Off Target: On the Underperformance of Target-Date Funds

June 3, 2023

Abstract

Target-date funds (TDFs) are popular retirement savings vehicles that provide diversification and dynamic asset allocation to investors. While TDFs hold trillions of dollars of assets, institutional features limit competition and make comparisons among funds difficult. We develop a novel means of benchmarking TDF performance that provides comparability and facilitates competition. Our benchmarking method utilizes mimicking portfolios of cost-efficient funds which we call *Replicating Funds (RFs)*. RFs substantially outperform TDFs and exhibit low tracking error. In 2019, excess costs to investors totaled $8.6 billion.

Keywords: Target-Date Funds, TDFs, Exchange-Traded Funds, ETFs, Financial Planning, Benchmarking

JEL Classification: G10, G11, G23, G50, G51, G53

1 Introduction

> In addition to reforming the laws governing traditional private pensions, the bill [Pension Protection Act of 2006] I signed today also contains provisions to help workers who save for retirement through defined contribution plans like IRAs and 401(k)s. These savings plans are helping Americans build a society of ownership and financial independence.
>
> —President George W. Bush, August 17, 2006

The Pension Protection Act of 2006 was established, in part, to reform defined contribution (DC) plans and assist American workers in saving for retirement.[1] A key feature of the legislation established safe harbor investments, also known as Qualified Default Investment Alternatives, to auto-enroll employees into. The most prominent of these safe harbor investments are Target-Date Funds (TDFs), which automatically adjust investors' asset allocations along a "glide path" to meet their needs at some future date (e.g., retirement). This feature of TDFs is attractive because it provides a convenient and diversified retirement savings strategy in a single fund; an investor simply chooses their retirement date (e.g., 2050). Since the Pension Protection Act of 2006, assets invested in TDFs have grown tremendously. From 2006 to 2020, the total assets in TDFs has expanded from $114 billion to almost $1.6 trillion (Investment Company Institute, 2021a). In 2019, according to Cerulli Associates, 39% of defined-contribution assets were held in TDFs, and nearly 60% of all ongoing 401(k) contributions were directed to TDFs.[2] Not only are TDFs large in terms of assets under management, TDFs are also popular on a per capita basis – 80% of all 401(k) participants are invested in TDFs.[3]

While TDFs are hugely popular and an important financial innovation to assist households in preparing for retirement, the TDF market is opaque. The opaqueness of the TDF market is due to its complexity; TDFs are a type of fund-of-funds in which the fund sponsor chooses anywhere

[1] Employer-sponsored retirement plans are divided into two major categories: defined-benefit (DB) plans and defined-contribution (DC) plans. DB plans, more commonly known as a pension, are retirement accounts sponsored and managed by an employer. As such, the employer's liability extends beyond employment and into retirement. Conversely, DC plans, for example a 401(k), are retirement accounts sponsored by an employer during a worker's career. In DC plans, the employee manages the retirement assets during employment and throughout retirement.

[2] See www.ft.com/content/3d85ebd4-65bb-4bcf-b687-4acabfa41420.

[3] See www.cnbc.com/2022/03/06/target-date-retirement-funds-work-up-to-a-point-when-to-reconsider.html.

between a handful of funds to over 35 funds to hold, the funds held are both actively and passively managed, and investors often pay multiple layers of fees (fees to the fund sponsor at the TDF level and more fees on the underlying funds). Fund sponsors also have discretion over glide paths (e.g., risk appetites, rebalancing frequencies, "to" versus "through" retirement mandates, and asset class inclusion/exclusion), adding additional complexity and complicating comparisons among TDFs. The complexity and opacity of TDFs calls for benchmarking methods to assess TDF performance and aid comparisons across TDFs.

We provide a novel benchmarking method for TDFs. Rather than relying on traditional factor-based models, we effectively benchmark each TDF to itself. Specifically, we match each TDF's quarterly holdings to highly-correlated exchange-traded funds (ETFs). We then rebuild each TDF using the matched ETFs to create *Replicating Funds (RFs)*. Our RFs can be thought of as investors' best outside option given a specific glide path. By construction, our benchmark is agnostic to the quality of a glide path, as glide path quality is subjective and depends on investors' preferences and beliefs. We measure performance by comparing TDFs' returns to their respective RFs' returns. Our main finding is that TDFs substantially underperform their RF counterparts with little tracking error. In our sample, the average TDF underperformed its RF by 1.03% per year.[4] Higher management fees account for 55% of the underperformance, while 40% is due to poor performance by TDFs' underlying funds. These results highlight that TDFs charge significantly higher fees than low cost alternatives and that TDF sponsors select poorly performing funds that underperform those low cost alternatives, *before fees*.

High fees and poor fund performance have imposed significant excess costs on investors. To provide a ball park estimate of these costs, we define **excess** costs as the return spread between a TDF and its matched RF minus a reasonable cost to operate a TDF, which we assume is 10 bps.[5] The 10 bps haircut accounts for the costs associated with actively managing a TDF's glide path, paying managers to choose which funds to include, and reflecting that a TDF's overhead costs may be relatively more expensive than that of an ETF. Figure 1 shows the cumulative excess costs to

[4]Throughout the paper, we calculate annual values by multiplying the monthly estimates by 12.

[5]10 bps exceeds the gap between Vanguard's TDFs and the aggregate cost of the least expensive share classes of their TDFs' underlying holdings.

