GMO White Paper

Feb 2017

Target Date Decisions, Decisions… Getting the Biggest Bang for the Buck


Peter Chiappinelli and Ram Thirukkonda

Executive Summary

There are a host of target date and plan design initiatives that the DC plan sponsor community is considering today that might, in theory, improve the retirement outcomes for their employees. This paper tests many of them by measuring the incremental impact each would have had over a specific 40-year time frame (the typical span of a working career), from 1975 to 2015. We systematically tested one variable at a time – passive vs. active, glidepath "riskiness," pre-determined vs. dynamic glidepaths, and deferral rates – and measured the incremental impact over a baseline. We were able to isolate the variables that had the "biggest bang for the buck." Importantly, we showed that incorporating a valuation-sensitive dynamic allocation piece into the glidepath, combined with efforts to boost deferral rates, would have delivered meaningful improvements in retirement wealth outcomes, significantly more than activities related to open-architecture ("best-in-breed"), active management, or different risk levels of glidepath design.

Introduction

The DC plan sponsor community today is confronting a cacophony of questions related to target date funds and overall plan design. These include, but are not limited to, the following: Active vs. passive? Custom vs. off-the-shelf? Proprietary vs. open-architecture ("best-in-breed")? Traditional vs. alternative asset classes? Glidepath risk, as in aggressive vs. conservative? Glidepath philosophy, as in pre-determined vs. dynamic? To vs. through? Also included are match design questions: Tiered vs. flat? To auto-escalate or not? Yikes. The list goes on and on. While these questions are all vying for attention, they all cannot possibly be equal in importance. We put together an experiment designed to cut through the noise and help the plan sponsor to focus on the key question: "What is the best way to help my employees?"

The experiment[1]

We went back in time and pretended that target date funds existed in 1975. We then created a baseline experience and measured the retirement wealth of a "model" employee over the next 40 years. In good scientific fashion, we then went back and changed one variable at a time, measuring the incremental impact on that wealth, determining which of those decisions provided the biggest bang for the buck.

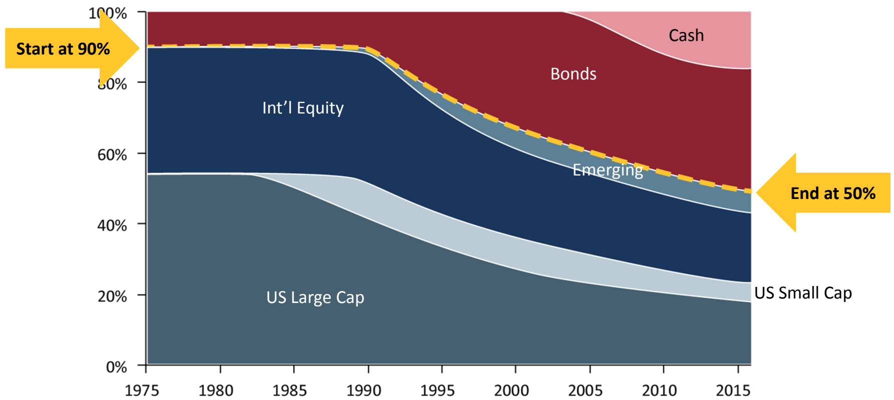
As noted, we created a baseline. There were some variables that we held constant, such as the time frame and some demographic and plan assumptions (e.g., salary, salary growth, match formula), but we changed other variables:

- **Passive.** The baseline had six different asset classes, all managed passively: US large cap, US small cap, international equity, emerging equity, US investment grade bonds, and cash. Each set of returns was represented by the common industry benchmark return with a modest fee deducted. In the experiment phase, we then replaced passive returns with active returns, using various methods of choosing underlying active managers.
- **Moderate glidepath "risk."** The baseline started with an equity allocation of 90% and then glided down to 50% over the course of the 40 years. In the experiment phase, we replaced the moderate glidepath with a more conservative one and then a more aggressive one.
- **Pre-determined glidepath.** The baseline started with a pre-determined glidepath (that is, allocations were determined four decades in advance, with no ability to adapt to new information). In the experiment, we replaced the pre-determined glidepath with one that used a simple valuation metric that allowed the allocation to change dynamically.
- **6% deferral rate.** The baseline case started with a 6% deferral rate, which has been the industry median for many years now. We then increased the deferral rate assumption. (To be clear, we knew, of course, that raising the deferral rate would have a positive incremental impact. Our intent was to measure the magnitude of the change, not the direction.)

[1] See Endnote 1 for acknowledgements.

Below is the detailed glidepath from 1975 to 2015. The final retirement wealth of the baseline experience, $629,000 (nominal), is also displayed.

