there are steps we can take to help retirees avoid misinterpretation of Monte Carlo analysis. From re-evaluating terminology - such as framing Monte Carlo analysis in terms of probability of "adjustment and excess" rather than "success and failure" - to

emphasising the importance of ongoing planning, and using "squishy" language to guard against clients being overly critical in their interpretation of advice. Ultimately, if we want to help retirees make better decisions when evaluating Monte Carlo analyses, the cognitive and behavioral biases which may interfere with proper interpretation should not be ignored.

THE "WRONG-SIDE-OF-MAYBE" FALLACY

The wrong-side-of-maybe fallacy refers to our tendency to assess whether a prediction is right or wrong based on which side of "maybe" (i.e., 50%) the prediction is on. In the world of making predictions, it can have a major impact on our ability to assess the accuracy of someone who offers up a forecast.

For example, suppose you turn on the radio and hear there is a 90% chance of rain today. You make it through the day and there wasn't a drop of rain. Was the meteorologist wrong? The answer is "no", or at least "not necessarily" - but most people say "yes". Why? Because the prediction turned out to be on the "wrong-side-of-maybe".

When the forecast for rain is greater than 50%, we tend to say the forecast was "right" if it does rain or "wrong" if it doesn't. Yet, this is the completely wrong way to look at forecasting. In reality, we can never assess the quality of a forecast based on a single observation. We can only evaluate probabilistic forecasts over many observations.

In their book, Superforecasting: The Art and Science of Prediction, Philip Tetlock and Dan Gardner explain that if a skilled meteorologist forecasts a 90% chance of rain, it should rain 90% of the time. Notably, that also means we should expect "no rain" 10% of the time, too. Or, stated another way, if it really rains every time there's a 90% chance of rain, that actually means the meteorologist was wrong – because it shouldn't have rained 10% of the time!

Unfortunately, though, our brains don't seem to like this probabilistic way of thinking. We prefer to boil complex things down to a "yes-or-no" forecast of whether rain is likely or not, see if it happens, and judge accordingly.

Another example was the most recent election. How many headlines followed claiming the pundits/models/pollsters were "wrong" for forecasting that Hillary Clinton would prevail over Donald Trump when she actually lost? That narrative was everywhere - but the reality is, Donald Trump's victory didn't prove the poll predictions wrong at all, as even those predicting a 95% likelihood for Clinton should still see Trump win 5% of the time... and maybe this was just his 5%. Similarly, Nate Silver of FiveThirtyEight - one of the most widely followed election forecasters - gave Hillary Clinton a 71.4% chance of winning the election as of 7 November, and has since been called "wrong" given Trump's victory. But again, Silver's prediction may have been high, low, or spot on. We simply don't know. Given that Silver gave Donald Trump a 28.6% chance of winning, he should have been expected to prevail 2 out of 7 times.

So why does this matter for practitioners? One of the most crucial things we do for our clients is help them forecast. Whether it's forecasting income and expenses, tax changes, the potential for a bull or bear market or, most commonly, a Monte Carlo probability of retirement success, forecasting plays a significant role in creating a financial plan.

And, as we make those forecasts, our clients are constantly evaluating us. Unfortunately for those of us who may wish to be evaluated based on the soundness of our forecasting and advice, there's little reason to believe that we'll be evaluated properly – that is, by the clinet looking at our cumulative forecasts and seeing how often our probabilities were right. Instead, the wrong-side-of-maybe fallacy looms large and our clients are likely to evaluate our forecasts the same way they evaluate a meteorologist – by looking at a single instance that is the opposite of the "most likely" outcome and deeming the forecast wrong, even if it really wasn't.

For instance, what will happen when a client was told they had a 99% chance of success and a 1-in-100 scenario occurs? Will a client appreciate the nuances of interpreting probabilistic forecasts when confronted with the possibility that they may run out of money? Probably not. More likely, the adviser is going to have a very dissatisfied client. And, if their forecast happened to impact many clients, then client satisfaction and retention could be negatively impacted.

