Introduction of Target-Date Fund and Application of Machine Learning to Optimize Glidepath

Yu-Sheng Chen, Yu-Hsing Chang, Levi Lan, Cheng-Yun Tsai

Abstract

Target-date funds have become an important asset in defined contribution retirement accounts, having gained huge popularity in recent years and now more than 70% of all 401(k) plans offer this option. Given its simplicity and clarity, it could be a satisfactory solution for the average retail investor; however, TDFs may not be the best investment product because it does not take into account factors such as risk aversion and shortfall risk. This project firstly presents a thorough introduction of target-date funds (TDFs) and the comparison between TDFs and traditional mutual funds. Next, we explore the core concept of the glidepath and examine the shortfall risk under different scenarios. The purpose of this project is to optimize an existing declining glide path informed by market environment analysis. Moreover, we also implement the reinforcement learning-based life-time portfolio selection method to develop an entirely new glide path that considers varying market scenarios.

Table of Contents

SECTION I: Target-Date Fund

A. Introduction

A target-date fund is an investment option that is available in many defined contribution plans. It targets the retirement age of investors and allocates investments according to the risk preferences of different life stages. In general, a target-date fund is named after the retirement date and is set up every five years. For instance, the 2025 TDFs are suitable for those who are roughly 60 years old and will retire around 2025. Most target-date funds are funds of funds, which hold a series of other funds: either publicly traded mutual funds or master trusts, and investors pay fees on the underlying funds and usually pay an added fee to the target date fund itself, the size of which depends on the share class that investors purchase.

TDFs allocate money to a combination of different kinds of equity, fixed income, and cash. This combination changes on a preset schedule called a glidepath, usually reallocating assets out of equities and into bonds over time. In detail, when the fund is many years away from its target date, the portfolio is highly concentrated in equities, enabling growth. As the fund approaches its target date, its portfolio rebalances to focus more on fixed income investing. Ultimately, when the target-date fund reaches its final asset allocation, target-date fund investors typically transition to a retirement income fund. The goal of the glidepath is to reduce the risk that investors will experience major losses in value as their retirement date approaches. Therefore, the glidepath of TDFs that intend to move to more conservative asset allocations as retirement approaches seems to be suitable for defined contribution plans since the research shows that older people tend to be more risk-averse than younger ones. The core concept of glidepath and its attractions and limitations will be detailed in the next session.

There are three ways to invest in a target-date fund. As mentioned above, target-date funds are a common preset choice for a 401(k). If investors have a 401(k) and never changed what's in it, there’s a good chance to have a target-date fund. People also can open a brokerage account with a fund manager or online broker to shop for target-date funds, or can purchase one directly from a fund provider like Vanguard, Fidelity or T. Rowe Price, but their choices may be more limited.

Although TDFs are a relatively good option for those having the inability to choose an asset allocation that is consistent with their retirement goals, or a lack of understanding of the importance of portfolio rebalancing, TDFs still have some disadvantages that make it not the best investment choice. First of all, the asset allocation of TDFs takes into account just one factor of risk aversion: the employee’s age, but people have a wide range of risk preferences based on their situations such as marital status, number of children, spouse’s employment, education, and so on. Hence, It's not a customized retirement option.

Second, TDFs are stuck with one provider. When choosing a target-date fund from company X, all investors going to see are company X’s funds as its holdings. While that may not be so bad at first blush, it doesn’t let investors pick and choose the best managers for each individual asset class. Investors could be leaving returns on the table. Using only one provider can also lead to a similar investment style across the underlying mutual funds. Also, the funds also differ in terms of investment style. For instance, investors can find a fund that is made up entirely of index funds. Based on algorithms, such a fund is likely to have lower fees. But investors who prefer active management, with actual human beings tracking market trends and making choices would, need to shop elsewhere.

B. Growth of Target-Date Funds

In 2006, the use of target-date funds accelerated because of a ruling by the Department of Labor that TDFs could be a Qualified Default Investment Alternative in tax-deferred retirement plans. Since that time, TDF assets under management have grown rapidly (see Exhibit 1). The bar in 2000 was too short to see, but TDF's total assets were $8 billion by the end of the year. As of December 31, 2021, target-date mutual fund assets totaled $1.8 trillion, representing an average compound annual growth rate of more than 29%. Retirement accounts held the bulk (85 percent) of target-date mutual fund assets, with 66 percent held through defined contribution retirement accounts and 19 percent held through Individual Retirement Accounts (IRAs).

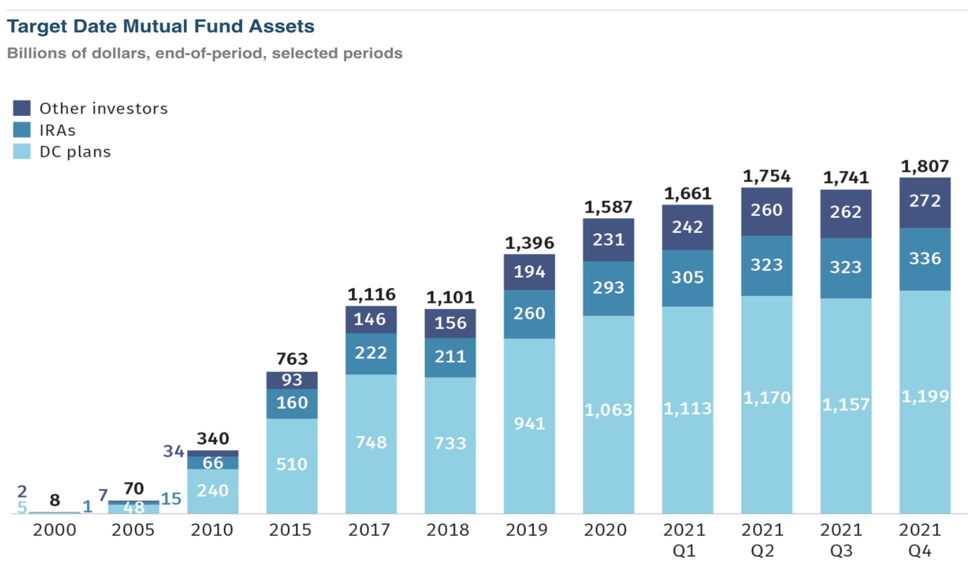


Exhibit 1. Target Date Mutual Fund Assets.
(Source: Investment Company Institute, 2022))

In addition, even though 401(k) plan participants are offered more than 20 investment options and have the opportunity of determining the appropriate asset allocation given their personal risk preferences and investing time horizon, target-date funds were more than one-quarter (27 percent) of 401(k) plan assets in the database (see Exhibit 2). More than half (56 percent) of 401(k) participants held target-date funds at year-end 2018. Looking into the 401(k) Plan Participants' Use of Target Date Funds, it can be even found that younger 401(k) plan participants and new hired 401(k) participants were more likely to hold target-date funds than older participants, indicating that more and more new 401(k) participants would consider target-date funds as their investment product. (Holden, VanDerhei, & Bass, 2021)

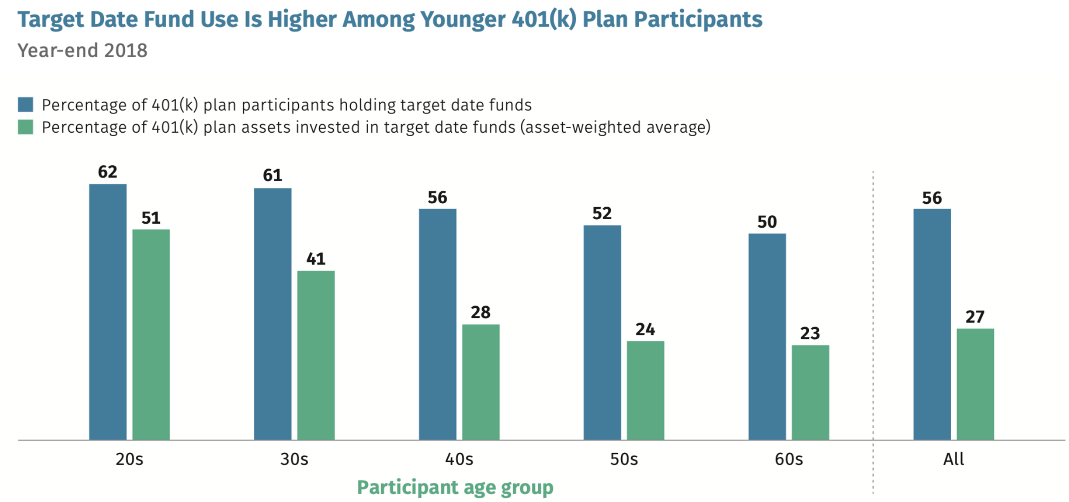


Exhibit 2. Target Date Fund Use Is Higher Among Younger 401(k) Plan Participants. (Source: Tabulations from EBRI/ICI Participant-Directed Retirement Plan Data Collection Project; see Holden, VanDerhei, & Bass, 2021)

C. Comparison between Target-Date Funds and Traditional Mutual Funds

1. Expenses ratio

In general, a reasonable expense ratio for an actively managed portfolio is about 0.5% to 1.5%. For passive or index funds, the typical ratio is about 0.2% but can be as low as 0.02% or less in some cases. However, due to the feature of fund of funds, expenses for a target-date fund include the expense ratio of the underlying mutual fund held by the target-date fund and the added expenses on the target-date fund itself. Thus, what we are curious about is the distribution of TDFs' expenses and how much is the investor paying in total expenses by holding a target-date fund rather than holding a matched set of mutual funds?

Overall, expenses of the underlying funds are low in TDFs because target-date funds often hold low expense mutual fund classes not available to any investor or only available to some investors. For example, 56% of the funds held by all TDFs are institutional class funds, 6.5% are retirement class funds, and 15.93% are master trusts. Besides, taking A share class and no-load as an example (see Exhibit 3), (Elton, Gruber, Souza, & Blake, 2015) showed that if buying mutual funds directly to duplicate the TDF portfolio of A share class, investors should pay fees of 102bps. In the same way, we see that an investor who could buy no-load shares needs to pay fees of 72.5 bps. When we make the comparison, the added expenses on the target-date fund itself are small since an investor who could hold A shares is only paying an additional fee of 9.6 bps for the services provided by the TDF management and who could hold no-load shares is only paying an additional fee of 4 bps.

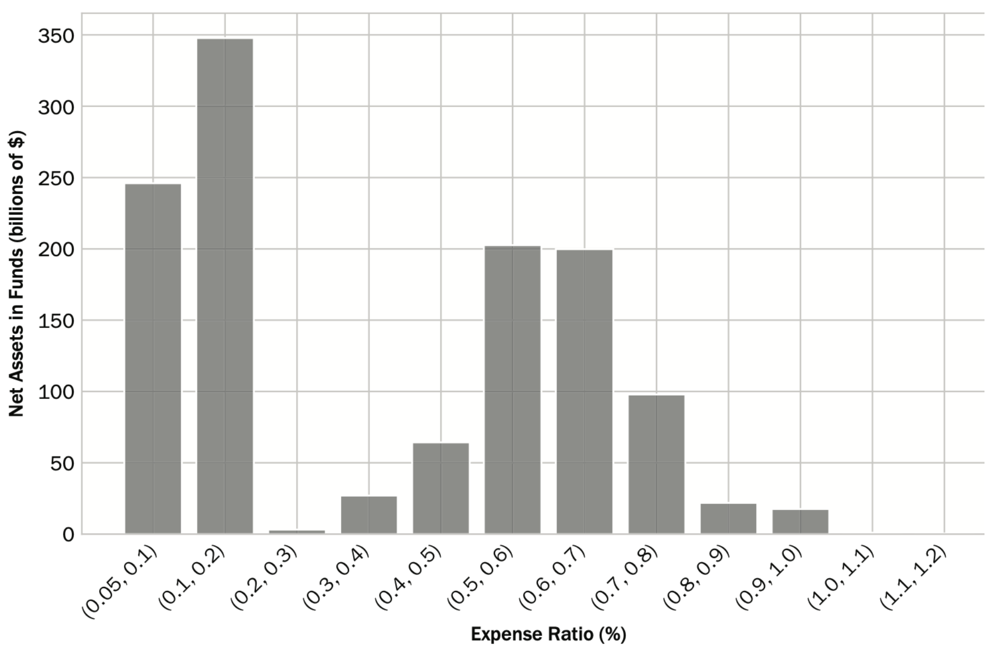
A class fees and no-load class fees for TDFs and fees to match the holdings of the TDF

	Total fees	Underlying fees	Investor matching portfolio fees
A shares (22 funds)	1.117	0.587	1.021
No-load shares	0.767	0.71	0.725

Exhibit 3. A class fees and no-load class fees for TDFs and fees to match the holdings of the TDF.
(Source: Elton, Gruber, Souza, & Blake, 2015)

The average target-date fund had an expense ratio of 0.52% in 2020, but these fees can range from as low as 0.1% to more than 1.5%. The difference in price often revolves around whether the fund leans on cheaper passive investing strategies or more costly actively managed accounts. Hence, according to (Shoven & Walton, 2021), the distribution of expense ratios is roughly bimodal. Half of TDFs have expenses of less than 20 basis points, and the other half have expenses between 50 and 70 basis points. They concluded that low-cost TDFs consist of passive (index) ingredient funds, while the more expensive TDFs consist of actively managed ingredient funds.