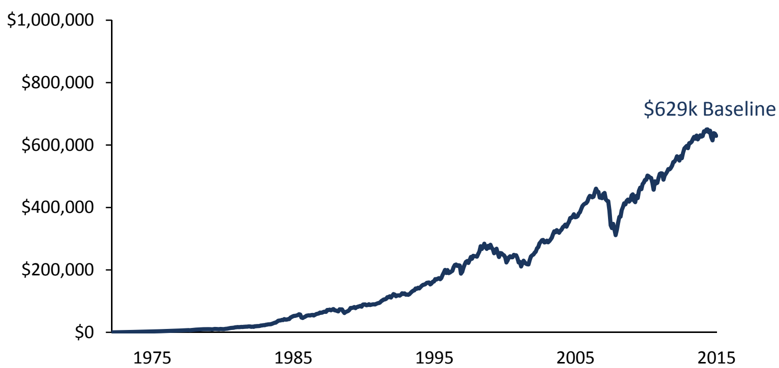
Exhibit 1: Passive + Moderate Glidepath + Pre-Determined + 6% Deferral



US Large	=	S&P 500	Int'l	=	MSCI EAFE	Bonds	=	Barclays Agg
US Small	=	Russell 2000	Emerging	=	MSCI EME	Cash	=	T-Bills

Source: GMO

Exhibit 2: Baseline Retirement Wealth



Source: GMO, Datastream

Now that we had our baseline wealth measurement, we started changing variables. First up was the swapping out of passive returns for active management.

Passive vs. active

The passive vs. active debate – Can portfolio managers add value through active security selection? – is a perennial one in the asset management industry, but it has particularly important meaning in today's target date environment where fee sensitivity is high. While many in the industry might concede that US large cap space is "efficient," target date funds comprise a variety of other less efficient asset classes that may provide opportunities for active managers to add value. In addition, our colleagues at GMO have written[2] about the apparent "cycles" of passive outperforming active (i.e., sometimes it does,

[2] Neil Constable and Matt Kadnar, "Is Skill Dead?" This white paper, published on February 11, 2015, is available with registration at www.gmo.com.

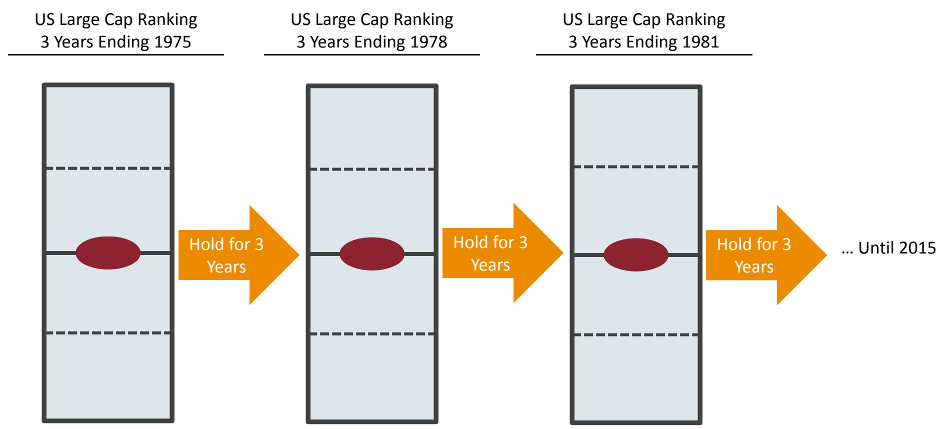
sometimes it doesn't), and the 40-year time frame would capture a few of those cycles. We wanted to explore what value active management would have added over this time frame in a target date setting.

First, we needed a set of active mutual fund returns that dated back to the 1970s. The Center for Research in Security Prices, or CRSP, is a highly regarded database of individual security and mutual fund data, maintained by University of Chicago. It has robust data in that historical returns never drop out, even when a fund closes, solving for the survivorship bias inherent in many other databases.

Second, we needed a method of choosing active managers. The following methods are by no means exhaustive, as most consultants and plan sponsors would use a mix of variables, quantitative and qualitative, to choose managers. But we narrowed the field to the following straightforward methods:

- **Median managers, returns.** For each asset class, we created a three-year look-back on raw returns and then chose those managers that ranked at the exact median of their respective asset class. We "held" those managers, experienced their returns for the ensuing three years, and then rotated into the next batch of managers who were also "median" on a three-year look-back. The diagram below illustrates this process.