POTENTIAL PROBLEMS INTERPRETING MONTE CARLO ANALYSIS

Unfortunately, the psychological challenges of Monte Carlo analyses are not constrained to succumbing to the wrong-side-of-maybe fallacy in evaluating whether the forecast was "right" or not. The increased prevalence of Monte Carlo analysis in retirement planning has certainly been an improvement relative to traditionally oversimplified straight-line projections. Yet clients may have difficulty properly interpreting such forecasts as applied to their own situation and drawing appropriate conclusions.

One concern is that it's possible that clients who are asked to express their comfort level with a particular Monte Carlo probability of success may instead be substituting in an expression of their intensity of feelings regarding the failure. In other words, we don't judge a 90% probability of success by our actual comfort with 90% probability but instead, by how intensely we dislike the potential of the other 10%. Yet saying we intensely dislike the outcome of the other 10% – for example, running out of money, or at least needing to make a material retirement adjustment – isn't the same thing as evaluating our comfort with taking the risk that it might occur.

Another communication concern identified by Tetlock and Gardner is that people both translate odds into ordinary language and ordinary language into odds, but they often don't do so consistently. An example of this referenced by Tetlock and Gardner is a story involving Sherman Kent – a former Yale professor who worked with the CIA and is commonly referred

to as the "father of intelligence analysis". Kent and his colleagues at the CIA would examine information and try to formulate predictions that might help the US government. In 1951, each member of Kent's team signed off on a National Intelligence Estimate sent to State Department officials for review. That report claimed there was a "serious possibility" of attack on Yugoslavia that year. When casually asked what exactly "serious possibility" meant, Kent replied he personally thought about 65%, but was troubled to hear that the State Department official thought the report meant much lower odds. He went back to his team and asked each member how they interpreted the statement they had signed off on. Their answers ranged from odds of 20-to-80 to 80-to-20. If a team of highly trained analysts making serious assessments regarding the risk of war can have this much misunderstanding in probabilistic communication about what a "serious possibility" means, it's highly likely that you and your clients aren't always on the same page, either.

In the context of financial planning, this means that using language like "your retirement plan has little risk of running out of money" could still be interpreted quite differently - some might assume "little risk" means just a 1-in-100 chance, while others might assume it means "just" a 20% chance of failure (which is a highly material difference).

One possible solution is to always express predictions as probabilistic forecasts - but problems can run the other direction as well. Most financial planning software does distil a Monte Carlo analysis down into a single number, but converting numbers to words won't necessarily result in less confusion. Not only may we apply different subjective meanings to different levels of confidence, but a singular percentage still fails to answer many important questions. What's the magnitude of failure? Does a plan failure mean merely cutting back on a few luxuries, or is a client left completely broke with no backup plan? How sensitive is a plan to various risks and trade-offs? Is the plan with the lowest risk of failure necessarily the best plan for a client if it has the worst magnitudes of failure when they do (rarely) occur? Unfortunately, oftentimes we don't even know the answers to these questions which is not necessarily our fault or an inherent flaw of probabilistic projections. In part, it's just a reflection of a current generation of Monte Carlo software excessively focused on reporting a single probability of success.

We also know that people display irrational discrepancies in interpreting probabilistic information based on how it is presented to them. For instance, one study found that framing risk information using mortality curves resulted in lower interest in preventative surgery relative to using survival curves, even though both were based on the same probabilistic forecasts. For instance, researchers have found that people make different medical decisions when told they have a 30% chance of survival (positive framing) versus a 70% chance of death (negative framing).