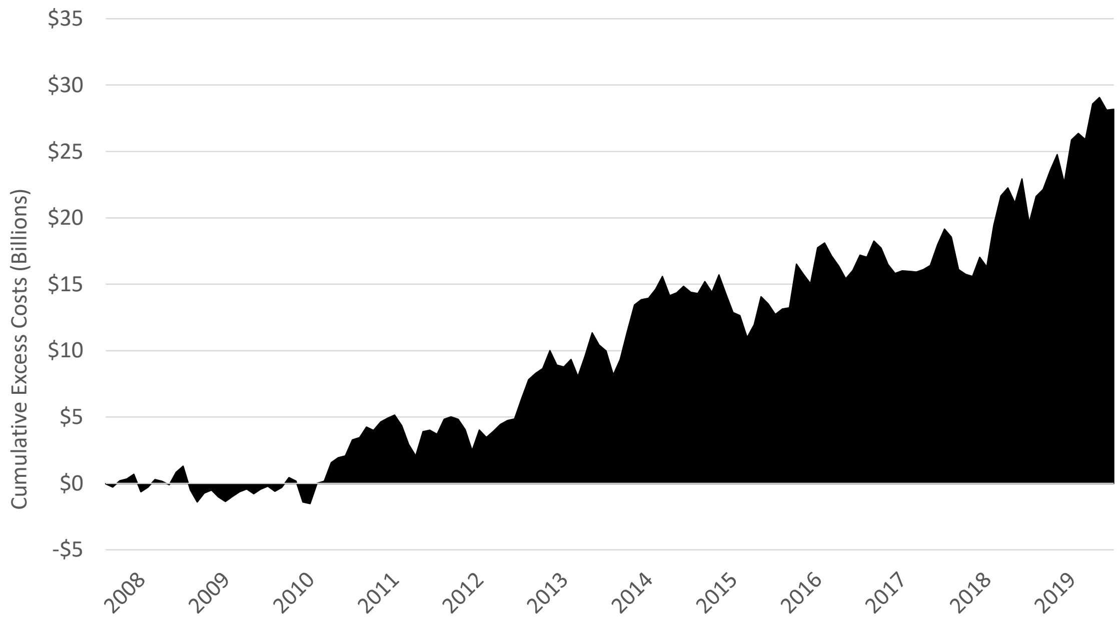

Figure 1: Cumulative Excess Costs. This figure plots the cumulative excess cost of all TDFs in our sample. We define excess costs as *Spread* minus an annual operational cost of 10 bps. 10 bps is likely an upper bound on operational costs, as it exceeds the gap between Vanguard's TDFs and the aggregate cost of the least expensive share classes of their TDFs' underlying holdings.

investors. From 2008 to 2019, investors have incurred $28.2 billion in excess costs, with $8.6 billion in 2019 alone.[6]

Our benchmarking method is intended to bring transparency to TDF performance and to reverse, or at least slow, the trend of growing excess costs to investors. While many target date indexes exist, in 2017 Morningstar stated that "... published benchmarks ... are all but useless in helping stakeholders assess the performance of the target maturity funds."[7] Not much has changed.[8] Our return spread measure allows individual investors – and people responsible for setting retirement plan menus – to assess how efficiently a TDF implements its glide path. Return spread is a transparent metric that stakeholders can use to hold TDF providers accountable for their performance.

[6]Our excess cost estimates are a lower bound as our data excludes TDFs held in collective investment trusts (CITs). Morningstar reported that in 2020, CIT target-date assets totaled $1.18 trillion. See https://www.pionline.com/defined-contribution/cits-become-preferred-target-date-series-choice.

[7]See www.cnbc.com/2022/03/06/target-date-retirement-funds-work-up-to-a-point-when-to-reconsider.html.

[8]Numerous "best of" lists offer TDF recommendations without clear, consistent and objective metrics. See, for examples, www.kiplinger.com/investing/mutual-funds/601381/best-target-date-fund-families and www.forbes.com/advisor/retirement/best-target-date-funds/.

Turning to the details of our analysis, we create RFs by matching TDFs' underlying holdings (mutual funds) to correlated ETFs. Our main analysis matches each mutual fund to one of 50 Vanguard ETFs based on the highest return correlation. These Vanguard ETFs are characterized by low fees and high liquidity and cover a wide range of asset classes and geographical regions. Focusing on a limited set of low-cost ETFs ensures that our strategies can be reasonably implemented. To create a RF, we weight each matched ETF based on the TDF's underlying holdings' portfolio weights.

To assess RFs' abilities to mimic their TDF counterparts, we perform two tests. First, we regress monthly RF returns on TDF returns. The regression results highlight several of our key findings: RFs earn excess returns relative to TDFs and RFs exhibit little tracking error. The regression intercept is 5 bps per month (t-statistic of 6.36), which amounts to excess returns of 0.60% per year. Additionally, across our full sample the adjusted R^2 is 96% and the RF's annual excess return per unit of risk (RMSE) is 0.28, consistent with low tracking error.

The second test is non-parametric and analyzes the difference in monthly RF returns and TDF returns, which we call *Spread*. In our full sample, average monthly *Spread* is 9 bps (1.03% per year). Using *Spread*, we are able to gain additional insight into the drivers of RF outperformance that are obfuscated in regression analysis. Specifically, we decompose *Spread* into three individual components: *Timing*, *Active*, and *Fee Gap*. *Timing* measures TDF managers' abilities to adjust asset allocations within a quarter (e.g., from equities to bonds), *Active* measures the relative performance of underlying holdings versus their matched ETFs (i.e., active management), and *Fee Gap* measures the difference between TDFs' total fees and matched RFs' underlying ETFs' fees. The decomposition shows that *Timing* accounts for the smallest portion at 0.4 bps of RF outperformance per month (5% of the average monthly spread), *Active* accounts for 3.4 bps (40% of the average monthly spread), and *Fee Gap* accounts for 4.8 bps, which is the the majority (55%) of the average monthly spread. An annual decomposition shows that *Timing* is relatively stable while *Active* is fairly noisy, adding to RF outperformance in some years and subtracting from it in others. *Fee Gap* exhibits considerable persistence and follows the overall trend in mutual fund fees from 2008-2019; *Fee Gap* begins around 7 bps per month and declines monotonically to 4 bps per month.

Having documented the excess costs present in the TDF industry, we also show that there is considerable variation in performance across TDFs and across fund families.[9] For example, the average difference between the 10th and 90th percentiles of monthly *Spread* is 43 bps (5.18% per year). As another example, across fund families, the average monthly *Spread* for Vanguard is 0.1 bps per month, while for Fidelity the number is considerably larger at 6.9 bps and is larger still for American Funds at 10.9 bps. Given the variation in RFs' performance relative to their respective TDFs, we test whether or not investors – or retirement plan sponsors – take *Spread* into account when choosing TDFs. We find that fund flows are not consistently sensitive to *Spread.* The large dispersion in TDF performance across funds and fund families, the apparent lack of awareness of these differences, and the lack of competitive forces to discipline poorly performing funds, all suggest that our benchmark can help investors and aid in the allocation of capital to better performing funds.