Exhibit 3: Median Managers Defined

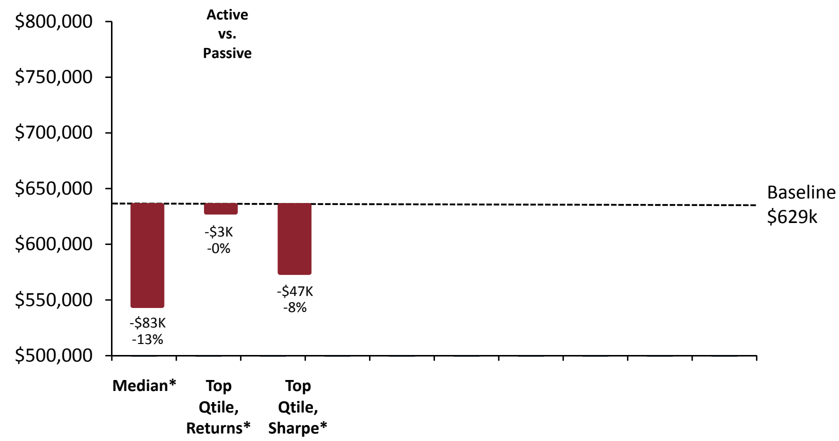


Note: We repeated the ranking exercise for each asset class (e.g., US small cap, international).

- **Top quartile, returns.** The next method of choosing managers was similar, but instead of choosing median managers we chose those in the top quartile. (We believe this was a more intuitive approach, as historically good performance is often an important criterion for manager selection.)
- **Top quartile, Sharpe Ratio.** Manager selection can also be done based upon return and volatility. So, we ran an exercise similar to the one above but this time chose managers that ranked in the top quartile in terms of Sharpe Ratio.

The purpose of this exercise was to measure the incremental impact of these various decisions. The results are displayed in the scoreboard on the following page.

Exhibit 4: Incremental Impact on Retirement Wealth



* Calculated (or Derived) based on data from CRSP SURVIVOR-BIAS-FREE US MUTUAL FUND DATABASE ©201606 Center for Research in Security Prices (CRSP®), The University of Chicago Booth School of Business (2016 being the year the database was published)

It is clear from the scoreboard that active management across this time frame, using admittedly simple backward-looking metrics, did not add value.[3] The take-away: Plan sponsors should not obsess about open-architecture or "best-in-breed" active frameworks.

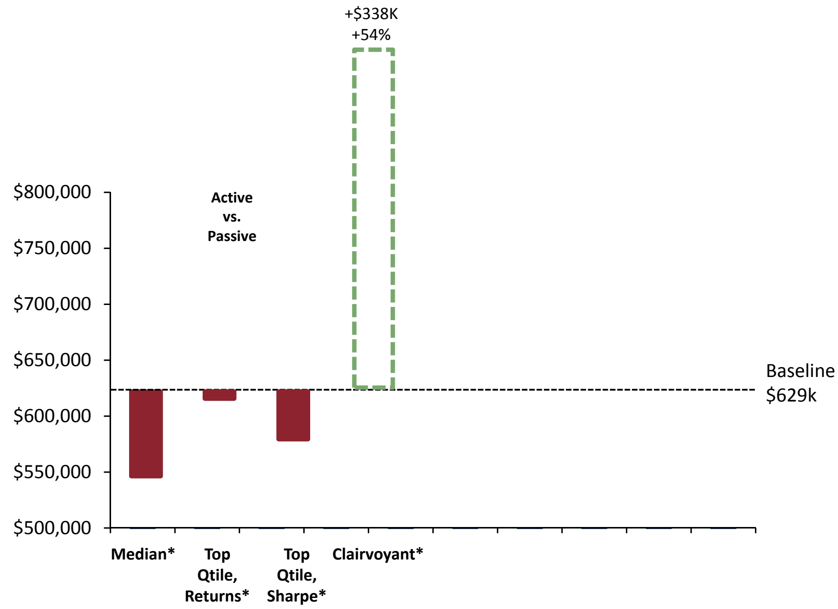
Before we leave the active vs. passive section of the paper, we want to add one more "method" of choosing managers, as all of the above experiments used backward-looking metrics. What if we were able to know, in advance, who the top quartile managers were going to be? What if every three years we rotated into a new set of managers for each asset class, knowing in advance that they would be top quartile three years hence? We call this the Clairvoyant Portfolio.