It's also important to recognise that it may not just be clients who struggle with probabilistic thinking - even the most technically competent practitioner may struggle to overcome the wrong-side-of-maybe fallacy and other probabilistic biases. While there is still a lot we don't know about cognitive bias mitigation, we do know that knowledge alone is not always

enough to overcome bias. In fact, in some instances, knowledge can make biases even worse. For instance, when it comes to framing medical outcomes, physicians can be impacted just as much as their patients, and one study found that physicians are more supportive of treatments when they are framed in terms of gains or relative reductions in risk rather than losses or as absolute reductions in risk. And the effect may well be present for practitioners too. Are you more comfortable if a client has a 90% probability of success, compared to a 1-in-10 chance of running out of money?

ADDRESSING PROBABILISTIC BIASES

You don't have to spend too much time reviewing the literature on probabilistic decision-making within the medical setting to realise that addressing probabilistic biases is complex and there's still a lot we don't know. But there may be some steps practitioners can take to try and help their clients (and themselves) make better decisions.

Reevaluate terminology

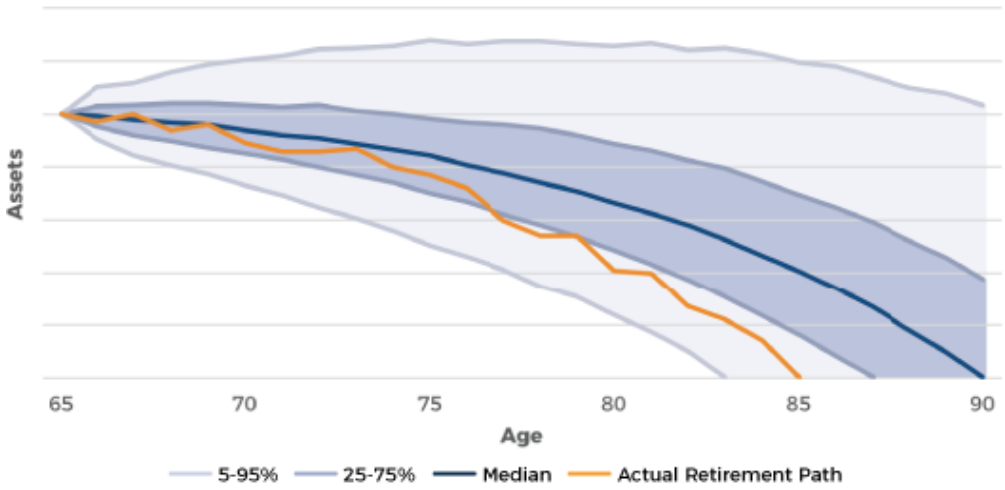
How we frame Monte Carlo results matters. As noted earlier, using actual percentages rather than words to describe results (like "very likely to succeed" or "little risk of failure"), can help. Reframing the description of outcomes can help, too. We most commonly talk about probabilities of "success" or "failure" - but saying "probability of excess" and "probability of adjustment" as alternatives can provide more practical insight into the actual consequences that result from various outcomes.

After all, while it's true that Monte Carlo analysis captures the percentage of plans that "fail" (run out of money), a Monte Carlo analysis also assumes clients will charge forward blindly and fail to make any adjustments along the way. In many cases, relatively minor adjustments can "save" a plan, and "probability of adjustment" may more accurately convey the consequences of encountering an unfortunate combination of circumstances.

Similarly, "probability of excess" better captures the reality that "successful" iterations of a Monte Carlo analysis often aren't just "successful" in not running out of money, but leave significant "excess" assets behind. This may be favorable as a means to hedge longevity risk or leave assets to heirs, but framing the probabilistic decision as one between "excess and adjustment" rather than "success and failure" arguably gives a client a clearer understanding of the alternatives ahead of them.

For instance, consider a scenario in which the initial Monte Carlo analysis for a retiree showed a 50th percentile outcome of having the portfolio last all the way to age 90. Further, suppose that an individual encounters an actual retirement path that results in running out of money at age 85, as indicated in Figure 1.