In addition, TDFs can be considered an innovation on traditional balanced funds, which hold roughly fixed proportions of asset classes, such as 60% stocks and 40% bonds, so it is interesting to compare each other (see Exhibit 4). For the expenses of balanced funds, it has a similar bimodal pattern to TDFs' distribution. Traditional balanced funds having expense charges between 80 and 120 basis points is much greater than with TDFs. The explanation is that the bulk of TDF assets are in defined contribution plans that tend to be offered in the less expensive share classes (such as R classes or an institutional class); however, balanced funds are more likely to be held in IRA accounts or taxable brokerage accounts and in more expensive retail share classes. From the cumulative distribution function, with TDFs, almost half of the total assets face expense charges of 20 basis points or less. For balanced funds, less than 30% of the assets face such low fees. At the other end of the spectrum, only a very small percentage of the assets in TDFs face expense ratios of greater than 80 basis points, while with balanced funds, about 25% of the assets face expense fees of 80 basis points or more.



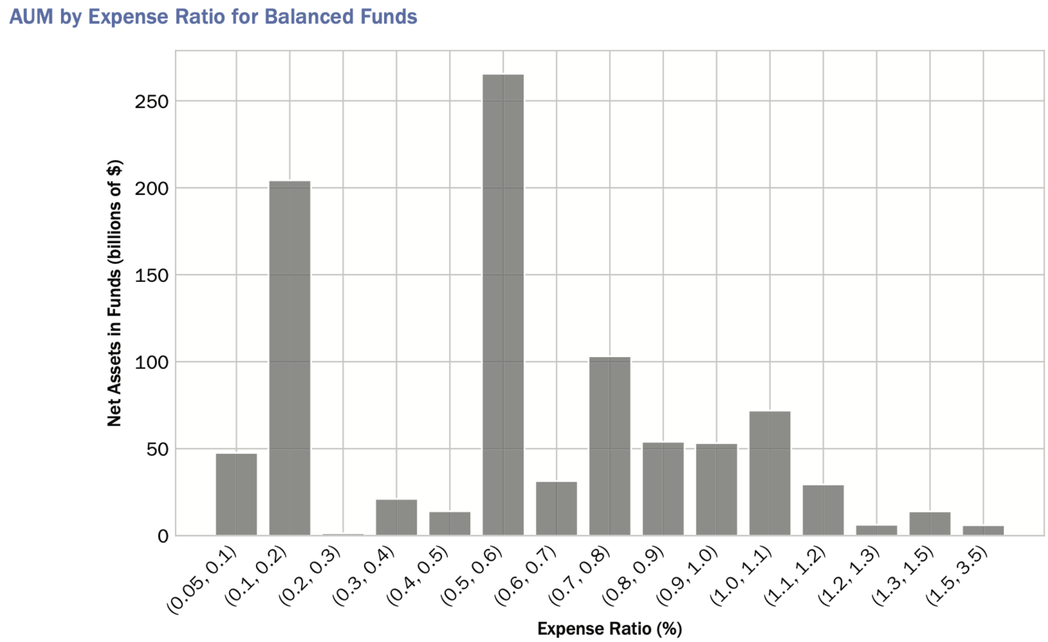


Exhibit 4. AUM by Expense Ratio for TDFs and Balanced Funds
(Source: Shoven & Walton, 2021)

2. Performances

From (Elton, Gruber, Souza, & Blake, 2015), the average alpha over the history across all target-date funds is a negative 20 bps per year and is significantly different from zero at the 1% level. This is the alpha across all the TDFs holdings and is after all expenses on the underlying funds but before the expenses added by the TDF.

Most studies looking at the alpha of the average mutual fund found that the average fund underperformed the index by about 70 basis points. Does this indicate that TDFs have better selectivity performance than other mutual funds? The answer is no. Examining the average expense ratio of the funds they hold shows an expense ratio of around 60 basis points. Most mutual fund research examines fund share classes with an average expense ratio of 110 to 120 basis points. Funds held by TDFs have better alpha mainly because of their ability to hold low expense ratio share classes. When you add the difference in expense ratios, the average alpha of the funds they hold is similar to the negative 70 basis points typically found in mutual fund research. We may conclude that people choose a TDF rather than traditional mutual funds may not because of its performance but because of TDFs' characteristics such as the design of comprehensive asset allocation.

To observe TDFs itself, (Shoven & Walton, 2021) categorized funds with annual expense ratios greater than 30 basis points as "high-cost" and those less than that as "low-cost." The constant scale is in percentage points per month, while the expense ratio scale is in percentage points per year (see Exhibit 5). The results show that the average constant alpha for a low-cost fund is -0.7 (-0.007%) basis points, while the average constant alpha for a high-cost fund is -3.9 (-0.039%) basis points. Because low-cost funds are typically passively managed index funds, it is not surprising that the estimated constant alpha is tightly around zero. In addition, the exhibit shows that the large majority of low-cost funds, especially those that charge less than 20 basis points annually, have constants within two basis points of zero on a monthly basis. Nevertheless, the distribution of constant alpha for the high-cost funds is highly dispersed; 25% have constants greater than 0, and they have a range of -16.7 basis points to +4.7 basis points.

A quarter of the high-cost funds are successful in at least meeting the market return, while the other three quarters fall short. Furthermore, the funds that appear to have even a chance of having a positive constant charge less than 80 basis points, with the most successful charging around 60 basis points. The negative relationship between constants alpha and expense ratios could be explained by the opinion that the various positions taken by actively managed funds likely will aggregate out to be a diversified portfolio, which will on average deliver the market return, scaled by the amount of systematic risk taken on, but with the additional disadvantage of removing high expenses from the fund's return. From the fitted slope of -7.9 basis points in the monthly constant alpha per 1 percentage increase in expenses, this translates to approximately 95 basis points per year. This means that the constant alpha decreases approximately 1 for 1 with an increase in net expense ratio. Therefore, once we control for asset allocation, high-cost funds produce gross returns equal to their low-cost counterparts, but net returns differ by the difference in net expense ratios, and it seems that expensive funds have difficulty in overcoming their expense disadvantage.

They also performed a similar exercise on balanced funds, and the results are shown in the right figure. Balanced funds generally have a wider range of constants than TDFs. This may be due to the wider coverage of the Balanced Fund category. This may be due to the fact that the class of balanced fund covers a wider range of funds with different objectives, including different active strategies, while TDFs have a narrower range of objectives. In the same pattern as the TDF data, low-cost funds show a larger clustering around zero than high-cost funds.

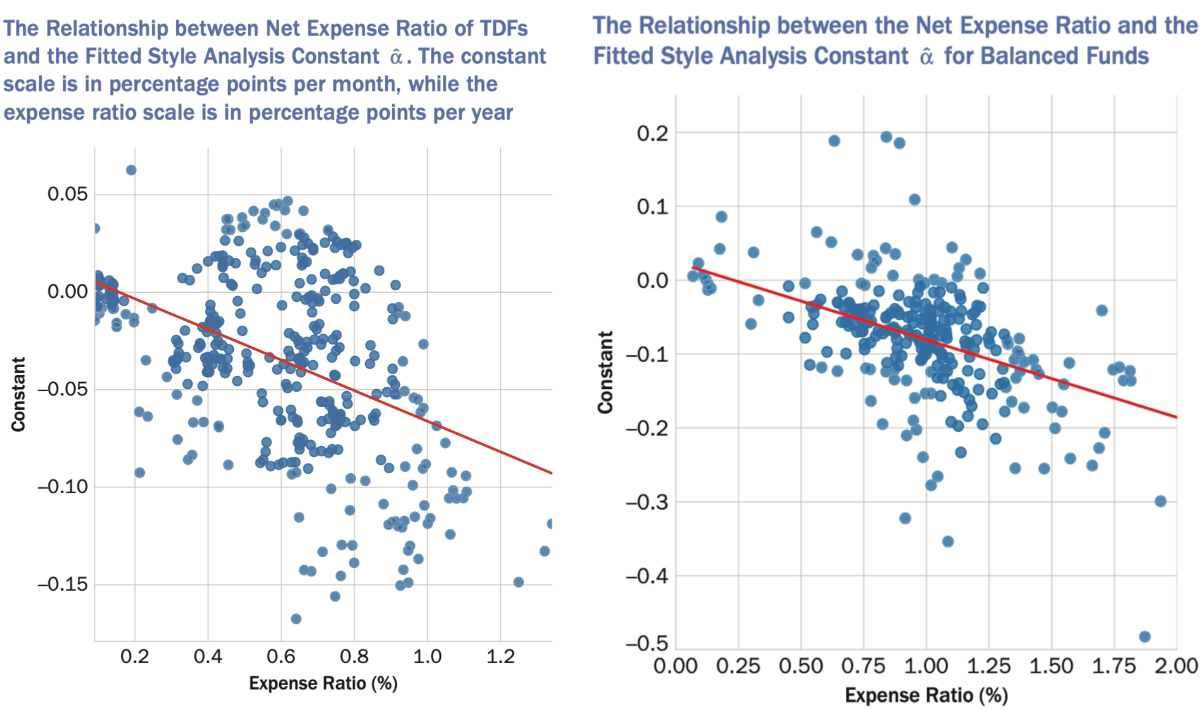


Exhibit 5. The Relationship between Net Expense Ratio of TDFs (Balanced Funds) and the Fitted Style Analysis Constant Alpha. (Source: Shoven & Walton, 2021)

3. Risk

Buried in some of the mutual funds in a target-date fund may be riskier securities than the investor can tolerate such as stock in companies that are in the emerging market, foreign bonds or junk bonds. For example, during the market crash from February 19 to March 23, 2020, the S&P 500 stock index fell 33.8%, while the Vanguard Total Stock Market Index fund(VTI) fell 35%. The Vanguard Total Bond Market Index ETF (BND) had a total return of about –1.5%. However, most of the distant future TDFs (the 2045, 2050, 2055, and 2060+) lost between 30% and 35% of their value over this five-week period, failing to do significantly better than an all-equity portfolio invested in the S&P 500. Most of the near future TDFs (the 2025 funds) lost between 20% and 25% of their value. While the 2025 funds, presumably held by workers within roughly five years of retirement, did better than the long-duration TDFs, they still lost more than one-fifth of their value in five weeks. This likely disappointed their shareholders, given the promise of the glide paths becoming safer and more conservative as retirement approaches. Therefore, although TDFs are considered a mix of equities and bonds, TDFs could be as risky as stock index funds.

SECTION II: Introduction of Glide Path and Analysis of Shortfall Risk

A. The core concept of the glide path

Human capital theory is the bedrock of the glidepath, the core concept of target-date funds. Human capital is commonly defined as the present value of a worker's projected labor income and is treated more like fixed income. As a result, the bulk of a worker's investments in the early career should concentrate on equity to achieve the desirable stock-bond ratio.

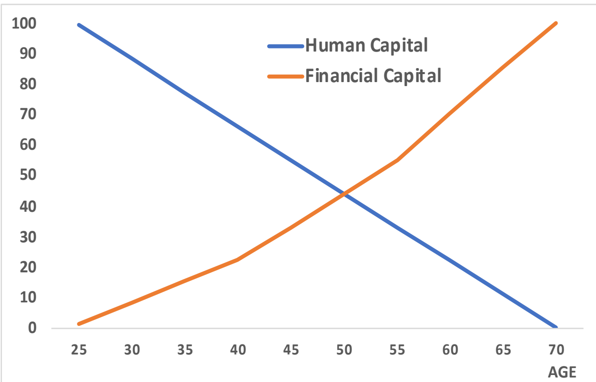


Exhibit 6. Human Capital vs. Financial Capital

(Source: Author)

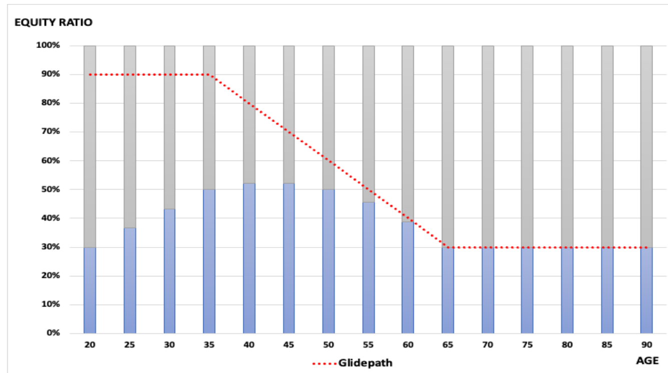


Exhibit 7. Extended Personal Portfolio

(Source: Author)

A glidepath is essentially the way the asset allocation within a target date fund changes over time; that is, the allocation between the risky asset and the reserve asset should follow this glidepath line to adjust as the investor ages. The underlying concept of this investment strategy is the human capital theory. To be more specific, the investment portfolio includes a large portion of equity in the early employment years because of the high value of human capital, the present value of future labor income. As the time passes, the human capital has been gradually depleted and thus the portfolio shifts from more aggressive but volatile investments (equities) to more stable investments (bonds) to help manage risk through time. There is, nevertheless, no universally accepted optimal answer as to what level of equity exposure is appropriate to diversify investors' human capital. Generally, the equity allocation begins at the level of 85% ~ 95% and ends up at the level of 20% ~ 30%. The extended personal portfolio, which accounts for human capital, shows how the investment composition will be like if the financial capital is invested along the glidepath.