Before we show you the Clairvoyant Portfolio results, we want to add some perspective to this experiment. Being able to predict which managers would be top quartile is not a trivial undertaking. Mathematically, there is a one in four chance of choosing a top quartile manager. However, doing this every 3 years, for roughly 13 cycles (13X3 = roughly 40 years) across 5 asset classes is really challenging. How challenging? The odds of doing this randomly are (1/4) to the 65th power. To put it another way, the odds of winning the Powerball Lottery jackpot are roughly 1 in 290,000,000; the odds of getting the results of the Clairvoyant Portfolio would be similar to winning the Powerball …five times in a row! Essentially impossible. But the exercise gave us a theoretical limit as to what picking the best active managers would have delivered in terms of incremental impact. With that as context, you will see below that the Clairvoyant Portfolio would have added $338,000 to the baseline retirement wealth, a 54% increase. We will plot it for now in the scoreboard below, but will not talk about it again until later as it is clearly an unrealistic yardstick given the astronomically poor odds against ever achieving this result.[4]

[3] Please note that this was not a broad assessment of active management, per se. There were numerous time periods when active management in small cap, international, and emerging equity added value over the respective passive benchmark. However, given the high weights of most glidepaths to US large cap and investment grade bonds, two admittedly "efficient" asset classes, the aggregated active portfolios did not add much value. Also, we used a three-year rotation cycle; further research on longer holding periods for managers might have different results.

[4] It is not surprising that we found the odds of achieving the results of the Clairvoyant Portfolio to be extremely poor given our assumption of the randomness of the exercise. To the extent that a plan sponsor or a consultant truly has skill in predicting which managers will be successful, it becomes more likely that some portion of the Clairvoyant Portfolio's benefits could be realized. For our experiment, however, we had no simple way of modeling this skill.

Exhibit 5: Incremental Impact on Retirement Wealth

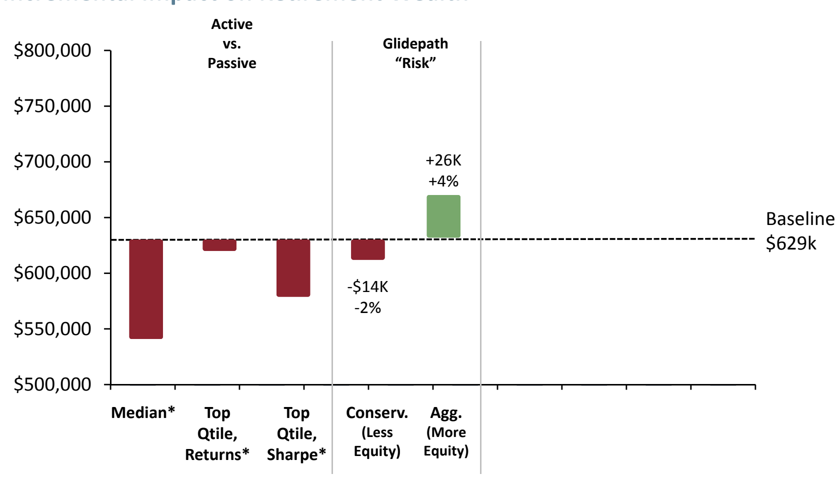


* Calculated (or Derived) based on data from CRSP SURVIVOR-BIAS-FREE US MUTUAL FUND DATABASE ©201606 Center for Research in Security Prices (CRSP®), The University of Chicago Booth School of Business (2016 being the year the database was published)

Glidepath "risk"

Another variable we tested was the "riskiness" of the glidepath as measured by the amount of equity held. We called one variant Conservative, characterized by the following: Instead of starting with 90% equity (the baseline case), it started with 85%; it held this amount of equity for a shorter period of time; and it ended its path at a lower amount of equity at retirement, 40%. The Aggressive variant did just the opposite: It started with 95% equity; it held onto equity longer; and it ended at a higher amount of equity at retirement, 60%. The results are shown in the scoreboard below.

Exhibit 6: Incremental Impact on Retirement Wealth



* Calculated (or Derived) based on data from CRSP SURVIVOR-BIAS-FREE US MUTUAL FUND DATABASE ©201606 Center for Research in Security Prices (CRSP®), The University of Chicago Booth School of Business (2016 being the year the database was published)

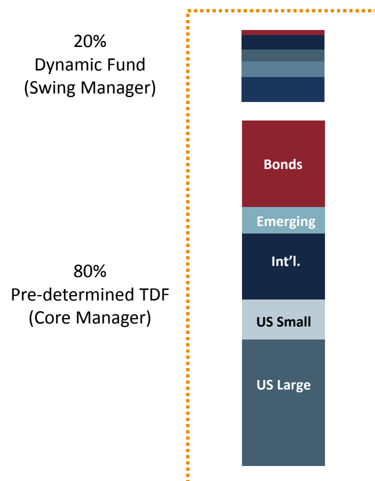
As you can see, the Conservative glidepath barely impacted the results. And the same could be said for the Aggressive glidepath. The take-away: Plan sponsors should not obsess about the slope of the glidepath.