While the main purpose of our study is to properly benchmark TDF performance, we conclude our analysis with normative recommendations allowing individuals to compete against TDF providers directly. Specifically, we provide boilerplate portfolios for investors who would otherwise choose a TDF simply based on its target-date (e.g., a 2030 fund). To do so, we aggregate the holdings of our RFs within vintage (e.g., the replicated holdings of *all* 2030 funds) and we call these portfolios *Passive Replications of Funds (PROFs).* We acknowledge that a drawback of our boilerplate portfolio recommendations is that a specific TDF (or its associated RF) may better fit a given investor based on her risk profile. Nevertheless, there is little evidence that investors have access to the "right" TDFs: Balduzzi and Reuter (2019) shows that it is unlikely that plan sponsors choose TDFs for their menus to offset participants' employment risk or overall risk preferences. As such, our PROF recommendations likely serve as an improvement over current options. Moreover, as investors may be overwhelmed by managing positions in 50 Vanguard ETFs, we construct consolidated PROFs that use only 20, 13, or 6 ETF positions.

PROFs outperform aggregate TDF portfolios and provide investors with practical alternatives

[9] Balduzzi and Reuter (2019) finds substantial heterogeneity in TDF returns and risk exposures.

to existing TDFs. Over our sample, PROFs outperform aggregate TDF portfolios by between 0.34% and 0.50% per year while limiting excess risk. When compared to individual TDFs, PROFs outperform between 87% and 91% of TDFs over our full sample, suggesting that most investors would be better off using our simple PROFs. This is particularly impressive because TDF returns are highly variable due to the different glide paths (and resulting asset allocations) used across managers.

For both individual investors and retirement plan sponsors, existing TDF benchmarks fail to provide accountability or aid in comparisons across funds. We introduce a new benchmarking method that assesses how well TDFs implement their given glide paths, enhancing accountability and comparability. Investors should select among TDFs with the lowest return spreads that fit their desired glide path profile. Our hope is that our benchmark can add competitive pressure to the TDF market, leading to lower fees and better fund choices for investors.

2 Related Literature

Our construction of RFs as benchmarks and feasible investment alternatives complements several literatures. First, several studies use low-cost index funds and, more recently, ETFs as benchmarks. Malkiel (1995) and Sharpe (1992) are early examples using stock indexes to benchmark active mutual funds. More recently, Berk and Van Binsbergen (2015) benchmarks active mutual funds to a set of 11 Vanguard index mutual funds. Similar to our study, Moraes, Cavalcante-Filho, and De-Losso (2021) uses ETFs to benchmark active mutual funds. While that study uses ETF benchmarks to assess funds' skill, we use ETF benchmarks to compare TDF performance to a relevant outside option. Our benchmarking methodology may be applied to fund-of-funds more generally, and in particular, to retirement-based funds such as balanced funds.

Benchmarking TDFs using their actual holdings distinguishes our paper from a number of studies that evaluate TDFs using factor models (Massa, Moussawi, & Simonov, 2022; Mao & Wong, 2022; Shoven & Walton, 2021).[10] While using factor models has lower data requirements,

[10] Parker, Schoar, and Sun (2020) also uses holdings data to assess TDFs' asset exposures. Rather than evaluating TDF performance, their analysis examines the impact of TDFs' aggregate rebalancing on stock returns.

that is not without costs. Since TDFs choose different asset classes to include in their portfolios (e.g., some TDFs include an allocation to real estate while others do not), it is unclear what type of risk model should be used. Furthermore, factor loadings are estimated with considerable noise, making factor-based benchmarks more volatile. In general, factor models are used to understand asset pricing when the assets cannot easily be decomposed into their component parts. For example, a firm cannot easily be valued based on the assets in place, growth options and intangible assets. However, when assets can be emulated by mimicking portfolios with nearly identical cash flows, those portfolios provide a better benchmark to evaluate the mispricing of assets (in fact, this notion underpins the textbook definition of arbitrage). In our setting, we show that such an exercise is possible, and thus we provide better benchmarks to employ for a large class of investment vehicles that are becoming increasingly prominent every day.

Second, our work relates to the literature on TDF performance evaluation. Balduzzi and Reuter (2019) shows large differences in realized returns among TDFs and attributes this dispersion to risk-taking by fund families with low market share. Massa et al. (2022) finds that TDFs experience lower performance due to investing in high-fee affiliated funds, particularly in later retirement date TDFs that have lower investor attention. We show that, even among the largest fund families, TDFs display large differences in performance after controlling for risk-taking through the glide path (via the RF benchmarks). Elton, Gruber, de Souza, and Blake (2015) finds that most TDFs invest in low-cost share classes of their underlying funds, so that investing in a TDF with its own fund-of-funds fee is only slightly more expensive than buying the underlying funds directly. We make a different point — investors can do better than the underlying funds themselves by replicating TDF portfolios with low-cost ETFs. From that perspective, TDFs are much more expensive than investors' outside options. Elton et al. (2015) also provides an alternative TDF strategy for investors, benchmarking each TDF to a static portfolio that mimics the fund's initial holdings using five index funds. Our PROFs utilize the aggregate TDF industry and low-cost ETFs to provide a dynamic, low-cost strategy that is easily implemented.

Third, our work relates to the extensive literature analyzing both excessive fees in the mutual fund industry (Cooper, Halling, & Yang, 2021) and limited investor financial literacy. Notably, Barber, Odean, and Zheng (2005) shows that investors are more sensitive to salient fees and largely ignore the fees that are less transparent. Our analysis complements the findings in Barber et al. (2005) by showing that investors do not respond to the underperformance of TDFs relative to RFs, which is less salient. Capon, Fitzsimons, and Prince (1996) and Alexander, Jones, and Nigro (1998) show that individual investors exhibit a limited understanding when choosing funds. In a similar spirit, Madrian and Shea (2001) shows that employees who are automatically enrolled in a default allocation retain that default. The study shows that retaining the default is a result of participants exhibiting inertia and/or interpreting the default as investment advice (see also Thaler and Benartzi (2004)). In this light, our finding that many investors are paying too much for TDFs may be an artifact of investors' limited financial literacy and interpreting TDF choices as financial advice.