B. Key factors in the design of the glidepath

At the beginning, the method of designing the glidepath merely considers one factor - an investor's age. Its simplicity and transparency have indeed helped target-date funds grow tremendously in popularity. As the market of target-date funds becomes more sophisticated and competitive, customized methods of planning the glidepath have been proposed. For instance, Vanguard's Life Cycle Investing Model allows the incorporation of several other factors, such as (1) net wealth, (2) labor income, (3) capital market risk-return expectations, (4) different degrees of investor risk tolerance, (5) investor aversion to myopic losses (negative returns), and (6) contribution rates required to successfully meet funding goals. Some funds, on the other hand, predominantly focus on the risk aspects. For example, a fund is likely to hold a relatively large portion of Treasury Inflation Protected Securities (TIPS) if it prioritizes inflation risk, while a fund will reduce its equity position slower if its priority is longevity risk. In this way, investors are allowed for a presumably best fit selection subject to their own expectations and limitations.

The glidepath universe can be broadly divided into two camps – to-date vs. through-date. A to-date glidepath stops shifting the equity allocation once the target retirement date is hit, whereas a through-date glidepath continues lowering the equity allocation for another 10 to 15 years past the target retirement date to de-risk.

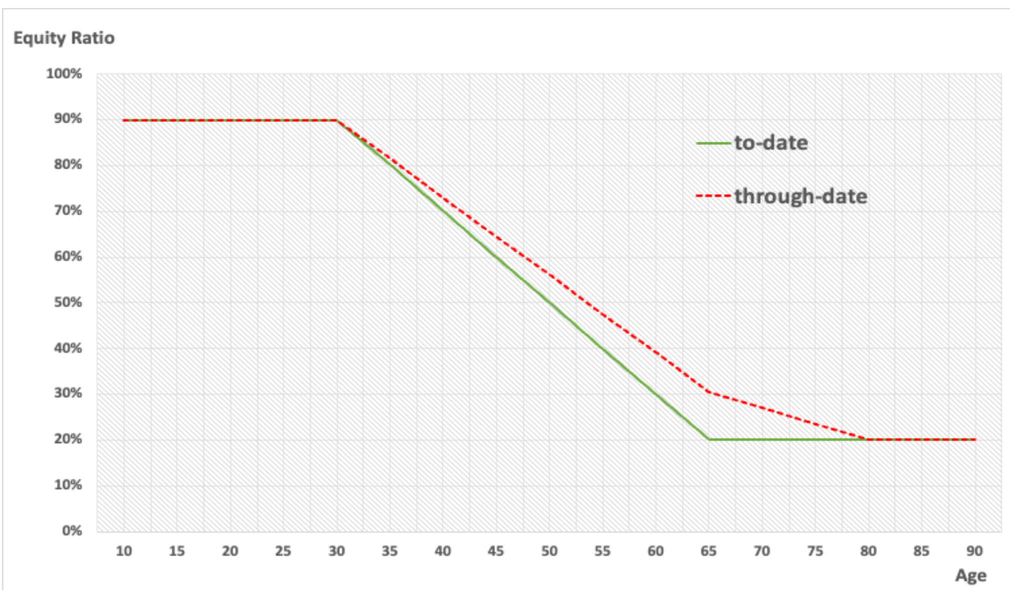


Exhibit 8. To-Date vs. Through-Date Glidepath (Source: Author)

C. The Prevailing Glidepath

Each kind of glidepath has its own benefits and shortcomings. The “to” glidepath is usually adopted to protect against sequencing risk and market risk, while the through-date glidepath is embraced when the focus is on longevity risk at the expense of increased market risk and volatility. In light of current trends — longer lifespans and the presence of pension benefits — the majority of assets under management are in through-date funds.

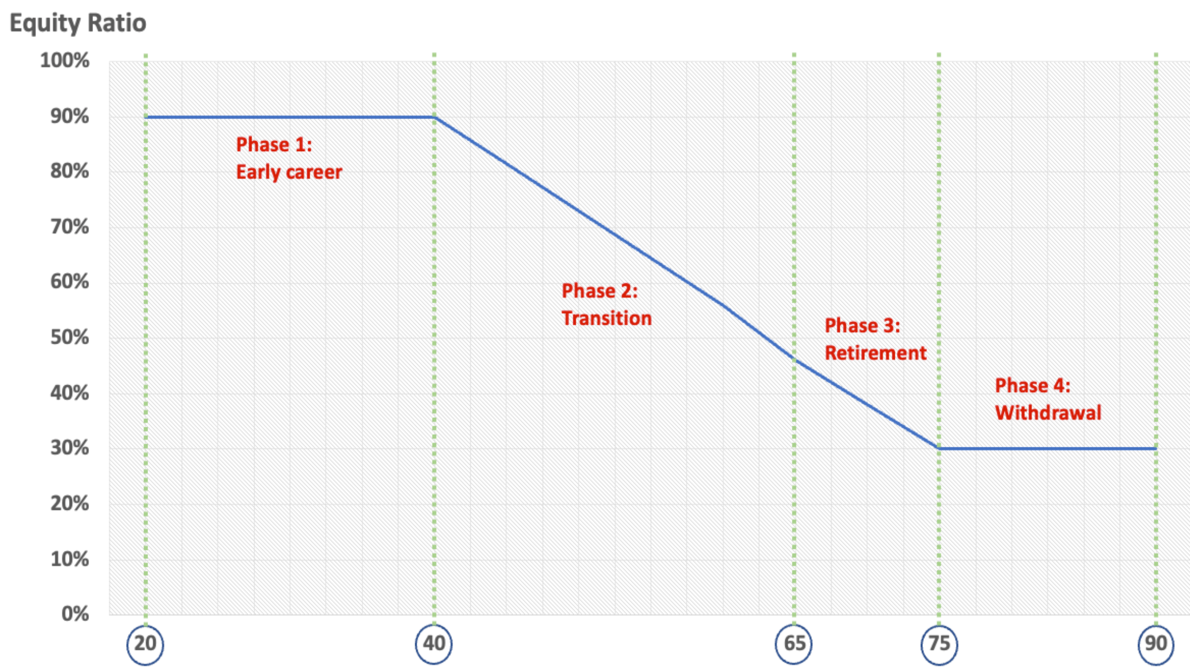


Exhibit 9. Phases of Through-Date Glidepath (Source: Author)

Over the past decade, the equity market has been on the overall upward trend and the treasury yields have been constantly declining. Accordingly, through-date funds have indeed performed better in comparison to to-date funds owning to the enhanced equity position. However, the current favorable market context won’t persist forever; as a result, whether the justification of the higher stock exposure is robust remains untested against the opposite market backdrop. There has been a heated debate over the suitability of incorporating longevity into the design of glidepath as living a long time is inherently not an investment risk. Besides, there are a variety of insurance contracts available to tackle this concern. Consequently, a few target-date fund providers deem capital preservation as the top priority and solely focus on to-date glidepath.

	Ticker	Fund Name	Equity Allocation *	Return 2021	Return Q1 2020
	VTWNX	Vanguard Target Retirement 2020	43.4%	8.17%	-10.74%
through-date	FFFDX	Fidelity Freedom 2020	48.5%	8.91%	-12.34%
	AACTX	American Funds Target Date Retirement 2020	46.3%	10.24%	-8.85%
	TRRBX	T. Rowe Price Retirement 2020	54.1%	10.47%	-14.21%
	Mean		48.1%	9.4%	-11.5%
	LIRIX	Blackrock LifePath Index	39.8%	6.86%	-7.59%
	ARTOX	American Century One Choice	44.7%	8.89%	-10.15%
to-date	JTTIX	JPMorgan SmartRetirement	32.6%	6.08%	-9.40%
	URTNX	USAA Target Retirement	38.1%	7.57%	-7.60%
	Mean		38.8%	7.4%	-8.7%

** as of Feb 28th 2022*

Exhibit 10. Through-date vs. To-date Performance During the Rising/Declining Market (Source: Author)

D. Real-world implementation of target-date funds

Several large wealth management companies, such as Vanguard, Fidelity and T. Rowe Price, offered a series of target-date funds with target retirement dates for every five years from 2015 to 2065. These target-date funds are largely structured as fund of funds (FoF). Taking Vanguard Target Retirement 2050 Fund (VFIFX) for illustration, its predefined glidepath indicates that the fund should possess roughly 90% of equities and 10% bonds in 2022. Instead of directly holding financial instruments, the fund holds 54% of Vanguard Total Stock Market Index Fund Institutional Plus Shares + 36.1% Vanguard Total International Stock Index Fund Investor Shares and 6.8% Vanguard Total Bond Market II Index Fund Investor Shares + 3.1% Vanguard Total International Bond II Index Fund.

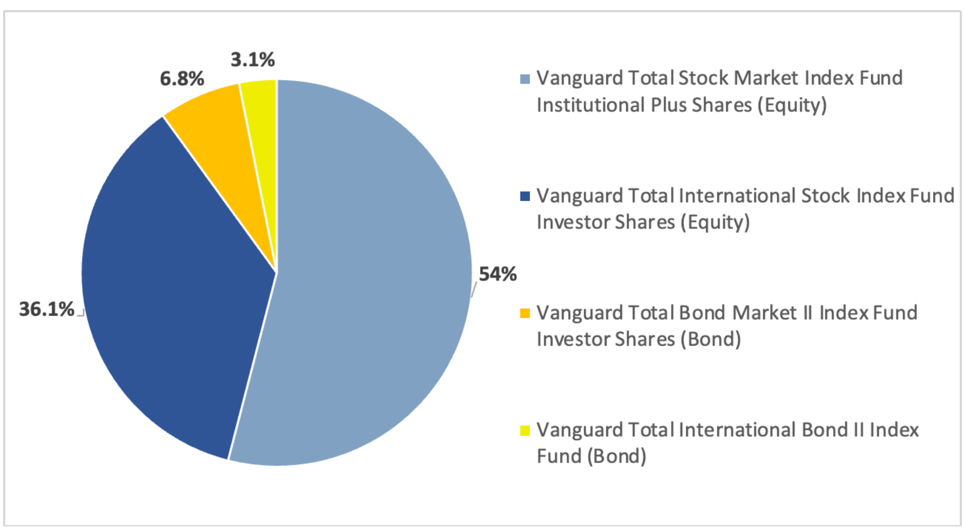


Exhibit 11. Decomposition of Vanguard Target Retirement 2050 Fund (Source: Vanguard)

Though the equity-bond allocation ratio is fixed by the predefined glidepath, there may be some discretion for target-date fund-managers to commit active investment within each asset class as the risk and return characteristics of different asset classes vary over time, such as market risk premiums, volatilities, and correlations. Instead of fully disregarding these subjective forward-looking estimates and merely taking on the broad market risk, the fund-managers may form views about the relative attractiveness of different asset classes and exploit them mainly through two ways: some are adopting a top-down tactical asset allocation to take advantage of strong market sectors; the others are using underlying fund building blocks like actively managed funds, which themselves have flexible asset allocations, to achieve relative allocation to sub-asset classes in pursuit of better performance.