Figure 1: Actual retirement path vs initial Monte Carlo projection

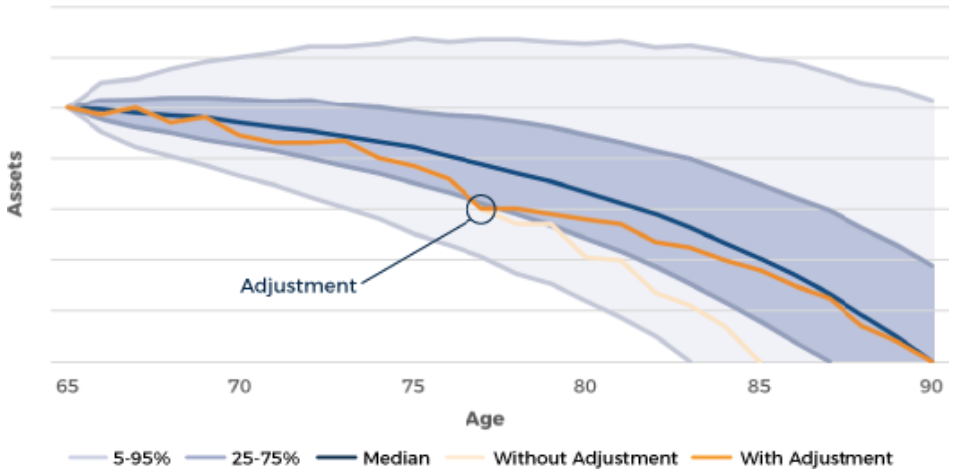


Sources: Michael Kitces, www.kitces.com

It's worth noting that just because the retiree ran out of money at age 85 does not mean the initial projection was "wrong". The original analysis did project a roughly 15% chance of depletion by age 85. A single instance of running out of money five years short of 90 doesn't mean it was wrong to project that the portfolio would last to age 90.

Of course, with an ongoing monitoring process, it would have been clear along the way that the portfolio was already veering off in the less-likely-but-possible direction of depletion by age 77 (Figure 2), when the portfolio that started out with a 50% chance of making it to age 90 dipping down to the 25th percentile. Rather than blindly staying the course, the retiree would have an opportunity to make an adjustment at that point, getting back on track to sustain the asset base longer.

Figure 2: Avoiding scenario failure with an adjustment



Sources: Michael Kitces, www.kitces.com

At a minimum, framing the outcomes in less extreme terms (adjustment vs excess rather than failure vs success) may reduce the intensity of retirement-related fear, and make it easier for clients to evaluate their situation more objectively, especially as highly adverse "running out of money early" scenarios rarely happen terribly suddenly. With ongoing monitoring and continuously updated Monte Carlo analyses, a portfolio that is veering off track will be evidenced by a probability of success that steadily declines over time (prompting spending to adjust before "failure" is assured).

Emphasise the importance of ongoing planning

Clients should understand planning is not a one-time occurrence. One advantage of focusing on "probability of adjustment" is that it implies that there's an ongoing need to monitor and evaluate potential adjustments. Additionally, monitoring the trend in the probability of adjustment (or success) over time may provide more valuable and motivating feedback for clients – particularly those who have significant time before retirement.

Suppose a couple in their 40s is way behind pace and their plan shows a 30% probability of success. Further, suppose that over the past year they paid down some debt and increased their savings. What's likely to be more motivating? Telling them that despite their efforts over the past year, they still only have a 35% probability of success? Or, instead, telling them that their probability of success increased 5%? Ultimately, some combination of both messages is probably needed as they need to know they are behind pace but moving in the right direction. But the point is, the trend may actually be more important than a client's position at any given time, and the trend can only be analysed through ongoing planning.

In addition, ongoing monitoring can help to spot potential problems before they become serious. For instance, if a plan has a 10% probability of adjustment and the next year it rises to 15%, the client has at least started down that 10% path – such that the risk is rising – but is also still early enough that a moderate adjustment can get them back on track.