3 Data and Methodology

3.1 Target-Date Fund Sample

We obtain data on mutual fund names, monthly returns, quarterly characteristics (e.g., expense ratios), and quarterly holdings from the Center for Research in Security Prices (CRSP) Survivor-Bias-Free U.S. Mutual Fund database. We flag mutual funds as TDFs if their names contain one of the following strings: *2000*, *2005*, *2010*, *2015*, *2020*, *2025*, *2030*, *2035*, *2040*, *2045*, *2050*, *2055*, *2060*, *2065*. As most TDFs have several share classes, we aggregate returns and expense ratios across share classes by weighting each share class's return and expense ratio by its end-of-month total net assets (TNA).[11] To augment the CRSP mutual fund quarterly holdings reports, we also obtain holdings data using the Thomson Reuters Mutual Fund Ownership (s12) database. The Thomson Reuters' data is utilized when CRSP does not report holdings for a given TDF. Both the CRSP mutual fund quarterly holdings reports and Thomson Reuters' s12 data provide quarterly snapshots of each TDF's holdings; that is, the assets held (typically mutual funds), their quantities,

[11] For example, in 2019 the Fidelity Freedom 2030 Fund had three share classes with three different fee levels: basic class shares (ticker "FFFEX", fee 0.69%), the K class shares (ticker "FSNQX", fee 0.59%), and the K6 class shares (ticker "FGTKX", fee 0.47%). As a result, our analysis overstates or understates the costs for some shares classes.

and their prices. We link the Thomson Reuters holdings data to CRSP data using the Wharton Research Data Services (WRDS) Mutual Fund Links (MFLINKs) database. A detailed discussion of our data set construction is outlined in the Data Appendix. Our sample begins in January 2008 and ends in December 2019.[12]

Panel A of Table 1 displays annual summary statistics for our sample. Between 2008 and 2019, the total net assets of TDFs in our sample grew from $145 billion to $1,299 billion. In addition to the massive asset growth, the number of funds grew steadily from 148 to 569. Total fees, composed of the TDFs' fund-of-funds fees and the underlying holdings' fees, decreased from 0.98% (0.66%) to 0.52% (0.40%) on an equal-weighted (asset-weighted) basis. Furthermore, the percentage of index funds held by TDFs steadily increased as well.

The TDF industry is highly concentrated. While our sample covers TDFs offered by 62 fund families, the vast majority of those funds are sponsored by one of six families: Vanguard (38% of 2019 TDF assets), Fidelity (21%), T. Rowe Price (13%), American Funds (12%), TIAA-CREF (5%), and JP Morgan (4%). The remaining 56 "Other" fund families collectively control just 9% of TDF assets. Panel B of Table 1 provides 2019 summary statistics for the six largest TDF families and the group of Other families. Notably, Vanguard has the lowest total fees (between 0.11% and 0.12%), the fewest holdings (approximately 4 funds in each TDF), and the highest percentage of index funds (100%). The other sponsors have consistently higher fees (between 0.23% and 0.66%), hold more funds (between 11 and 25 funds) and hold fewer index funds (between 0% to 49%). In the Internet Appendix, Table IA6 reports summary statistics by fund vintage and shows that 2030 funds are the largest by AUM.

Table 1 shows that TDFs' total fees can be broken down into fund-of-funds fees and underlying holdings' fees. Fund-of-funds fees are charged at the TDF level and can be thought of as the cost of adjusting asset allocations among the underlying funds to follow a TDF's glide path.[13] The underlying funds are almost exclusively mutual funds themselves, with their own management fees

[12] We have limited data on some TDFs from 2006 and 2007 but elect to begin our sample period in 2008 when the TDF industry was more mature and TDF holdings were more consistently reported. Including data from 2006 and 2007 introduces additional noise into the empirical analysis but leaves the results qualitatively unchanged.

[13] Note, while we measure the performance impact of TDF managers' intra-quarter rebalancing later in the analysis, our study is agnostic regarding deviations from funds' stated glide paths. See, instead, Elton et al. (2015) which analyzes TDF managers' glide path deviations and finds that performance is negatively impacted by such deviations.

that are paid from the TDF assets and thereby passed through to investors.[14] Typically, underlying holdings' fees make up 60-90% of total fees, with the remainder attributable to fund-of-funds fees. There are several notable exceptions. First, Vanguard does not charge a fund-of-funds fee. Because Vanguard is the largest fund family by assets, annual asset-weighted fund-of-funds fees are much lower than their equal-weighted counterparts. Second, some fund families have moved to charging only fund-of-funds fees. For example, in 2017 Fidelity began waiving underlying holdings' fees in some of their TDFs, and instead adopted flat (and higher) fund-of-funds fees.[15]

In addition to setting fund-of-funds fees, fund families control total TDF fees by choosing the underlying mutual funds. Fund families do not hold just any underlying mutual funds in their TDFs; at the end of our sample in December 2019, there are 8,502 unique TDF-holding pairs and 6,615 (77.8%) of these TDF-holdings are in the same fund family as the TDF itself. On an asset-weighted basis, 98.6% of TDF-holdings belong to the same fund family. In other words, TDFs almost exclusively hold mutual funds in their own families, allowing fund sponsors to collect fees on both the TDF itself and on the mutual funds held by the TDF. Moreover, given that mutual funds often have multiple share classes charging different levels of fees, some TDFs elect to hold the relatively *more* expensive share classes. Even Vanguard, typically viewed as the poster child of low-fee financial services, is not immune to this practice. Specifically, Vanguard funds have a unique feature that their ETFs are a share class of its index mutual funds.[16] As such, Vanguard ETF shares are claims to the very same pool of assets as the Vanguard index mutual funds. Despite this feature, Vanguard has a tendency to hold the relatively more expensive mutual fund shares in their TDFs rather than the less expensive ETF shares.[17]

[14]Technically, TDFs are hybrid vehicles that may hold both individual securities and positions in other funds. However, in practice, TDFs are essentially funds-of-funds as 97% percent of TDF assets are other mutual funds (Investment Company Institute, 2020b).

[15]See https://www.investmentnews.com/fidelity-american-century-adopting-new-tdf-fee-tactic-as-cost-pressures-grow-71695.

[16]See "ETF Scorecard: Exploring the ETF share-structure debate" (https://personal.vanguard.com/pdf/icrsc.pdf).

[17]For example, at the end of 2019, 24% of Vanguard TDF assets were held in the Vanguard Total Stock Market Index Fund Investor Share Class (ticker "VTSMX") with an expense ratio of 0.14% per annum. Vanguard could have instead held the Vanguard Total Stock Market ETF Shares (ticker "VTI") which charged an expense ratio of 0.03% at that time. Interestingly, as of July 2020, Investor Shares were no longer accessible to individual investors. However, Vanguard continued to hold these relatively more expensive share classes in their TDFs. Even if there were institutional needs to hold an open-end fund rather than an ETF, Vanguard could have also elected to hold the Vanguard Total Stock Market Admiral Share Class (ticker "VTSAX") which charged an expense ratio of 0.04% as of April 2021.