Still, the glidepath is the most important determinant of the return variability and long-term performance of a target-date fund due to limited market-timing opportunities. A slight difference in the glidepaths can lead to significant differences in the performance

E. Rebalancing Policy and Its Market Impact

Some crucial questions regarding the rebalancing policy includes how often the portfolio is rebalanced, what the hard limit that will trigger rebalancing is and what rebalancing strategy is adopted. After reviewing prospectuses and annual reports of several large target-target funds, it is interesting, albeit unsurprising in retrospect, to find out that little information has been disclosed about the rebalancing policy. This kind of obscurity probably serves the purpose of averting the potential abuse. Take Vanguard for example, its total AUM under those target-date funds within the transition and retirement periods is around USD $500 billion and the slope of its glidepath is - 2% / year, implying that Vanguard alone needs to clip roughly USD $10 billion equity position each year for rebalancing. It is a bulky amount and definitely leaves some impact on the market price. If the detailed rebalancing schedule and process are released, competitors and hedge funds are bound to take advantage or even manipulate the prices to benefit. As a result, it's reasonable for the large target-target fund providers to hide their own practices.

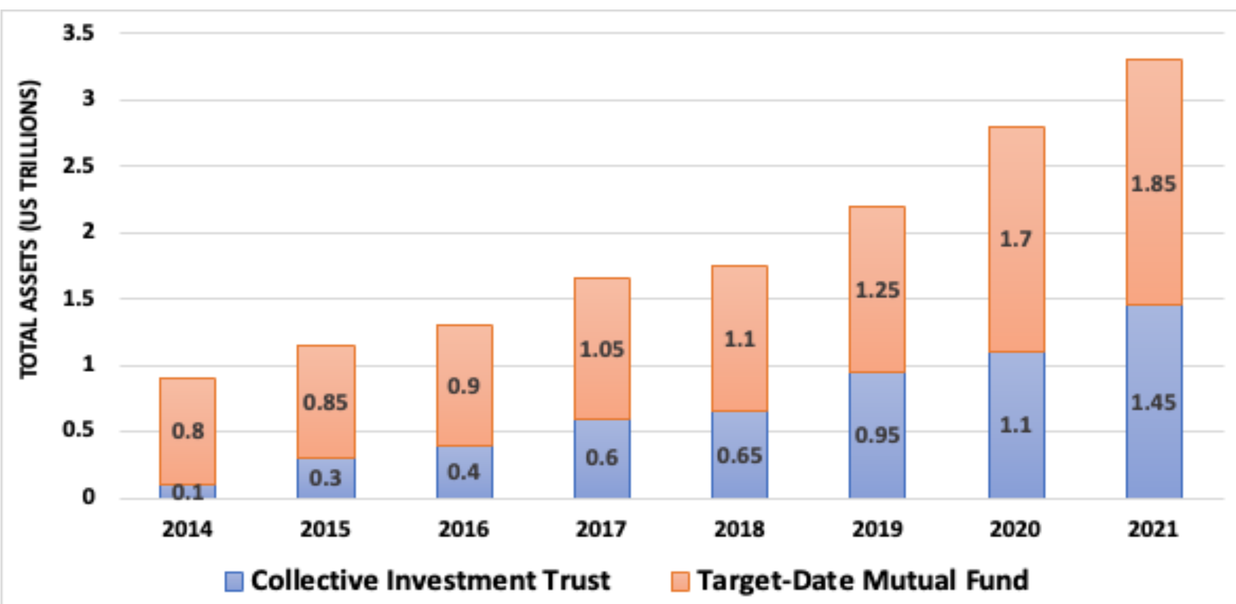


Exhibit 12. Assets under Glidepath-Based Investment Vehicles (Source: Morningstar)

F. Shortfall Risk Analysis of Different Glidepaths and Withdrawal Rate

Longevity risk is always a genuine concern for retirement savers and hence the annual withdrawal amounts for retirement portfolios have been a crucial issue. The thought of being many years into retirement and seeing an account with a zero balance is a reasonably tangible concern. The optimal path of the personal asset should be the one like the exhibit 13.

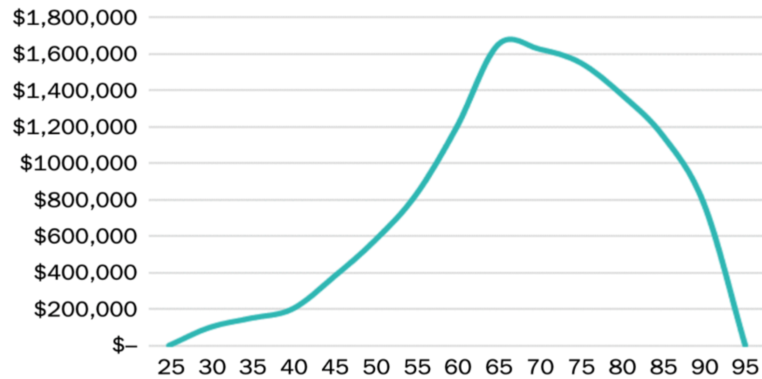


Exhibit 13. The Sample Path of Personal Asset Value (Source: Author)

In this part, Monte Carlo simulations are used to examine the potential impact of the different withdrawal rates and glidepaths between to-date and through-date on the shortfall risk, namely, the probability that the asset value drops into the negative territory.

The model specifications are as follows:

1. Income process:

The income process before the retirement is modeled as stochastic: $\log(I_t) = f(t) + P_t + U_t$

- $f(t)\ is\ the\ baseline\ path\ of\ earnings\ and\ is\ a\ deterministic\ function\ of\ age$
- $P_t\ is\ the\ permanent\ component\ of\ income\ shock: P_t = P_{t-1} + N(0, \sigma_P^2)$
- $T_t\ is\ the\ transitory\ component\ of\ income\ shock: T_t \sim N(0, \sigma_T^2), uncorrelated\ with\ P_t$

The income process after the retirement (pension benefits) is modeled as deterministic:

$$\log(I_t) = \log(\lambda) + f(K) + P_K$$

- $K\ is\ the\ retirement\ age$
- $\lambda\ is\ a\ fixed\ percentage\ of\ final\ period\ income$

2. Returns of each asset class:

Not to complicate the model and incorporate too many assumptions, only S&P500 and Bloomberg US Aggregate Bond Index are used to represent the equity and bond position respectively. Historical statistics for return, standard deviation and correlation are computed using the past 33 years daily data.

3. Death rate of each age:
The life table from SSA database is used to calculate the conditional death rate for each age.

The Exhibit 14 shows what the simulation looks like — 100 sample paths of the asset value were produced and drawn. Each path represents the net asset value of a person's lifetime portfolio. If a path drops below the zero, that path indicates that an investor outlives his/her assets. Given that the conditional death date is incorporated into the simulation, some paths may end early, either because the shortfall happens, or the person dies prematurely.

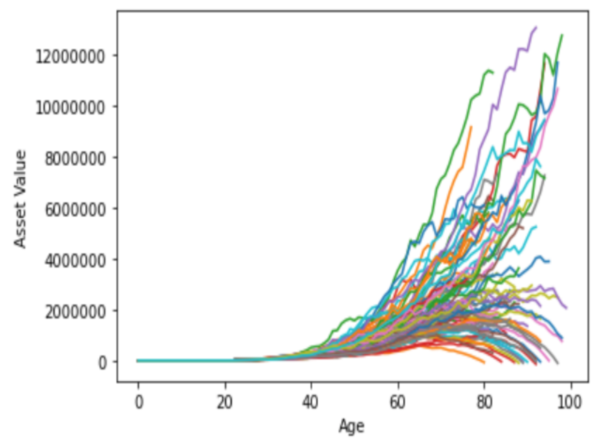


Exhibit 14. 100 Simulations of Asset Value Path

(Source: Author)

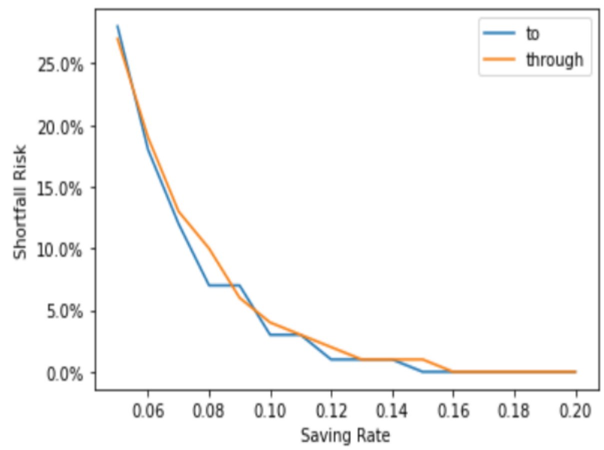


Exhibit 15. Saving Rates vs. Shortfall Risk

(Source: Author)

In order to calculate the probability of the shortfall risk, 5000 simulations have been conducted for each scenario. Through-date funds are dominant lately because they are touted for its particular focus on the longevity risk, but the result of the simulation (Exhibit 15, 16, 17) demonstrates that they perform quite similar from the perspective of shortfall risk. Admittedly, the results of the simulation are pretty sensitive to several assumptions; however, whether higher levels of stock in the portfolio clearly provided better retirement outcomes seems to be questionable.

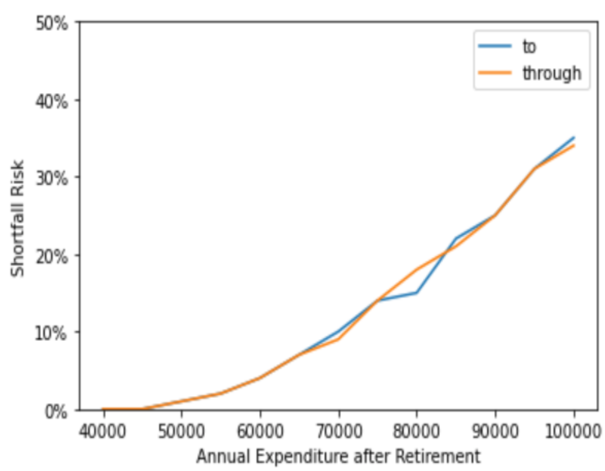


Exhibit 16. Annual Expense Rates vs. Shortfall Risk

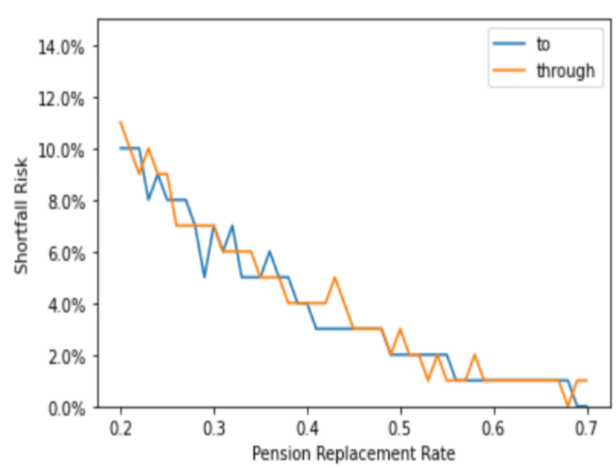


Exhibit 17. Pension Rep Rates vs. Shortfall Risk

G. Stepping into the Future

A fixed glidepath, admittedly, is an imperfect strategy, but its imperfections are easily overweighed by the great simplicity and efficiency it brings to large numbers of investors. Even so, financial institutions continually devise new variants of target-date funds to distinguish from the pack in this competitive target-date fund arena. Some providers attempt to constructing more complex glidepaths for an allegedly better optimization; some providers, on the other hand, focus on achieving better returns through short-term tactical asset allocation to attract investors. The others are devoted to presenting customized or personalized target-date structures. While these augmented offerings may appear to solve additional problems or accommodate extra objectives, at the meantime these added features somewhat muddle the original allures of target-date fund – its simplicity and convenience – as well as imperil the attainment of the primary goal to some extent. In addition, as the products become more and more specialized, some benefits of aggregation may shrink or even vanish. For instance, there are nearly USD $600 billion under the Vanguard's target-date funds' management and this gargantuan scale leads to the relatively low expense ratio, 0.08% vs. 0.5% in the industry. The newly launched target-date funds usually charge even higher fees than the average.

Therefore, in selecting a suite of target-date funds, investors must keep a close eye on the ball and assess carefully whether the extra management fees for the distinguishing characteristics are genuinely worth incurring. Cost advantage is particularly important for target-date funds. Even though a dozen basis points of annual extra expense may seem negligible in the first place, the compounded cost can be terrifyingly significant after several decades pass as target-date funds are in nature ultra-long, life-cycle investment.