Pre-determined vs. dynamic glidepath

Virtually all the major glidepath designs in the industry were originally designed as "pre-determined." That is, their asset allocations would change through time, but they would change in a pre-determined manner, regardless of market conditions at any given moment. This strikes us at GMO as a bit silly, given that destructive asset class "bubbles" have formed through time, and are, we believe, detectable and largely avoidable. There are signs. Yet a pre-determined glidepath tragically ignores these signs. So we designed a test that incorporated two simple metrics the glidepath might have used to alter the pre-determined path dynamically – the Shiller Cyclically Adjusted Price to Earnings ratio, or CAPE, for equities, and current real yields for bonds.[5]

We tested two ways of implementing this dynamism. The first is what we call "Dynamic Swing." In this structure, we take the original baseline glidepath, but instead of allocating 100% of the capital (employee and employer cash contributions), we allocated only 80%. The remaining 20% went to a Dynamic Fund, which was "paired" with the baseline as shown in the diagram below.

Exhibit 7: Dynamic Swing: Building a Paired Portfolio

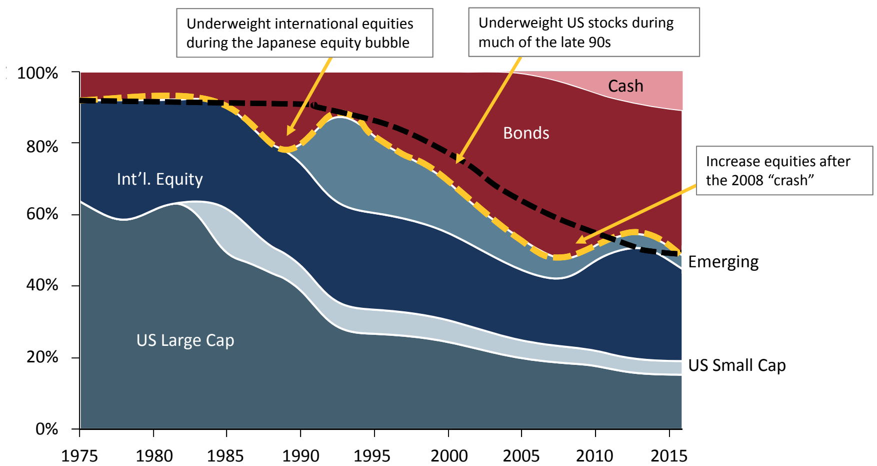


In the Dynamic Fund, the simple valuation tools were used to change the asset allocation through time. Each month, the Dynamic Fund simply took the backward-looking Shiller CAPE (no judgment or knowledge of the future was required) for the different equity asset classes and real bond yields for bonds, translated them into expected returns for each of the various asset classes, and then conducted a straightforward mean-variance optimization to create the asset mix. In other words, it used market valuation as its asset allocation guide, not time. Because we constrained the amount of capital to the Dynamic Fund to 20% (remember, 80% of all assets would still be following the original baseline glidepath), we let the Swing Manager be unconstrained. (We set up the first experiment this way to demonstrate how easily a plan sponsor could have implemented this concept. Specifically, a plan sponsor need not have created a customized, open-architecture structure; he or she could have implemented this structure simply by unitizing two distinct funds.)

[5] Please note that we used the Shiller CAPE because over time it has proven to be highly correlated with future equity returns. The same is true for real bond yields. For more detail on the detectability of dangerously overvalued markets please see Endnote 4.

The diagram below shows how the allocation path would have changed during the 1975 to 2015 time frame, with the Dynamic Fund reacting to changing valuation signals. We have plotted the original baseline glidepath as a dotted black line for purposes of comparison. We have highlighted, with the yellow arrows, some interesting "decisions" that the valuation tools would have made. First, in the mid to late 1980s, the glidepath would have tilted away from international equity, driven by the Japanese equity bubble that was forming. Second, we highlighted the long-term underweight to equities, specifically US equities, during the TMT bubble formation of the late 1990s and the debt bubble of 2007-08. Finally, we highlighted the post 2008 time frame, when the glidepath would have added to equities.