3.2 ETF Sample

We use a sample of low-cost ETFs to replicate TDFs. In our main analysis, we focus on Vanguard ETFs for several reasons.[18] First, Vanguard ETFs are known for low fees and relatively high liquidity.[19] Second, focusing on a relatively small set of ETFs allows for practical implementation of our Replicating Funds by asset managers and individual investors. We obtain ETFs' monthly returns and quarterly characteristics from CRSP. To mitigate the impact of illiquidity and possible non-synchronous prices due to infrequent trading, we include ETFs in our sample after they have achieved at least $50 million in assets.[20]

We restrict our universe of ETFs to 50 of the 80 available Vanguard ETFs (as of December 31, 2019). For a Vanguard ETF to be included in our sample, we require that it is part of a larger fund that also has the Admiral Share Class. The requirement serves two purposes. First, Vanguard funds that have both ETF shares and the Admiral Share Class are generally larger funds (and more liquid). Second, as a robustness check, we also create RFs using Vanguard Index Funds. As such, to maintain a similar menu of choices in constructing those funds, requiring that Vanguard ETF shares match to an Admiral Share Class ensures consistency. While the restriction lowers the number of Vanguard ETFs in our sample from 80 to 50, the 50 Vanguard ETFs account for 96.6% of the assets held in all 80 Vanguard ETFs (as of December 31, 2019). Finally, the 50 Vanguard ETFs' entries were staggered and most funds entered between 2006 and 2017. Internet Appendix

[18]For robustness, we replicate our analysis using a broader set of 2,412 ETFs, which we detail in the Internet Appendix. The results are qualitatively unchanged.

[19]In our broader set of ETFs, many ETFs are dominated by highly-correlated ETFs with lower fees and higher liquidity (see Brown, Cederburg, and Towner (2021)). Some ETFs charge expense ratios north of 70 bps per year (in 2019, the average expense ratio for Vanguard ETFs was 8 bps) and many ETFs are relatively small. For example, of the 289 ETFs used in our RFs (formed from the broader ETF sample) for the fourth quarter of 2019, 20 of them held less than $100 million in AUM as of December 2019. As a point of comparison, the smallest Vanguard ETF in the fourth quarter of 2019 held over $1.4 billion in AUM and 46 out of the 50 Vanguard ETFs were larger than $2 billion in AUM. Thus, given our desire to focus on low-cost ETFs with high liquidity for practicality, focusing on Vanguard ETFs is a broad stroke means of doing so. We note, however, that the outperformance we document using RFs formed from Vanguard ETFs is likely a *lower-bound* on what could be achieved. That is, while constraining ourselves to Vanguard ETFs is a simple and innocuous means of identifying cheap and liquid ETFs, one could likely replace some of the Vanguard ETFs in our sample with even cheaper ETFs with ample liquidity. For example, the Vanguard U.S. Small-Cap ETF (VB) could be replaced with the Schwab U.S. Small-Cap ETF (SCHA): as of August 2021, the Schwab ETF charged a smaller expense ratio (4 bps as compared to 5 bps on the Vanguard fund), the Schwab ETF was a relatively large fund ($\approx$$9 billion in AUM at the end of 2019) and the return correlation between the Schwab and Vanguard funds was 0.998. Nevertheless, we refrain from adding or deleting funds from our Vanguard ETF sample to avoid the appearance of padding our results.

[20]The $50 million threshold is common in the literature. See, for example, Brown, Davies, and Ringgenberg (2021).

Table IA1 lists the 50 Vanguard ETFs and shows that 33 Vanguard ETFs were available at the beginning of 2008 (i.e., the ETFs had met $50 million in AUM). One additional Vanguard fund became available in 2008, two in 2009, six in 2010, and the remaining eight became available at different times between 2011 and 2017. Thus, the set of available Vanguard ETFs is smaller during the early years of our sample.

3.3 Replicating Fund (RF) Construction

We follow a systematic method for constructing a Replicating Fund. First, for each TDF-quarter, we obtain that TDF's holdings report of underlying funds. To add concreteness and provide an illustration, on September 30, 2019, the Fidelity Freedom 2030 Fund held 29 different mutual funds (the Fidelity Freedom 2030 Fund is the largest TDF that holds both index funds and actively managed mutual funds). Some of the funds held by the Fidelity Freedom 2030 Fund were index funds, while others were actively managed mutual funds. For example, about 2% of the TDF was the Fidelity Series Large Cap Value Index Fund (ticker "FIOOX") which passively provides exposure to the Russell 1000 Value Index and about 5% of the TDF was the Fidelity Series Stock Selector Large Cap Value Fund (ticker "FBLEX") which is actively managed and focuses on large cap value stocks. That is, both "FIOOX" and "FBLEX" provide exposure to large cap value stocks, with one serving as an index fund and the other serving as an actively managed mutual fund.

We map each mutual fund in the TDF-quarter holdings report to an ETF based on (i) return correlation and (ii) an ETF's availability on that date. Each TDF holding-ETF return correlation is measured using all in-sample monthly returns of both the individual holding and the ETF.[21] We compute pairwise correlations for each TDF holding with each of the 50 Vanguard ETFs. An ETF can only be matched to an individual holding if the ETF was traded at the time of the holdings report and had reached its $50 million AUM threshold date.

Table 2 shows the matches of the Fidelity Freedom 2030 Fund to the 50 Vanguard ETFs. Across the 29 holdings, the mean correlation is 0.90 and the median correlation is 0.95. The difference

[21] While using in-sample data to compute return correlations provides stable, and potentially better, matches, it also introduces a look-ahead bias. As an alternative, we measure return correlations using daily returns over the prior year. We require at least 120 overlapping daily returns to calculate return correlations. Tables IA9 and IA10 in the Internet Appendix show that our main results are unchanged using rolling return correlations.

between the mean and median reflects several low correlations for underlying funds focused on commodities, money markets and floating-rate high income securities, which are not represented by the 50 Vanguard ETFs. Multiple holdings of the Fidelity Freedom 2030 Fund map to the same ETFs. For example, the Stock Selector Large Cap Value Fund, the Value Discovery Fund, the Large Cap Value Index Fund, and the Large Cap Stock Fund are all mapped to the Vanguard Value Index Fund ETF ("VTV") with correlations of 0.98, 0.98, 0.99 and 0.97. TDFs often hold highly correlated underlying funds – in this case the holdings include both active and passive funds.