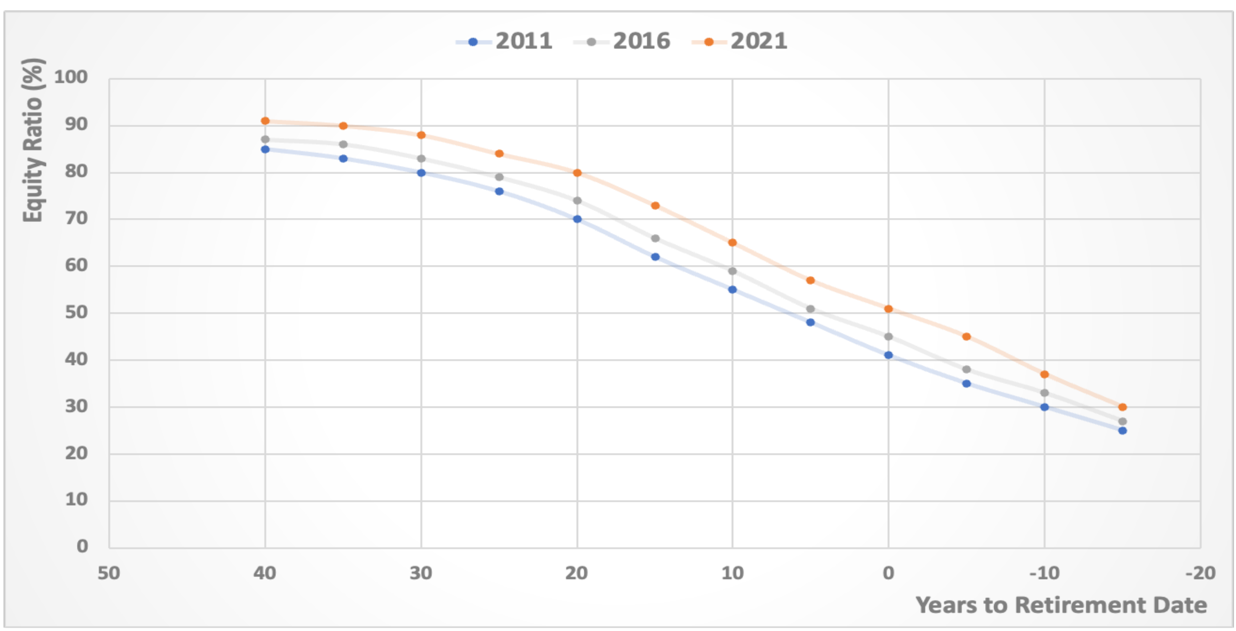


Exhibit 18. Comparison of Glidepaths between Different Vintage Years (Source: Author)

SECTION III: Optimize the Existing Glide Path by Predicted Market Movement

A. Introduction

Target date strategies are now the primary retirement investment vehicle for a large percentage of defined contribution (DC) workplace retirement plans. Given this popularity, plan sponsors, advisors, and consultants should evaluate a target date provider's glide path.

The glide path is an important investment driver of whether an investor will achieve their desired retirement outcomes. Glide path strategies mainly based on investor's time horizon, investor risk tolerance and investor's target income goal, while there is more to consider. Take 1290 Retirement 2020 Funds (1290 Funds, 2021) as instance (see Exhibit 19), as time approaches retirement year (2020), the fund will slide down the glide path, allocating fewer assets to equities and more to fixed income 80% as the allocation becomes more conservative.

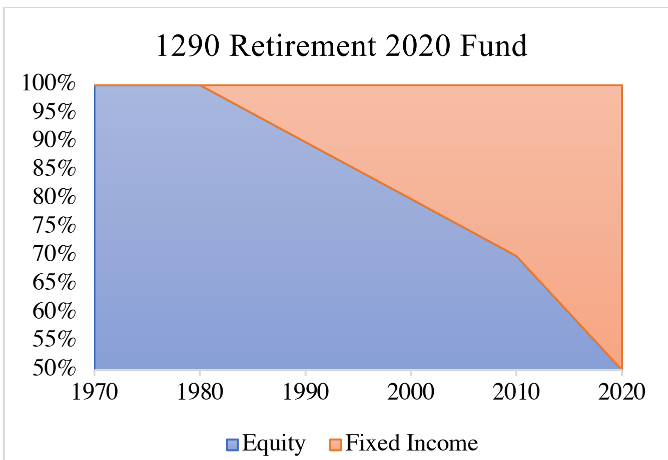


Exhibit 19. 1290 Retirement Fund (Source: 1290 Funds, 2021)

Apparently, the glide path is typically based on the time from the retirement year, while ignoring market environments and potential movement. Market environments can shift, and positioning a target date strategy for any one market environment based only on the investor's time horizon could lead to a substandard outcome. Hence, inspired by the Fidelity report (2020), we propose to utilize machine-learning frameworks to predict the market movement, informing the glide path optimization. The goal is to optimize the glide path of 1290 Retirement 2020 Funds for resiliency to different market environments while seeking to enhance the outcomes.

B. Data and Analysis

Intuitively, the American equity market is a long-term bull market accompanied with several short-term bear markets. To find out those short-term recessions, we choose S&P 500 Index for the machine-learning model. It is a stock market index based on the capitalization of 500 American large companies, and widely followed by worldwide market participants. Also, it is the leading economic indicator with the best track record at identifying recessionary troughs before they occur. We use S&P 500 daily index time series from January 1949 to December 1969 containing 4,935 observations to train the XGBoost model, and from January 1970 to December 2020 to conduct prediction.

For backtesting and comparing the performance of original glide path and new glide path informed by the market movement, we choose S&P 500 Index (from January 1970 to December 2020) to stand for equity position and Fidelity Investment Grade Bond Fund (from January 1980 to December 2020) to stand for fixed income position. Admittedly, target date funds typically hold multiple equity-related and fixed-income-related funds, we choose only one for each to simplify the problem and focus on revealing the difference between time-horizon-based and market-analysis-based glide path. In addition, there is a ten-year gap (from 1970 to 1980) between S&P 500 Index and Fidelity Investment Grade Bond Fund. To address it, we revise the glidepath of 1290 Retirement 2020 Funds to keep 100% equity holdings for the first ten years.

	Predicting Market Movement by XGBoost		Backtesting and Comparing the Performance of Glide Paths
S&P 500 Index	Training Period	1949 – 1969	-
	Testing Period	1970 – 2020	1970 – 2020
Fidelity Investment Grade Bond Fund	-		1980 – 2020

Exhibit 20. Data and corresponding time period used in this section (Source: Author)

C. Proposed Methodology

1. Machine-learning model: XGBoost

Recently, new machine learning techniques such as decision trees, support vector machines and neural networks among others, followed by alternative datasets and cheap computational processing power became available, allowing for alternative ways to determine (or predict) the market movement.

Mamed's works (2018) develop a supervised machine learning classifier using Random Forest technique to identify economic regimes using the S&P 500 stock market index series. In the present work, we use XGBoost instead for the reason that XGBoost is better to deal with an imbalanced dataset.

XGBoost stands for eXtreme Gradient Boosting and is developed on the framework of gradient boosting. The sequential ensemble methods, also known as "boosting", creates a sequence of decision trees that attempt to correct the mistakes of the decision trees before them in the sequence (see Exhibit 21). The first decision tree is built on training data, the second decision tree improves the first one, and so on. Gradient boosting is an approach where new decision trees are created that predict the residuals or errors of prior decision trees and then added together to make the final prediction. In gradient boosting descent while combining the multiple decision trees, the loss function is minimized.

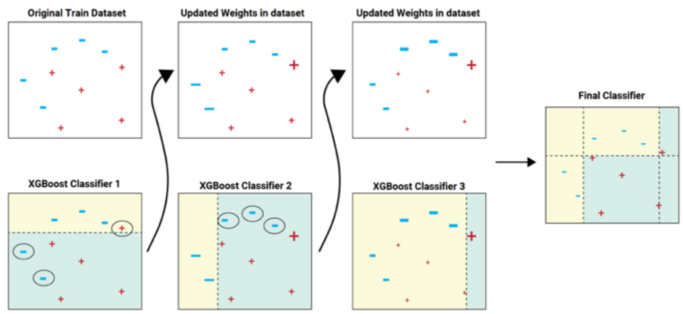


Exhibit 21. Gradient Boosting (Source: Quantinsti)

In applications like market movement prediction, the classes will almost certainly be imbalanced where the number of bull market movements will be huge when compared with bear market movements. In XGBoost, when the model fails to predict the anomaly for the first time, it gives more preferences and weightage to it in the upcoming iterations thereby increasing its ability to predict the class with low participation; but we cannot assure that random forest will treat the class imbalance with a proper process.

2. Target Value and Features

In the two-state market movements representation, the outcome is represented by (i) state 0 or (ii) state 1 for each day t:

$$FutureReturn(FR)_{t,T1} = \begin{cases} 0, & if\ (\dfrac{Price_{t+T1}}{Price_t} - 1) \geq 0 \\ 1, & if\ (\dfrac{Price_{t+T1}}{Price_t} - 1) < 0 \end{cases}$$

$$Market\ Movement\ _{t,T2} \begin{cases} State\ 0, & if\ \ Mode(FR_t, FR_{t+1}, \ldots, FR_{t+T2}\ ) = 0 \\ tate\ 1, & if\ \ Mode(FR_t, FR_{t+1}, \ldots, FR_{t+T2}\ ) = 1 \end{cases}$$

, where $T1 = 5$ and $T2 = 63$. That is, state 0 stands for bull market movement, meaning most of weekly returns are positive in the following quarter; state 1 stands for bear market movement, meaning most of weekly returns are negative in the following quarter.

There are three kinds of features, predictor variables, for each day t. First one is returns under past determined time period (see first equation below). Second one is volatility under a past determined time period (see second equation below). Third one is technical indicators, including Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), On-Balance Volume (OBV), Accumulation Distribution Indicator (ADI) and Force Index (FI).

$$Return_{t,T3} = \frac{Price_t}{Price_{t-T3}} - 1$$

$$Volatility_{t,T3} = std\{Price_{t-T3}, Price_{t-T3+1}, \ldots, Price_t\}$$

, where $T3 = 10$ for two-week, $T3 = 21$ for monthly, $T3 = 63$ for quarterly, $T3 = 126$ for semi-annually and $T3 = 252$ for annually.

3. Tuning Hyperparameter

Using the previous configurations, we perform a grid search for the maximum tree depth parameter in the interval between 2 and 5, number of decision tree parameter in the interval of 2 to 20, and parameter controlling the balance of positive and negative weights in the interval of 1 to 3.75.

It is worth noticing that it is an imbalance dataset, parameter controlling the balance of positive and negative weights is useful to deal with imbalanced issues. Also, considering the imbalanced dataset, Receiver Operating Characteristic Curve (ROC AUC) is selected as scoring strategy to compare performance of each parameter set.

4. Testing Various Sets of Decreasing Rates for each Market Movement and Backtesting

Under the original glidepath, the percentage of equity goes down 1% each year from 1980 to 2010, and accelerates to 2% from 2010 to 2020. The target-date fund adjusts its portfolio annually, while we prefer to adjust semi-annually to make the difference between two glide paths more obvious.

There are two decreasing ratios for equity holdings: decreasing_ratio_0 for state 0 (bull market movement) and decreasing_ratio_1 for state 1 (bear market movement). Since it is better to decrease the equity positions to reduce losses during the bear market, decreasing_ratio_1 is larger than decreasing_ratio_0, and vice versa. We test multiple sets of decreasing ratios: decreasing_ratio_1 in the interval of 0.125% to 1.75%, decreasing_ratio_0 in the interval of 0.125% to decreasing_ratio_1. The suitable set of decreasing rates is chosen to build the glidepath.

Then, we evaluate the performance of the produced glide path contrasting the original glide path by the statistics of cumulative return from 1970 to 2020, Sharpe ratio and Drawdown.

D. Results

1. Prediction of Market Movement

It is obvious that the dataset is imbalanced (see Exhibit 22). The true number of State 0 is significantly greater than state 1, which corresponds with the nature of long-term bull U.S. equity market. As can be seen in the confusion matrix, XGBoost model does well on predicting bull market movement, while missing some bear market movements. Accuracy score is 67% (see Exhibit 23). It is acceptable, because it is better than 50%, the probability of correct prediction by choosing one of two market movements randomly.

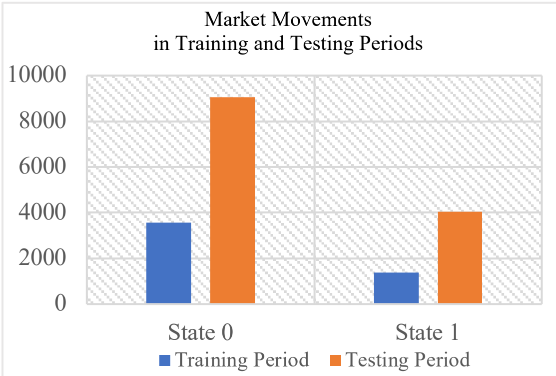


Exhibit 22. Market movements in training and testing period (Source: Author)

	Precision	Recall	F1-score	Support
State 0	0.71	0.90	0.79	9058
State 1	0.42	0.16	0.23	4030
Accuracy			0.67	13088

Exhibit 23. Classification Report for XGBoost Model (Source: Author)

To evaluate the prediction of market movement (see Exhibit 24), we compare its cumulative return under switching strategy (long for predicted state 0, and short for predicted state 1) with buy and hold strategy (always long). As can be seen in Exhibit 25, switching strategy outperforms buy and hold strategy, meaning XGBoost model somehow correctly distinguishes the bull and bear market movements and prevents from losses in downside market movement.