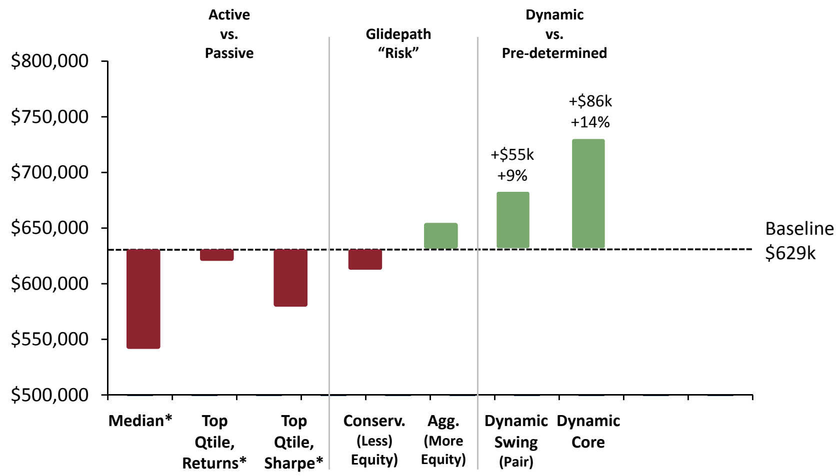
Exhibit 8: Dynamic Swing: How Did the Glidepath Change?
Allocations in line with historical valuations



Source: GMO

The second variation of the dynamic glidepath, which we call "Dynamic Core," would have been a bit more disruptive because it would have entailed hiring an entirely new manager to manage 100% of the assets in a dynamic fashion. We constrained the optimization exercise so that the valuations-based allocation could have deviated from the original baseline glidepath no more than 0.5 to 1.5 times the original weight. As an example, if, let's say, a point along the baseline glidepath had had a 6% allocation to emerging equities, the "new" allocation would have been range-bound between 3% and 9%. But even with these constraints, the Dynamic Core experiment would have had even more freedom to shift assets than the Dynamic Swing. The results are shown on the following page.

Exhibit 9: Incremental Impact on Retirement Wealth



* Calculated (or Derived) based on data from CRSP SURVIVOR-BIAS-FREE US MUTUAL FUND DATABASE ©201606 Center for Research in Security Prices (CRSP®), The University of Chicago Booth School of Business (2016 being the year the database was published)

As you can see, the Dynamic glidepaths had meaningful impacts. With the Dynamic Swing, a $55,000 incremental impact represents close to a 9% increase in retirement wealth, and the Dynamic Core adds $86,000 to the baseline wealth. These are interesting results, especially given how admittedly crude and simple the valuation tools were. The take-away: Plan sponsors should explore dynamic allocation options because they appear to show promise.

Deferral rates

The final experiment dealt with deferral rates. As an aside, we knew a priori that increasing deferral rates would, by definition, increase retirement wealth; the objective here, however, was to measure the magnitude of this change relative to other variables. For this analysis we simply assumed that the plan sponsor – through cajoling, education efforts, and an assortment of behavioral "nudging"[6] tricks – was able to boost the median deferral rate of the employee base from 6% to 7%. As it turned out, a 1% boost provided an 11% increase in retirement wealth, or about $70,000; this was in line with the increase that our simple dynamic allocation approach would have delivered. The take-away: Boosting deferral rates showed meaningful improvement, certainly relative to other variables.

Encouraging news

Based on our results thus far, the two most promising levers appear to be adding a dynamic component to the glidepath and boosting deferral rates. But, it actually gets better. These two approaches need not be mutually exclusive. A plan sponsor could do both, combining a dynamic approach with efforts to boost deferral rates. The scoreboard below shows that such a combination adds substantially to retirement wealth: a +20% to a +30% improvement over the baseline. Just for reference, we are replotting the impact of the Clairvoyant Portfolio. It's striking that two relatively simple changes helped to achieve a meaningful portion of the essentially impossible odds of the Clairvoyant Portfolio.

[6] Please see Endnote 5 for a detailed discussion on behavioral "nudging."

Exhibit 10: Incremental Impact on Retirement Wealth

The biggest bang

Many voices are shouting for the attention of the DC plan sponsor community today, offering ways to enhance target date programs. What we've demonstrated here, though, is the importance of cutting through the noise and focusing one's time, effort, and resources on the initiatives that appear to show the most promise. Adding a dynamic component to the glidepath combined with an earnest effort to boost deferral rates appears to have the potential to deliver the biggest bang for the buck.

Endnotes

1) The authors would like to acknowledge the prior support of Mr. Chiappinelli's research by his former employer, Putnam Investments, where he carried out related research in 2005. While Mr. Chiappinelli's and Mr. Thirukkonda's current analysis and conclusions differ in substantive respects from Mr. Chiappinelli's prior work, and from continuing work by others at Putnam Investments, one can find the current research in this field published by Putnam Investments in W. Van Harlow, "Defined Contribution Plans: Missing the forest for the trees?" http://www.empower-institute.org/media/PDF/EmpowerInstitute_Research_Forest.pdf, accessed January 18, 2017.