Table 3 displays the distribution of correlations between each utilized mapping of an underlying holding to an ETF in our data, showing that similar correlation patterns exist throughout the sample. Overall, 59% (83%) of TDFs' underlying holdings are more than 0.95 (0.80) correlated with their mapped ETFs, and only 20% of TDFs' underlying holdings are correlated less than 0.50. Correlations are particularly strong for underlying index funds, with 76% (93%) having correlations more than 0.95 (0.80) with their mapped ETFs. Conversely, the non-index funds have lower correlations, but still 54% (81%) have correlations of more than 0.95 (0.80) with their mapped ETFs. While a small fraction of underlying funds have relatively poor correlations with their mapped ETFs, these holdings are relatively small and the poor individual correlations have little effect on the overall RF portfolios.[22]

Once each holding from a TDF-quarter holdings report is mapped to an ETF, we then form the RF's asset weights using the TDF's weights on its holdings. For example, since "FIOOX" represented 2% of the Fidelity Freedom 2030 Fund based on AUM, we use an asset weight of 2% for the matched ETF holding. We rebalance each RF on a quarterly basis using the most recent TDF's holdings report and fix those portfolio weights for the following quarter. We calculate monthly returns, net of fees, for each RF as a weighted average of the individual ETFs' returns in which weights are determined by asset weights. That is, the return of RF j in month t and quarter τ is,

$$r_{j,t}^{RF} = \sum_{i=1}^{N} r_{i,j,t}^{etf} \times \omega_{i,j,\tau}, \tag{1}$$

[22]Table IA4 in the Internet Appendix shows that correlations are stronger when we use a more inclusive set of ETFs. This is unsurprising as the Vanguard ETFs are a subset of the more inclusive set of ETFs. The stronger individual correlations do not lead to meaningfully better matching at the portfolio level.

in which the TDF holds N underlying funds, $r^{etf}_{i,j,t}$ is the return in month t of the ETF matched to fund i and $\omega_{i,j,\tau}$ is the weight of fund i in TDF j at the beginning of quarter τ.

To add additional concreteness to the construction of a RF and how it compares to its TDF, consider a hypothetical married couple, Ross and Rachel, in January 2006. Ross and Rachel are in their early 40s and expect to retire in 2030. As such, they put their 401(k) savings of $1 million into the Fidelity Freedom 2030 Fund. In December 2019, Ross and Rachel's savings would have grown to just over $2.33 million. However, had Ross and Rachel replicated the Fidelity Freedom 2030 Fund using our RF, their saving would have grown to nearly $2.60 million, an outperformance of over $267 thousand or 11%! Moreover, their replication of the Fidelity Freedom 2030 Fund would have exhibited little tracking error: the monthly return correlation is 0.993. Figure 2 illustrates the cumulative performance of one dollar invested in the Fidelity Freedom 2030 Fund and one dollar invested in the RF formed from a portfolio of ETFs. The figure shows the two funds' performances move in tandem with a growing spread due to primarily fee savings. This example puts investors' potential gains from using our RFs rather than TDFs into context. While a few basis points of monthly savings aggregates to substantial savings over our twelve-year sample period, young investors have even more to gain.

4 Replicating Fund Performance

In this section, we study the performance of RFs relative to the target-date funds they are designed to benchmark. To begin, we regress the monthly performance of our RFs on the monthly performance of their respective TDFs. That is, we perform univariate time-series regressions of the form,

$$r^{RF}_{j,t} = \alpha + \beta r^{TDF}_{j,t} + \epsilon_{j,t}, \tag{2}$$

in which α is the intercept, $r^{RF}_{j,t}$ is the return on the RF, $r^{TDF}_{j,t}$ is the reported return of the TDF, β is the coefficient on $r^{TDF}_{j,t}$, and $\epsilon_{j,t}$ is the error term.[23] Several estimated variables are of interest in understanding the relative performance of RFs versus TDFs. First, a positive intercept (i.e., $\alpha > 0$)

[23] We use standard errors clustered by fund family and quarter.

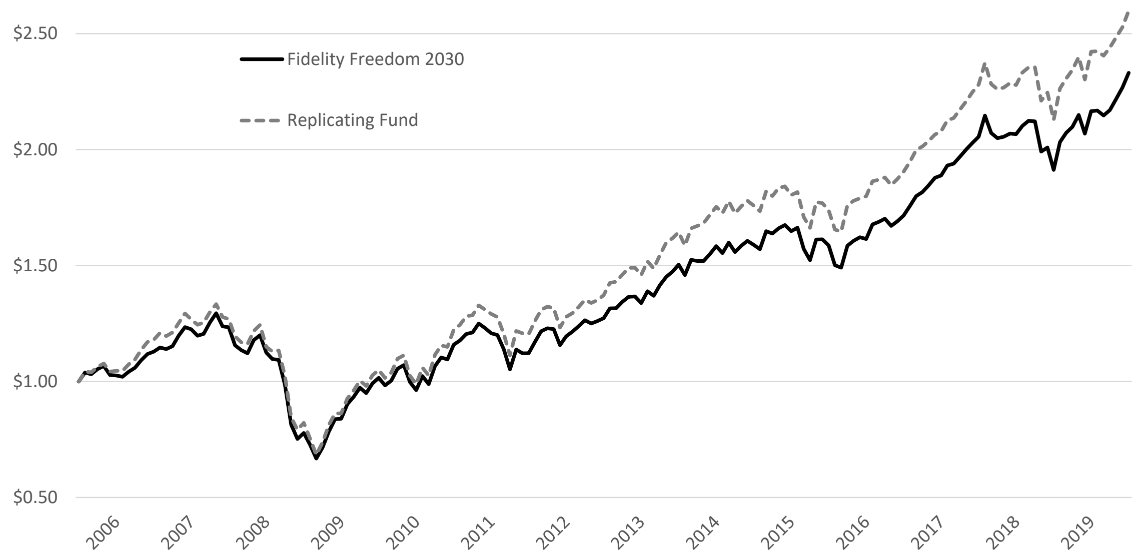

Figure 2: Fidelity Freedom 2030 Replicating Fund Performance. This figure shows the performance of one dollar invested in one of two portfolios: the Fidelity Freedom 2030 TDF or the Fidelity Freedom 2030 RF formed from Vanguard ETFs. The sample performance covers January 2006 through December 2019. The RF outperforms the TDF by 11.5% over the sample period and the correlation between the two funds' performances is 0.993.

has the economic interpretation that the RF improves an investor's Sharpe Ratio, assuming (i) the investor only faces market risk, and (ii) the investor's "market" portfolio is the asset allocation of the target-date fund. The null hypothesis is that α is not statistically different than zero. Second, the regression coefficient β gives a sense of "fit," that is, a good RF should have a regression coefficient near one. Under the null hypothesis, β equals one.