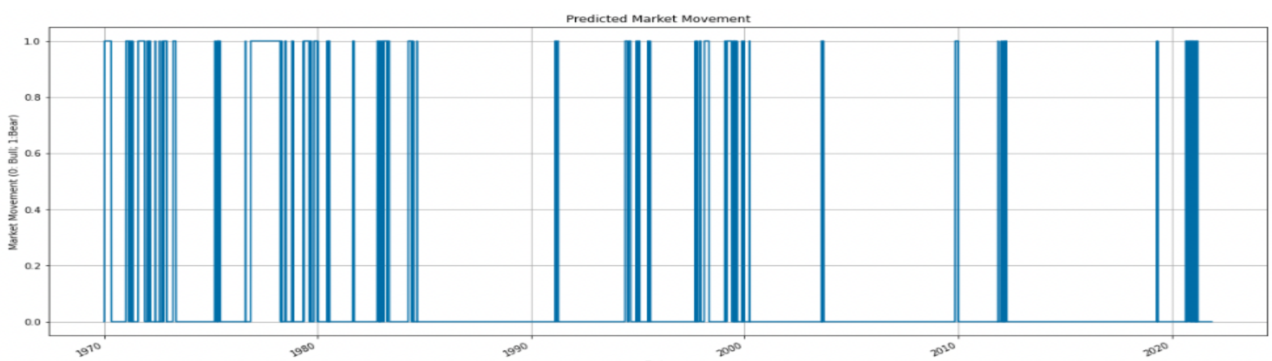


Exhibit 24. Predicted market movement in testing period (Source: Author)

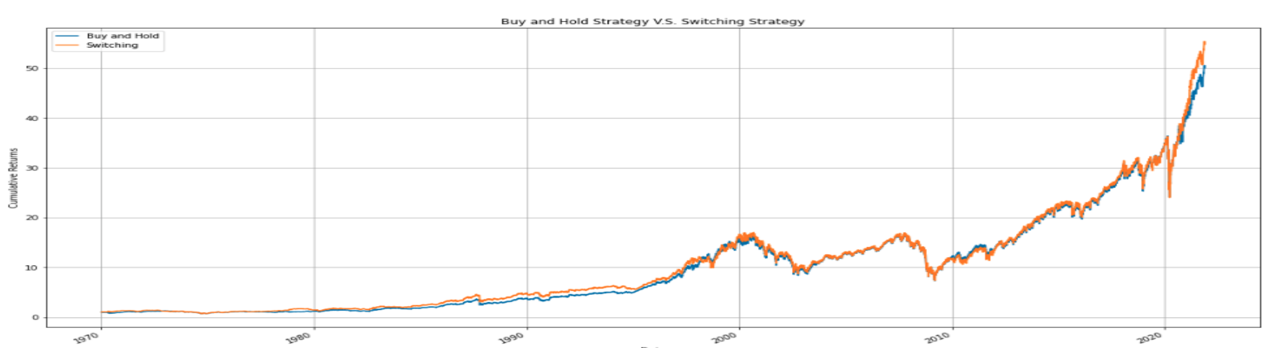


Exhibit 25. Cumulative Returns in testing period (Source: Author)

To further explain why outperformance implies the efficiency of market movement prediction, we dive into the number of correct transactions (long for bull market movement or short for bear market movement) and incorrect transactions (short for bull market movement or long for bear market movement). Compared to Buy and Hold Strategy, Switching Strategy does 726 more correct short transactions while doing 768 more incorrect long transactions. The overall percentages of correct transactions are similar (53.1% for Buy and Hold Strategy; 52.8% for Switching Strategy). However, the equity market is negative-skewed. Geometric mean of those 768 incorrect long transactions is 0.35%, which is the punishment for switching strategy. Geometric mean of 726 correct short transactions is 0.38%, which is the reward for switching strategy. Greater reward and lower punishment lead to higher cumulative return for switching strategy. To conclude, it is the XGBoost model that somehow successfully recognizes bear market movements from long-term bull market movements, resulting in the outperformance of switching strategy.

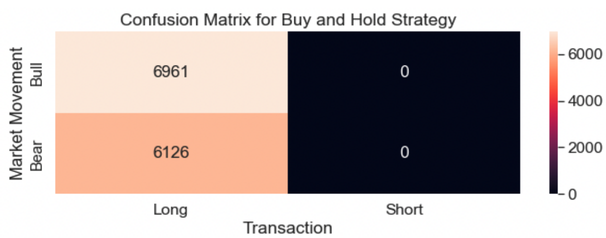


Exhibit 26. Confusion matrix for buy and hold strategy (Source: Author)

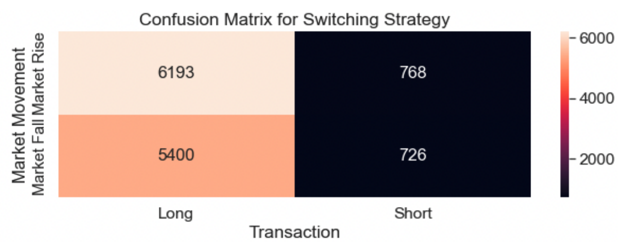


Exhibit 27. Confusion Matrix for Switching Strategy (Source: Author)

2. Construct Optimized Glide Path

Different sets of decreasing ratios would result in different percentages of equity at the end of the glide path and performance of cumulative return. Exhibit 28 shows cumulative return at the end of the glide path decreases with higher decreasing ratios. It is not surprising, because equity provides overall better profits than fixed income. With higher decreasing ratios, the portfolio earns less from equity position, resulting in lower cumulative return. Exhibit 29 shows the positive relationship at the end of the glide path between the percentage of equity and the cumulative return. The original glidepath ends with the percentage of equity in 50%, and the cumulative return reaches 35.66 (see red dotted line in Exhibit 29). Apparently, all of the glide paths ending in approximately 50% equity holdings outperform the original glidepath. New glide paths do improve the returns.

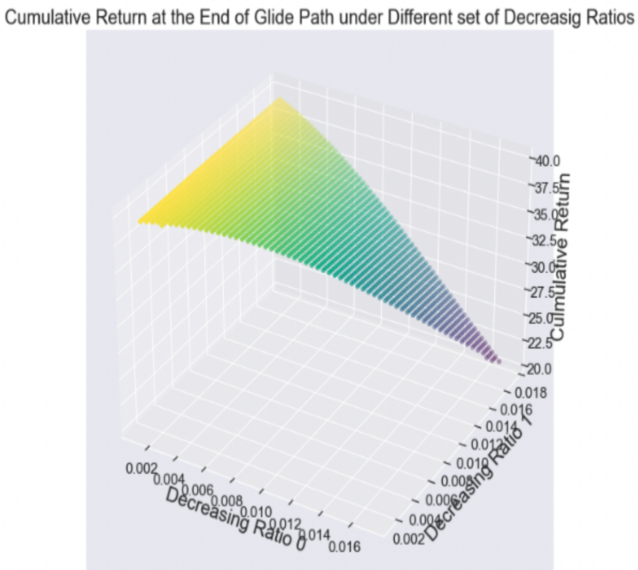


Exhibit 28. Cumulative return at the end of glide path under different set of decreasing ratios (Source: Author)

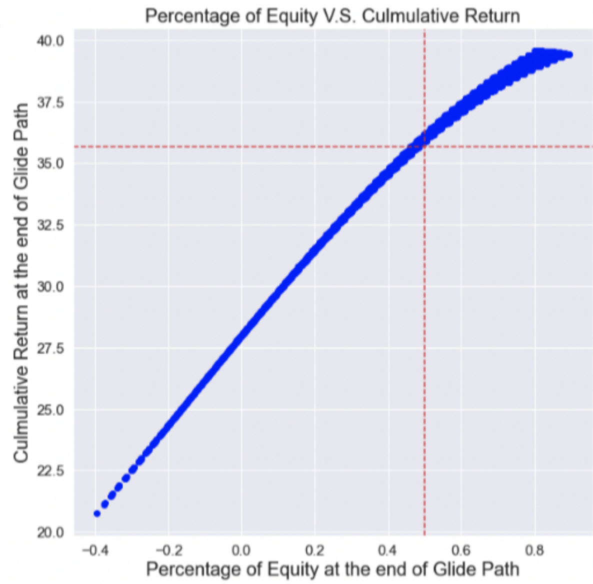


Exhibit 29. Percentage at Equity V.S. Cumulative return at the end of the glide path (Source: Author)

Now, we dive into details how the optimized glide path outperforms the original glidepath. We select one glide path ending with 50.8% equity holdings, which is more reasonable to compare with original glide path ending with 50%.

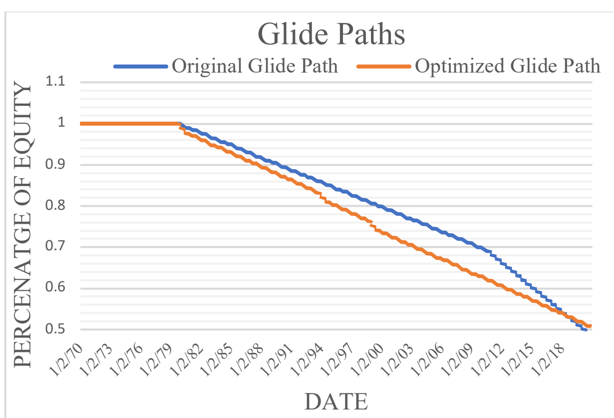


Exhibit 30. Original glide path V.S. optimized glide path (Source: Author)

Exhibit 31 shows that the new glide path optimizes the return. At the end of the glide path, the optimized glide path reaches 36.14, while the original glide path has only 35.66. Over 85.7% of time (from 1970 to 2020) optimized glide path outperforms original glidepath. Moreover, as could be seen in Exhibit 32, the optimized glide path provides less downside volatility than the original one. Optimized glide path, without a doubt, achieves a better Sharpe ratio of 59.8%, comparing to the original one of 57.7%.

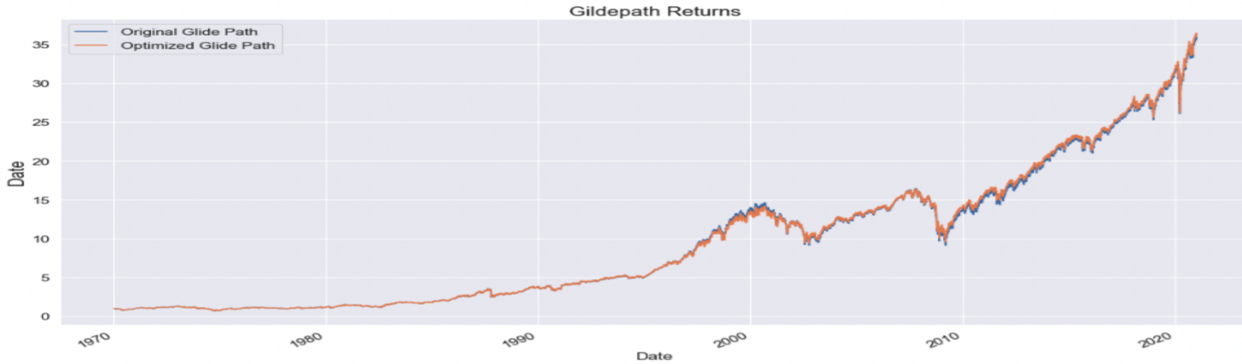


Exhibit 31. Cumulative returns in original glide path and optimized glide path (Source: Author)

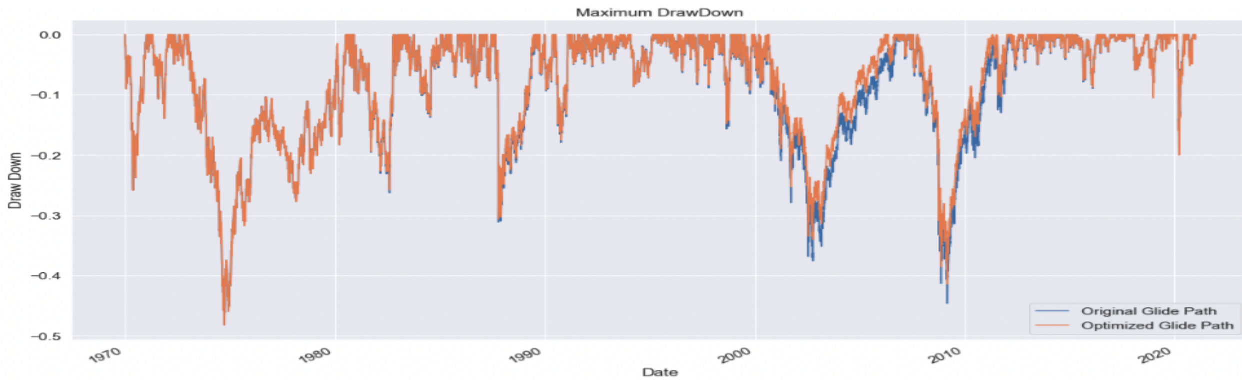


Exhibit 32. Maximum drawdown in original glide path and optimized glide path (Source: Author)

E. Discussion and Conclusion

In the original glide path, the proportion of equity positions and fixed income positions could be determined in advance. However, in this optimized glide path, we predict the market movement semi-annually and decide the corresponding decreasing rate later. That is, we won't know the whole glide path until last year. Chances are the percentage of decreasing ratio goes under zero (see Exhibit 33).