Present information in more than one way

In a review of published literature on describing treatment effects to patients, researchers found that methods for describing probabilities of side effects and treatments need careful consideration. Their main takeaway for physicians was that treatment effects need to be communicated to patients in multiple ways. They found that presenting information in more than one way increased the likelihood that patients would understand the information being presented and make informed decisions.

In the context of retirement planning, this might mean presenting probabilities of adjustment and excess, but also a visual look at the range of retirement outcomes and potential paths that might occur. It could also include scenario planning to look at how making upfront or mid-course adjustments can improve the situation or a look at how sensitive the outcomes are to adjusting the planning assumptions in the first place.

Consider using squishy language

Tetlock and Gardner highlight the fact that many people use "squishy language" to avoid being held accountable for their predictions. We see this often from people who make bold assertions in the media. For instance, in January 2012, economic historian Niall Ferguson claimed that "The Greek default may be a matter of days away." Yet, Tetlock and Gardner note that Ferguson's statement provides many ways to avoid criticism and being proven "wrong". First, "default" has both colloquial and technical definitions. Second, Ferguson's use of the word "may" also necessarily implies that it may not. While this isn't a problem in itself, without any actual probability assigned, Ferguson could retroactively claim that he either meant there was a high or low likelihood of occurrence, depending on the outcome. Finally, "matter of days" provides no definitive and measurable guidelines. He could mean three days, 30 days, or 300 days. Without a more precise statement, we have no way of knowing.

The point isn't to pick on Ferguson's squishy language (which may equally be the fault of the interviewer for not demanding a more precise answer), but to acknowledge that forecasters can and often do get away with fuzzy predictions using squishy language, and minimise their responsibility for being "wrong" about their prediction.

In the context of Monte Carlo analysis and retirement planning, explaining risks as a probability of "adjustment" instead of "failure" is one form of squishy language, as it dampens the severity (adjustment is ostensibly less severe than failure), although it doesn't necessarily explain what the adjustment might be (and whether it might still be quite severe).

Ongoing updating of the plan, where the probabilities are only likely to move slightly from year to year and therefore can merit an intervention before a seriously bad outcome occurs, is another way to minimise the risk that the forecast is 'proven' wrong.

Obviously, these techniques also present some serious ethical considerations. While squishy language may help guard against clients' fallacious use of wrong-side-of-maybe reasoning, there's a clear risk of breaching of fiduciary duty by misleading clients if the practitioner is trying to avoid accountability altogether, particularly when that ambiguity interferes with a client's ability to make an informed financial decision. While helping clients avoid wrong-side-of-maybe reasoning is certainly a legitimate goal, it shouldn't be used to pursue self-serving ends.

Ultimately, the role of the practitioner in "de-biasing" a client's various behavioral biases is both important and under-researched when it comes to the best communication techniques to do so effectively and ethically. Nonetheless, the first step is to at least recognise and understand the kinds of behavioral biases that may be at work, to spot opportunities to clarify that the client really understands the information and decision being presented to them - and, perhaps, to reduce the risk that the practitioner is inappropriately blamed for the client's biased misunderstanding of that information or decision.

So what do you think? Have you ever had a client fall victim to the wrong-side-of-maybe fallacy? Do you think clients struggle interpreting Monte Carlo analysis? What can we do to communicate a Monte Carlo analysis more clearly? Please share your thoughts in the comments below!

Michael Kitces is a Partner and the Director of Research of Pinnacle Advisory Group, a US-based private wealth management firm that works with over 700 families and manages close to US$1 billion in assets for clients in the US and around the world. Michael is a member of PortfolioConstruction Forum's core faculty of leading investment professionals.

Derek Tharp is Research Associate at Kitces.com and is currently finishing up his Ph.D. in the Personal Financial Planning program at Kansas State University, and assists clients through his RIA Conscious Capital. Derek is a Certified Financial Planner.

The above article is reproduced with permission from Michael's blog "The Nerd's Eye View".