In addition to the estimated values of α and β, we are also interested in the regressions' tracking errors, measured by adjusted R^2 and root-mean squared error ($RMSE$). Furthermore, to the extent that RFs outperform their TDF counterparts, we are interested in evaluating the outperformance relative to tracking error. As such, we also calculate annualized appraisal ratios (i.e., excess return Sharpe Ratios) for each of our RFs. The annualized appraisal ratio is calculated as,

$$\text{Appraisal Ratio} = \frac{\alpha\sqrt{12}}{RMSE}, \tag{3}$$

and is interpreted as the replicating portfolio's excess return per unit of tracking error.

Table 4 estimates Eqn. 2 using our full sample and the first half (2008-2013) and second half

(2014-2019) sub-samples.[24] In the full sample, α is 5.0 bps per month (60 bps/year) and is statistically significant (t-statistic of 6.356). Across our two sub-samples, α is relatively stable; from 2008-2013 α is 5.2 bps per month and from 2014-2019 α is 5.4 bps per month.[25] The economically and statistically significant α estimates demonstrate significant underperformance by TDFs relative to their Replicating Fund benchmarks.

Turning to β, our estimates are all significantly greater than 1: β is 1.06 in the full sample and is 1.07 and 1.04 in each of the sub-samples. These estimates suggest that RFs are "scaled-up" versions of TDFs. This may be a function of two factors: the quality of mappings between underlying holdings and ETFs and the cash holdings of TDFs. The fit of the regressions, measured by adjusted R^2, suggests that poor matching is not driving the high β estimates. Adjusted R^2 is 96% in the full sample and ranges between 96% and 97% in the sub-samples. High adjusted R^2 values imply that RFs' outperformance comes with little tracking error: RFs' appraisal ratio is 0.28 for the full sample, 0.21 from 2008-2013, and 0.42 from 2014-2019. The other potential explanation for our high β estimates is differences in cash holdings. TDFs, and their underlying funds, often hold a significant fraction of AUM in cash. Holding cash allows a fund to more easily satisfy redemptions. ETFs, however, do not hold meaningful amounts of cash. Therefore, there is a performance wedge between ETFs and TDFs due to cash drag.[26] Given the high adjusted R^2 values of our regressions, our β estimates suggest that cash drag imposes a significant cost on TDF investors.

Because RFs are built to emulate TDFs' returns (which Table 4 shows that they do well), we can also evaluate RFs' performance by directly comparing their returns to TDFs' returns. We

[24]Some investors may have an aversion to ETFs and prefer open-end index funds (e.g., trading ETFs introduces an additional source of risk as investors may face a premium or discount to NAV when transacting). In Internet Appendix Table IA2 we form RFs using Vanguard Index Funds, rather than Vanguard ETFs. The results are qualitatively similar.

[25]We replicate the analysis in Table 4 by vintage and by fund family in Internet Appendix Tables IA7 and IA8. The main conclusions are unchanged.

[26]Cash drag refers to the attenuating effect holding cash has on performance. For example, in an S&P 500 index fund, cash drag will typically result in the fund's performance having a beta with the market smaller than one. Cash drag is typically disadvantageous in up-markets and advantageous in down-markets. The impact of cash drag on performance could be estimated using data on TDFs' and their underlying funds' cash holdings. Unfortunately, CRSP data on cash holdings (per_cash) are unreliable. For details, see Chernenko and Sunderam (2016) and Balduzzi and Reuter (2019).

define *Spread* as the monthly outperformance of a RF relative to its TDF, calculated as,

$$Spread_{j,t} = r_{j,t}^{RF} - r_{j,t}^{TDF}, \tag{4}$$

in which $r_{j,t}^{RF}$ is the return of the associated RF in month t and $r_{j,t}^{TDF}$ is the monthly return of TDF j in month t. *Spread* is highly related to the α estimates in Table 4 as positive values of *Spread* indicate that TDFs underperform their RF counterparts. However, *Spread* and α differ in that our calculation of *Spread* implicitly assumes $\beta = 1$. So while *Spread* does not account for differences in risk exposures, it better reflects investors' relative returns from replicating TDFs using ETFs.

Table 5 displays annual average *Spread* and its distributional characteristics from 2008 through 2019. On an equal-weighted (asset-weighted, based on end-of-year AUM) basis, average annualized *Spread* is 1.08% (0.53%) per year. On both an equal-weighted and asset-weighted basis, average *Spread* has moderately decreased over our sample period. Further, the asset-weighted average *Spread* is substantially lower than equal-weighted average *Spread* in most years. The difference is partially attributable to Vanguard TDFs, which have substantially lower fees than other fund families (see Table 1) and make up just over 10% of sample assets in 2006 and grow to nearly 38% by 2019. While *Spread* has tended to decrease, the costs to investors are substantial, and increasing due to the growth of the industry. Figure 3 shows the cumulative monthly *Spread* over our sample, which is denoted by the solid line. From 2008 to 2019, asset-weighted aggregate *Spread* totals $35.2 billion. In 2019 alone, *Spread* totals $9.7 billion.[27]

The distributional characteristics in Table 5 show significant variation in *Spread*. Average *Spread* varies considerably across years. For example, the annual equal-weighted average *Spread* is 3.34% in 2010, -0.25% in 2011, and 1.76% bps in 2012. *Spread* also varies considerably within year. While median *Spread* is generally in-line with equal-weighted average *Spread*, the 10^{th} and 90^{th} percentiles show high variation, ranging from slightly negative (indicating the RF underperformed the TDF) to significantly positive. For example, in 2013 the 10^{th} percentile spread indicates annual RF underperformance of 0.50% and the 90^{th} percentile spread indicates annual RF outperformance

[27]We note that, unlike Figure 1 in the Introduction, Figure 3 does not include the 10bps haircut (which accounted for possible additional overhead for TDFs relative to ETFs).