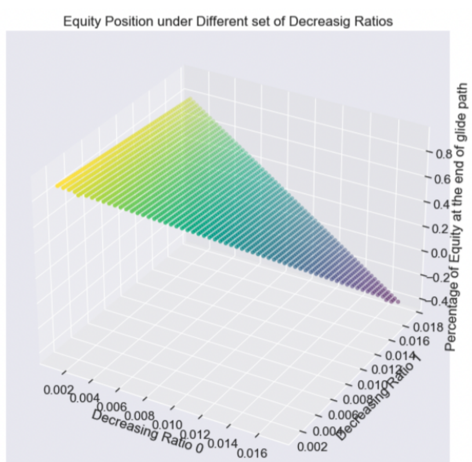


Exhibit 33. Equity positions under different sets of decreasing ratios (Source: Author)

Though we won't know the exact number of bull and bear market movements, they could be estimated based on the historical data. The market cycle is regular rather than promiscuous. As a result, we could conduct simulations with different sets of decreasing ratios to make sure the proportion of equity and fixed income positions would always be in an appropriate range in most cases. To address the worst case, a correction mechanism could be designed.

In conclusion, there are possibilities that glide path informed by predicted market movement could improve the return and volatility of time-horizon-based glide path. Using modern artificial-intelligence and machine-learning frameworks, we could conduct the market environment analysis and build an ideal asset allocation for a target date strategy—one that seeks the higher expected return or lower risk (i.e., volatility) over time—depending on what market environment exists and evolves. Therefore, a glide path informed by market environment analysis may help target date strategies be more resilient in dynamic and changing markets.

SECTION IV: TDF Glide Path Optimization with Reinforcement Learning

Besides supervised and unsupervised learning, reinforcement learning is also an exciting area of research for financial problems. Reinforcement learning allows us to solve complex dynamic optimization problems in an almost model-free way, by optimizing its reward under environment and actor spaces (Kolm & Ritter, 2020). In this section, we apply reinforcement learning (Policy Gradient) for constructing optimal TDF glide paths based on market scenarios. First of all, we reviewed the related research and the basic concepts of reinforcement learning. Secondly, the reinforcement learning model of this research is constructed. Thirdly, we implement Monte Carlo simulation for evaluating our model, compared with the benchmark of the Vanguard glide path. Finally, we conclude this section with our training and simulating results, and the future improvements of our model.

A. Literature Review

1. Relevant Research

Reinforcement learning has many applications in finance, including hedging, pricing, asset allocation, asset liability management, etc. For instance, Buehler et al. (2019), Kolm & Ritter (2020), and Cao (2021) applied reinforcement learning to derivatives hedging; Fontoura et al. (2019) implemented Deep Deterministic Policy Gradient in Asset-Liability Management, while Abe et al. (2010) develops a framework on Markov Decision Process to solve the debt collections problem.

Asset allocation applications, where the goal of reinforcement learning is to obtain the weights of the assets that maximize the rewards in a given state of the market considering risk and transaction costs, are more relevant to our glide path optimization problem. Moody & Saffell (2001) use Recurrent Reinforcement Learning (RRL) on optimizing portfolios. In their research, risk-adjusted returns such as maximum drawdowns and differential Sharpe ratio are included as a part of their reward function. Almahdi & Yang (2017) also adopt RRL in their portfolio optimization research. They show that the maximum drawdown risk-based objective function yields superior return performance compared to previously proposed RRL objective functions (i.e., the Sharpe ratio and the Sterling ratio).

Besides maximizing the return-based reward function, a Q-learning approach is used to maximize the expected utility of consumption in Weissensteiner (2009), to optimize consumption and asset allocation decisions. The objective function in Irlam (2020) is also designed to maximize the expected utility of consumption with Proximal Policy Optimization (PPO). Using Monte Carlo simulation, Irlam (2020) simulates stock prices, bond prices, and client's financial scenarios under different relative risk aversion (RRA) assumptions for lifetime portfolio selection. However, although Irlam (2020) provides the performance comparison between the traditional glide path and the reinforcement learning result under Monte Carlo simulation, they do not guarantee the shape of their lifetime asset allocation conforms to the basic rule of the traditional glide path.

Lastly, Wang et al. (2021) optimize their asset scoring unit and market scoring unit with policy gradient to construct a risk-return balanced portfolio. Also based on policy gradient, i Alonso & Srivastava (2020) provides a simple deep reinforcement learning method with portfolio vector memory to optimize portfolio within 24 US equity securities. Convolutional neural network (CNN), recurrent neural network (RNN), and Long Short Term Memory Neural Network (LSTM) were included as a component of their agent network. The LSTM model outperforms other models in terms of total returns, Sharpe ratio, and maximum drawdown.

In this research, following the approach by i Alonso & Srivastava (2020), we implement the deep policy gradient with the LSTM agent network. To take the portfolio volatility into account, returns and downside risk are both adopted in our reward function. Moreover, to deal with the glide path problem discussed earlier, a special glide path term is included in our reward function to make sure the decreasing asset allocation path.

2. A Brief Introduction to Reinforcement Learning

To make readers comprehend our model more easily, we briefly introduce the basic concepts of the reinforcement learning algorithm here. As we mentioned above, reinforcement learning is a machine learning approach concerned with solving dynamic optimization problems in an almost model-free way by maximizing a reward function in state and action spaces (I Alonso & Srivastava, 2020). That is to say, reinforcement learning has three key components: Agent (Policy), Environment (State), and Reward. The goal of a reinforcement learning algorithm is to maximize the total reward, using the data generated by its agent, environment, and reward function. The environment of a reinforcement learning algorithm could be a game, a simulation, or a series of financial data which are usually undetermined (a black box). As shown in Exhibit 34, the agent receives states generated by the environment and transforms the received states into actions. After each action, the environment responds with another state and reward generated by the reward function (according to the action in the environment, for example, killing a monster in a game might gain positive rewards). Hence, the sum of each reward after each action until the episode end (being killed or win), the total reward, is the objective to be maximized. Different from supervised learning, target labels are not needed in reinforcement learning. And similar to supervised learning in which the goal is to minimize its loss function, the goal of reinforcement learning is to maximize its objective function (total reward), which can be considered as minimizing a negative loss.

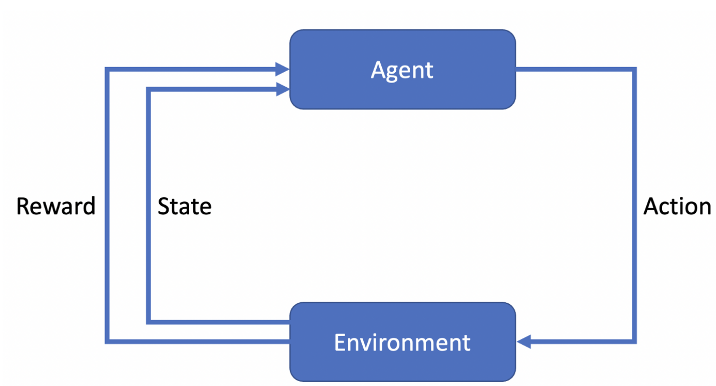


Exhibit 34. Reinforcement Learning Structure (Source: Toward Data Science)

B. Methodology

In this research, our goal is to construct optimal TDF glide paths based on given market scenarios. We follow the approach by i Alonso & Srivastava (2020) to construct the agent network and train the model with the Deep Policy Gradient algorithm. To design a new glide path for TDF under market scenarios, the input data of our model includes bond returns, stock returns, and the investor's age. The contribution of this research is that we propose a reward function for applying reinforcement learning to design TDF glide paths. Following is the concept of our environment, agent, reward function, and the training algorithm.

1. Environment

As mentioned above, we adopt market scenarios as our environment. We use 41 years (from 1981 to 2021) of daily bond returns of Fidelity Investment Grade Bond Fund, and stock index returns of the S&P 500 index as our training environment to train our reinforcement learning model as the simulation of TDF from 25-year-old to the target date of retirement (65-year-old). Furthermore, to make sure the input data of the agent network is in the same shape, we adopt the least yearly trading date of 41 years as our yearly trading date, which is 248 days during the data period.

2. Agent

The function of the agent in reinforcement learning is to transfer information given by the environment (states) into actions. In our case, we need to transfer the market scenarios, which are the returns of stock index and bond, into asset allocation strategies. In order to construct the agent that behaves differently under different ages of clients, we also add a feature that is composed of the client's current age as the input environment data. Similar to the agent network of i Alonso & Srivastava (2020), as shown in Exhibit 35, our agent network is composed of an LSTM layer and two dense layers. The output of our agent network is a 1*2 array, representing the weight we are going to invest in the equity market and the bond market next year. Note that the sum of two weights is equal to one (no cash position) in our case.

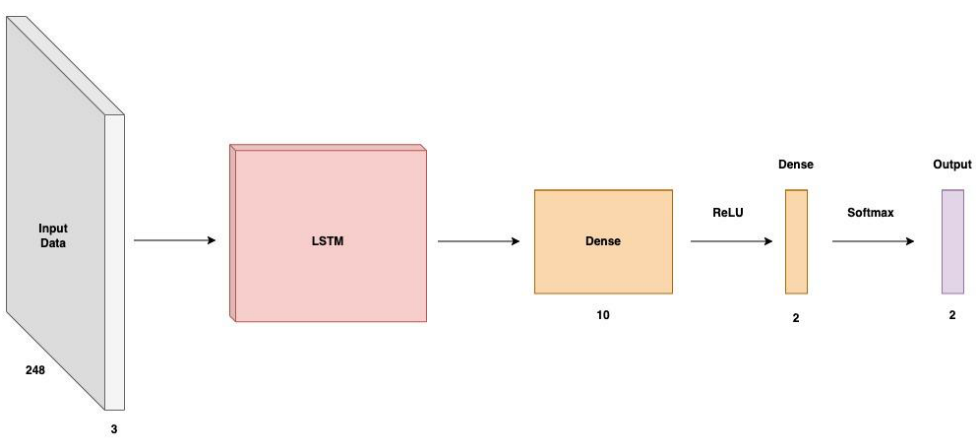


Exhibit 35. Structure of Agent Network (Source: Author)

3. Reward Function

Reward function plays the most important role in our research. Our reward function is composed of three components: return, risk, and the special term for designing a glide path. First, as proposed in previous papers, returns of assets are the most popular feature being used in asset allocation reinforcement learning research. Hence, the first term of our reward function is the cumulative return R_c^i of the year i (we have 40 years in an episode). Secondly, to consider the portfolio volatility, risk-adjusted return is also very famous in reinforcement learning reward. Considering previous research, we decided to adopt drawdown risk:

$$DD^i = \sqrt{\frac{1}{T}\sum_{t=1}^{T} min\,\{R_t\,, 0\}^2}$$

where R_t is the daily return at time t of the year i, and T equals 248 in our research. Lastly, we design a term for the reward function in order to make sure the asset allocation path of our training result conforms to the basic glide path rule: begin with 80% to 90% of equity position and remain 30% to 50% of equity position at the target date. The glide path term is

$$G^i = \sqrt{Age^i} \times W_S^i$$

where Age^i is the age of the investor at year i, and W_S^i is the weight of the equity position at year i. The term G^i increases when the investor's age increases, hence the algorithm will automatically decrease W_S^i to find the expected decreasing glide path. To conclude, our reward function at year i is:

$$Reward^i = R_c^i - \alpha DD^i - \beta G^i$$

where α and β are hyperparameters. Downside risk and the glide path term negatively impact the reward of the year. In our model, α is set as 0.01 and β is 0.0045. A larger β makes the model prefer a greater bond position to avoid the punishment from the glide path term, and a smaller β makes the model prefer a greater equity position for greater returns. Also, to make the model converge more easily to the expected glide path, we limit the minimum weight of the positions to 10%, and the maximum to 90%. Furthermore, Gaussian random is added to the weight to prevent the algorithm from sticking to a specific strategy.