2) Below are some important observations regarding our methodology:

    a. **Target Date Framework.** The use of the Target Date Framework was chosen to represent the current state of the defined contribution landscape.

    b. **Forty-Year Time Frame.** The time frame we chose, 1975 to 2015, was important because we needed a 40-year span of time to match the typical Target Date "accumulation" period.

    c. **Database Survivorship Bias.** This research was conducted using data from the CRSP SURVIVOR-BIAS-FREE US MUTUAL FUND DATABASE ©201606 Center for Research in Security Prices (CRSP®). This database is careful not to lose the performance record of funds, thus solving for survivorship bias. Survivorship bias in any research study of historical returns is an important consideration, as it creates an upward drift in reported returns. Removing this drift provides a more accurate assessment of past performance.

    d. **Clusters.** In order to construct the "first quartile" or "median" manager returns, we used a cluster method. That is, rather than using a single manager to represent any quartile, we created a cluster of four to five managers, which lent stability to the return streams and removed the idiosyncratic returns of any one manager. This cluster methodology was also used to create the Clairvoyant Portfolio.

    e. **Sharpe.** The usage of Sharpe Ratios, to supplement raw returns, was done using a three-year look-back on the volatility of returns of the cluster of managers. This method tried to emulate a model of choosing managers that plan sponsors or consultants routinely use.

    f. **Dynamic.** Introducing the dynamic part of the research was important, as most Target Date funds were originally designed as pre-determined glidepaths. We wanted to isolate the effect of introducing dynamic components. Moreover, the two different methods of introducing dynamic portfolios represent two very distinct ways a plan sponsor might actually implement such a program: first, through a "pairing" concept, which could be done with a modest disruption to the plan; and second, through hiring an entirely new manager. We wanted to show the results of both, as it was an important part of the study.

g. **Global Equity Asset Classes.** We chose four distinct types of major equity sub-asset classes for the study, as this is the way many plan sponsors break out categories: US large cap, US small cap, international (developed countries outside of the US), and emerging equities.

3) The assumptions used for this experiment are provided below.

   a. The actual monthly returns of active managers and relevant indexes were used for the stated time frame of 1975 to 2015.

   b. US large cap returns were represented by the S&P 500 index; US small cap returns were represented by the Russell 2000 index; international equity returns were represented by the MSCI EAFE index; emerging equity returns were represented by the MSCI Emerging Equity index; US investment grade bond returns were represented by the Barclays US Aggregate Bond index; and cash returns were represented by US T-Bills. We deducted a fee of 16 basis points annually from the total passive portfolio.

   c. The model employee and plan design assumptions that follow were based upon standard industry data:

      i. Employee began working at age 25 and retired at age 65. Employee participated and contributed throughout the entire 40-year career;

      ii. Salary began in 1975 at $10,000 and grew at a real rate of 1.5% annually;

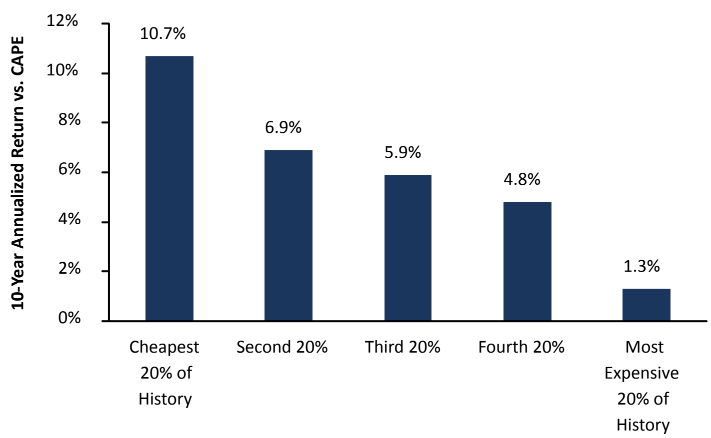
      iii. Deferral rate was 6% in the baseline;

      iv. Company match was 50 cents on the dollar, up to 6%;

      v. For additional information on defined contribution statistics see https://pressroom.vanguard.com/nonindexed/HAS2016_Final.pdf

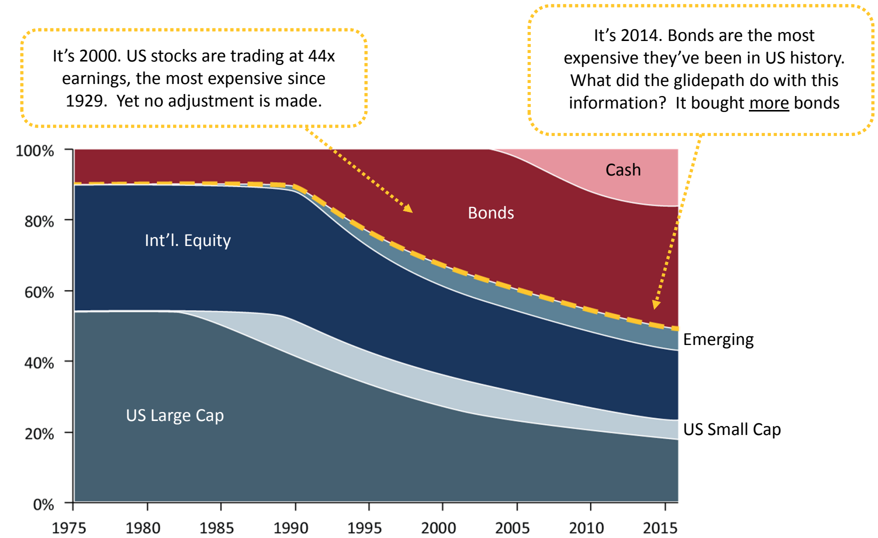
4) The Shiller P/E, or CAPE, was used to create expected return assumptions in the mean-variance optimization exercise because it has shown to be correlated with future long-term returns. It also provides, at times, examples of over- and underpriced markets that pre-determined glidepaths would typically ignore. The same holds true for bond yields.

Starting Valuations Tell a Lot About the Future



CAPE: Cyclically adjusted price to earnings

Pre-determined Glidepaths: Potential for Irresponsible Decisions



Source: GMO

5) By "behavioral nudging," we mean that behavioral finance studies have shown that employers can signal, or "nudge," employees to make better decisions subconsciously through, first, the choice of default deferral rate, and, second, through the match threshold. An employer may, for a variety of reasons, have chosen a default deferral rate of 2% and a match formula of 50 cents on the first 6%. But the employee often interprets these numbers as the "right" ones, even though this was not necessarily the intent. Studies have shown, however, that an employer, at no additional cost, can "nudge" its employees to defer more money simply by shifting the two variables. In the stylized example below, the red dots represent where employees tend to "cluster" – at the default deferral rate and at the match threshold. By shifting both upwards, employees will tend to cluster around these new, higher deferral levels, regardless of what the actual match amount is.

Endnote:Trying to Boost Deferral Rates

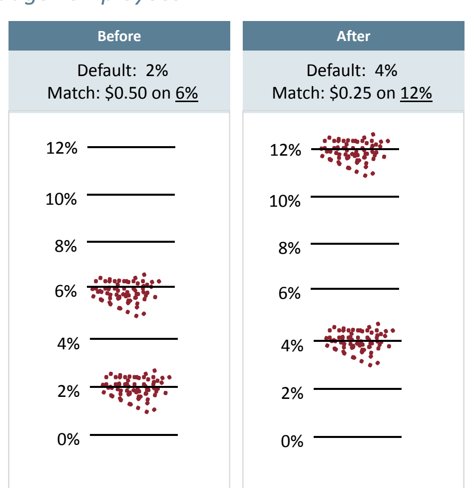
A stylized example of methods to "nudge" employees

Defaults

- Set default rate higher
- Default auto-escalation
- Let inertia do its job

Signaling/Anchoring

- Employees subconsciously look for signals
- Employees tend to "cluster" around match level



For further reading on how match thresholds can influence behavior, see work by James Choi, Harvard University: https://www.shrm.org/ResourcesAndTools/hr-topics/benefits/Documents/04-08LaibsonFinal.pdf

6) We fully acknowledge some of the limitations of this experiment, the most significant being that the time frame we chose used the actual capital market experience but it represented only one "run" of history. Through future research papers, we hope to supplement the work we've done by using stochastic modeling techniques (i.e., "Monte Carlo" simulations) to further test our conclusions.

Peter Chiappinelli. Mr. Chiappinelli is a member of GMO's Asset Allocation team. Prior to joining GMO in 2010, he was an institutional portfolio manager on the asset allocation team at Pyramis Global Advisors, a subsidiary of Fidelity Investments. Previously, he was the director of institutional investment strategy and research at Putnam Investments. Mr. Chiappinelli earned his MBA from The Wharton School at the University of Pennsylvania and his B.A. from Carleton College. He is a CAIA charterholder, and was the founding President of the CAIA Boston chapter. He is a CFA charterholder.

Ram Thirukkonda. Mr. Thirukkonda is a member of GMO's Asset Allocation team. Prior to joining GMO in 2014, he was a quantitative analyst at Batterymarch Financial Management. Previously, he worked at Fidelity Investments, most recently as the Director of Architecture. Mr. Thirukkonda earned his Bachelor of Technology from the Indian Institute of Technology, his MS in Transportation Engineering from MIT, and his MS in SDM focusing on Financial Engineering from MIT. He is a CFA charterholder.