4. Training Algorithm

Combining the environment, agent, and reward function presented above, here is our training algorithm for constructing optimal TDF glide paths. The first step of the algorithm is to set up the input data for the agent network, that is, the $State^0$ generated by the environment. The $State^0$ in the training process is the year 0 (age 24) stock return data and bond return data with an investor's age of 24. The data $State^0$ is the input data of the agent network, and the network will generate the weight of the equity position next year, W_S^i, and the weight of the bond position next year, W_B^i. After we get the asset allocation weights from agent network, we can calculate the reward using the $State^1$ (the stock returns and the bond returns of year 1 (age 25)) and the weights. The process continues until we reach the investor's retirement (age 65). Hence, at the end of this episode, we calculate the reward using $State^{40}$ and weights generated by $State^{39}$ and the agent network. Summing up all 40 rewards, we get the total reward of this episode which can be used to calculate gradients and update our agent network. Using Adam Optimizer to update the parameters in our agent network, this is our first training epoch. Repeat the steps multiple times, and the model for generating an optimal TDF glide path is prepared. To avoid overfitting and converging to an abnormal glide path, we do not set the training number too large. We only trained the model for approximately 60 epochs. Exhibit 36 shows our complete reinforcement learning training algorithm.

Algorithm 1 Reinforcement Learning Training Algorithm

while $n < TotalEpoch$ **do**
    Initialize Environment
    $i = 0$
    while $i < 40$ **do**
        $State^i = \{StockPrice^i, BondPrice^i, Age^i\}$
        $State^{i+1} = \{StockPrice^{i+1}, BondPrice^{i+1}, Age^{i+1}\}$
        Input $State^i$ into Agent Network
        Agent output W_S^i, W_B^i
        Input $State^{i+1}$ and W_S^i, W_B^i into Reward Function
        Calculate $Reward^{i+1} = R_C^{i+1} - \alpha DD^{i+1} - \beta G^{i+1}$
    end while
    Calculate $TotalReward = \sum_{i=1}^{40} Reward^i$
    Update Parameters in Agent Network
end while

Exhibit 36. Reinforcement Learning Training Algorithm (Source: Author)

C. Monte Carlo Simulation

To validate our model, we implement the Monte Carlo simulation on both the traditional benchmark glide path and the glide paths generated by our reinforcement learning model.

1. Monte Carlo Method

To implement the Monte Carlo simulation in our research, the simulation period and features of simulated data must be equal to our training data. Hence, our simulation period is also 41 years per simulation, and the features of simulations are also extracted from our training data, S&P 500 index, and Fidelity Investment Grade Bond Fund from 1981 to 2021. The mean of returns and covariance of equity and bond data are extracted to implement the Monte Carlo simulation. The simulation processes are similar to the algorithm presented in Exhibit 36. We first generate the 40-year equity and bond return data by the Monte Carlo method and input the data into our trained reinforcement learning model. A series of weights will be generated by our agent network, which composes our optimal glide path for the simulation. We then calculate the return and risk-adjusted returns that symbolize the performance of the simulation. We repeat the simulation 1,000 times and average the results of each simulation as the performance of our model.

2. Vanguard Glide Path
Vanguard is well-known for its variety of low-cost index mutual funds and ETFs and captured most investor money in its target-date funds in 2021 relative to other asset managers. According to Morningstar, retirement savers invested a net $55 billion in Vanguard's Target Retirement Funds last year. To choose a benchmark to compare with our model, the Vanguard glide path is selected as the traditional benchmark glide path. According to the report by Vanguard Research in 2015, the glide path for Vanguard TDF is designed beginning with a 90% equity position from 25-year-old to 40-year-old and decreasing to 50% at the target retirement date (65-year-old) as shown in Exhibit 37. The same 1,000 simulations are also executed on the Vanguard glide path.

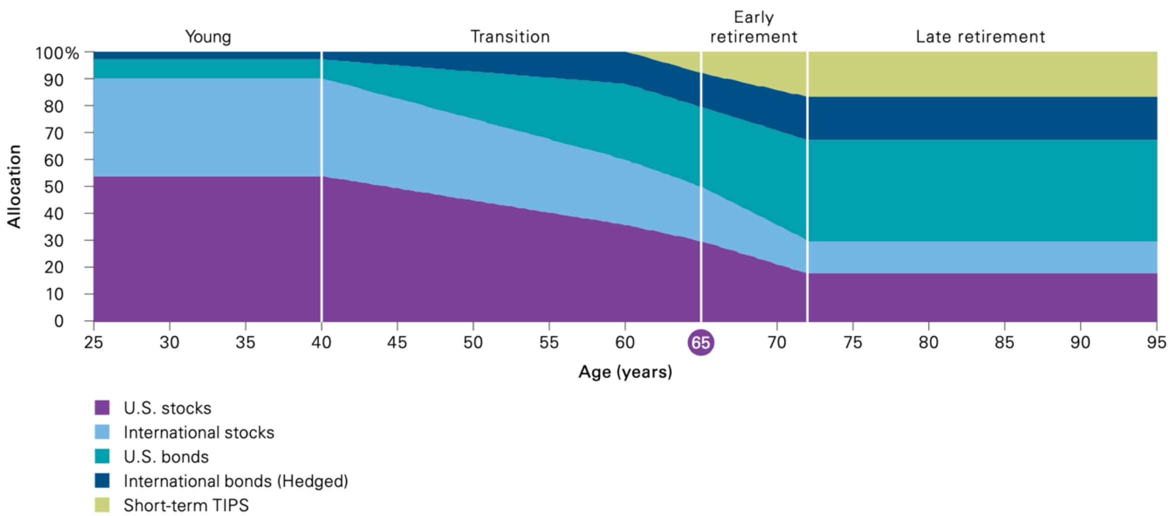


Exhibit 37. The Benchmark Vanguard Glide Path (Source: Vanguard Research)

D. Training and Simulation Results

1. Training Result
The initial output of the agent network is designed as uniformly distributed. Hence, in Exhibit 38, we can observe that the initial weights of the glide path are approximately 50% equity and 50% bond. After we trained the model for several epochs, the decreasing asset allocation path along with the increasing age of the investor demonstrates the effect of the glide path term in our reward function. At the end of the training process, an incurve appears between 25 to 35-year-old. In terms of rewards, the total reward increases from 2.93 to 3.12 during the training process. Exhibit 38 shows the evolution of the reinforcement learning glide path.

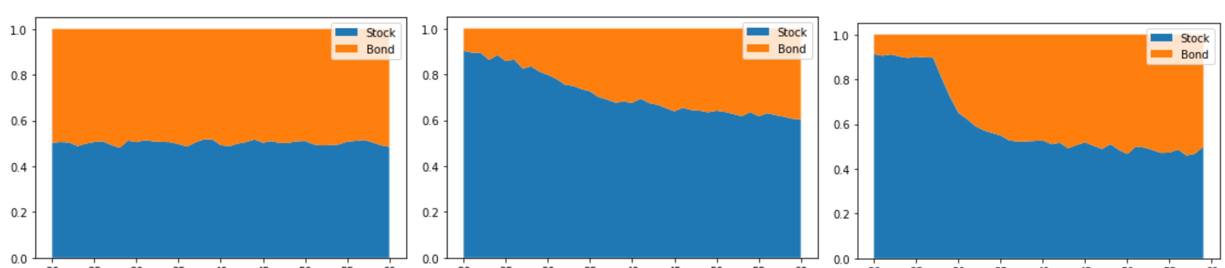


Exhibit 38. Evolutions of the Optimal Reinforcement Learning Glide Path (Source: Author)

2. Monte Carlo Result

After each simulation of the Vanguard glide path and the optimal glide path, we calculate the Sharpe ratio, Sortino ratio, Maximum Drawdown, Calmar ratio, and average yearly return. After 1,000 simulations, we calculate the arithmetic average of the ratios. As shown in Exhibit 39 the result shows that our model outperforms the Vanguard glide path in terms of all risk-adjusted returns, and slightly underperforms in terms of average yearly return.

	Sharpe Ratio	Sortino Ratio	Maximum Drawdown	Calmar Ratio	Average Yearly Return
Vanguard Glide Path	0.589	1.654	-0.263	0.466	**0.101**
Reinforcement Learning	**0.679**	**1.963**	**-0.200**	**0.618**	0.095

Exhibit 39. Results of Monte Carlo Simulations (Source: Author)

E. Discussion and Conclusion

1. Result Discussion

First of all, although the risk-adjusted returns of our model outperform the Vanguard glide path, the average return is lower. One reason for the lower average return is the shape of our final model. Compared to the Vanguard glide path, our model holds more bond position when the investor's age increases. While the Vanguard glide path holds approximately 80% equity position at 45, our model only holds 50% of the equity position. Hence, even if the initial and final positions of both models are close, the shape of our model has a significant disadvantage concerning returns.

Secondly, as we mentioned in the previous paragraph, our model has an unusual concave appearance between 25 to 35-year-old. During our training processes, the drop in investment ratio appears more rapidly as we train the model with more epochs. In the end, if we train the model for 1,000 epochs, the model will converge to the extreme value we set (10% and 90%) with a single cliff that appears depending on the β we set. The larger the β is, the earlier the cliff appears. Changes in hyperparameters and the reward function do not influence the appearance of the cliff, as the training times increase. Hence, our solution is to not train the model for too many epochs to avoid the problem.

Moreover, unlike most asset allocation research that focuses on short-term and mid-term investments, TDF requires at least tens of years of data for one training episode. To make sure the veracity of the training data, we can only train the model with the same data set again and again. It may easily cause the overfitting of our model. Although we can assume that the future price paths have the same features as our training data and apply the Monte Carlo simulation for model testing, it does not solve the problem of lacking training data.

Lastly, the model is designed to change its asset allocation strategy based on different market scenarios. Whereas the result of the Monte Carlo simulation shows that the optimal glide paths for every simulation are close. Considering the varying simulated market scenarios, the only fixed term in the reward function is the increasing punishment from the glide path term. Hence, weakening the influence of the glide path term may be necessary to increase the influence of market scenarios.

2. Future improvements

Besides the points discussed above, there are some directions for potential improvements to this model. First, in our research, we only consider the returns of equity and bond. To provide the model with more market information, other features like trading volume could be added to the model once the problem of lacking data is addressed. Also, we only consider an equity index and a bond in our asset allocation of glide paths. Whereas most TDF providers include multiple assets in a single TDF. Hence, considering multiple assets will not only make the research closer to reality but also bring the power of reinforcement learning into play. Finally, again, although our reward function provides a way for applying reinforcement learning to designing a TDF glide path, there are many problems that occur while training the model. Hence, a redesigned reward function that could deal with the above problems will be a significant improvement.

3. Conclusion

In this section, we implemented Deep Policy Gradient to design optimal TDF glide paths under market scenarios and the investor's age. A new reward function was developed to make sure the decreasing asset allocation paths conform to the basic rules of the traditional TDF glide path. Using Monte Carlo simulation, our optimal glide path model outperformed the Vanguard glide path in terms of all risk-adjusted returns, but slightly underperformed in terms of the average yearly return.

Reference

John B. Shoven and Daniel B. Walton. “An Analysis of the Performance of Target Date Funds.” The Journal of Retirement, 2021.

Edwin J. Elton, Martin J. Gruber, Andre de Souza and Christopher R. Blake “Target Date Funds: Characteristics and Performance” 2022.

Ilia Lanski, Raj Paramaguru, Wesley Phoa, Yung Wang and P. Brett Hammond. “Using a Life Cycle Model to Design a Target Date Glidepath.” The Journal of Portfolio Management, 2022.

Irlam, G., “Lifetime Portfolio Selection: Using Machine Learning.” 2020.

Weissensteiner. A., “A Q-Learning Approach to Derive Optimal Consumption and Investment Strategies.” IEEE transactions on neural networks, 2009.

Fidelity. “Creating a Resilient Glide Path for a Target Date Strategy” 2017.

ICI Research. “Target Date Funds: Evidence Points to Growing Popularity and Appropriate Use by 401(k) Plan Participants”, 2021

Mamed Caki. “Economic Regimes Identification Using Machine Learning Technics” 2018.

1290 Funds. “Finding Your Target Retirement Date” 2021.

Julianna Paterra. “To or Through? Evaluating TDF Glide Paths”, 2019.

Morningstar. “2022 Target-Date Strategy Landscape”, 2022.

Saud Almahdi, Steve Y. Yang, “An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown” 2017.

Dr Miquel Noguer i Alonso and Sonam Srivastava, “Deep Reinforcement Learning for Asset Allocation in US Equities” 2020.