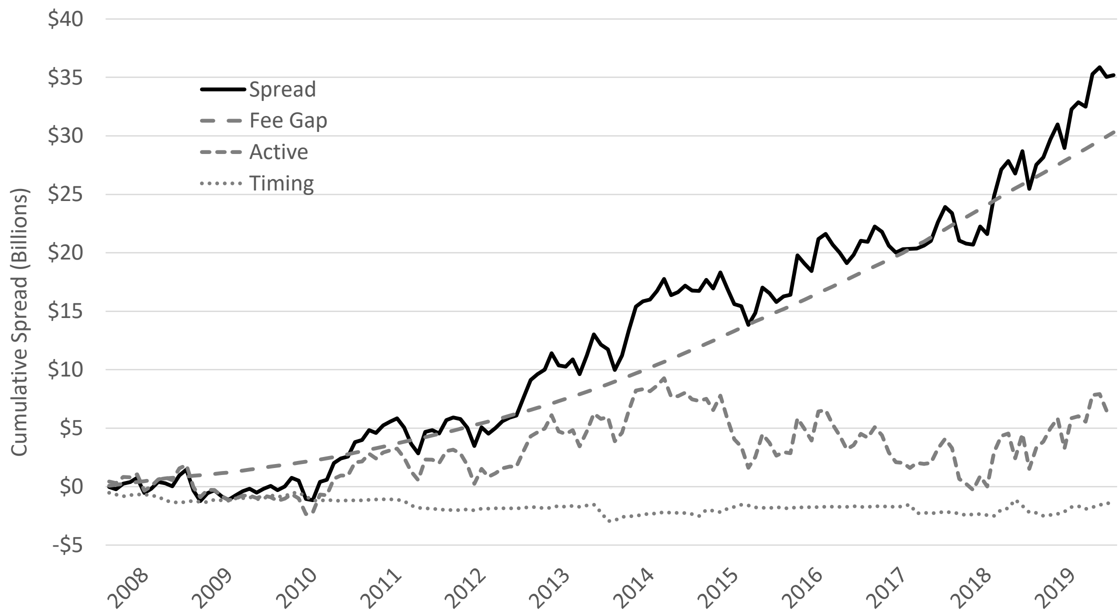

Figure 3: Cumulative Dollar Value of Spread and Spread Components. This figure plots monthly cumulative *Spread*, *Timing*, *Active* and *Fee Gap* of all TDFs in our sample on a dollar basis.

of 5.43%. In the first half of our sample, the average range between the 10th and 90th percentiles was 7.12% per year. Since 2014, the average 10th-90th percentile difference has decreased, but is an economically significant 3.24% per year. The substantially heterogeneity across TDFs indicates that there are significant differences across funds.

The construction of RFs implies that *Spread* takes glide paths as given, making *Spread* better suited than returns for comparing among TDFs. This is because conditioning on glide paths removes a major source of return dispersion due to differences in risk-taking (Balduzzi & Reuter, 2019). The differences in *Spread* highlight the dispersion in the implementation costs of TDFs. Thus, our benchmark serves as a means for identifying underperforming funds that can be obfuscated by only looking at realized returns.

4.1 Performance Decomposition

To understand the volatility of *Spread* across and within years, we decompose it into *Timing*, *Active*, and *Fee Gap*. Positive values for any of the three components indicate underperformance of TDFs relative to their RF counterparts along that dimension. *Timing* is the difference between what a TDF actually earns and what it would have earned if its underlying holdings were constant throughout the quarter. To calculate this implied fund return (denoted IF), we use the returns of the funds' individual holdings and beginning-of-quarter weights on those holdings,

$$r_{j,t}^{IF} = \sum_{i=1}^{N} r_{i,j,t}^{fund} \times \omega_{i,j,\tau}, \tag{5}$$

in which all parameters are the same as those in Eqn. 1 except that $r_{i,j,t}^{fund}$ is the reported monthly return of the held mutual fund. Because TDFs are funds-of-funds and most collect TDF-level fees, the actual TDF return may be mechanically smaller than the implied fund return. To account for TDFs' fees (we denote a TDF's explicit expense ratio, i.e., the fund-of-funds fee, as $exp\_ratio_{j,\tau}^{TDF}$), *Timing* is computed as,

$$Timing = r_{j,t}^{IF} - \left(r_{j,t}^{TDF} + \frac{exp\_ratio_{j,\tau}^{TDF}}{12} \right), \tag{6}$$

which is the difference between the TDF's gross performance (i.e., before fees) and its implied fund return $r_{j,t}^{IF}$. Similar to the return gap measure of Kacperczyk, Sialm, and Zheng (2008), *Timing* measures the performance impact of intra-quarter rebalancing (either across mutual funds or into new mutual funds). If a TDF's managers are skilled at adjusting asset allocations intra-quarter, *Timing* will be negative because the RF does not adjust holdings throughout the quarter.[28]

Active is the difference between the gross return of the underlying holdings and the gross return of the matched ETFs. To calculate gross returns, we calculate two measures of expense ratios. We calculate the expense ratio for the underlying assets actually held by TDF j as,

$$exp\_ratio_{j,\tau}^{IF} = \sum_{i=1}^{N} exp\_ratio_{i,j,\tau}^{fund} \times \omega_{i,j,\tau}, \tag{7}$$

in which all parameters are the same as those in Eqn. 1 except that $exp\_ratio_{i,j,\tau}^{fund}$ is the reported

[28] Because the portfolio weights for the quarter are fixed in Eqns. 1 and 5, both replicating returns and implied fund returns implicitly assume monthly rebalancing back to start-of-the-quarter weights. Thus, intra-quarter trading measured by *Timing* is evaluated relative to a monthly-rebalancing benchmark.

expense ratio of holding i, in quarter τ.[29] We also calculate the expense ratio of the RF as,

$$exp\_ratio^{RF}_{j,\tau} = \sum_{i=1}^{N} exp\_ratio^{etf}_{i,j,\tau} \times \omega_{i,j,\tau}, \tag{8}$$

in which $exp\_ratio^{etf}_{i,j,\tau}$ is the reported expense ratio of the ETF matched to fund i. Using the calculated expense ratios, *Active* is computed as,

$$Active = \left(r^{RF}_{j,t} + \frac{exp\_ratio^{RF}_{j,\tau}}{12} \right) - \left(r^{IF}_{j,t} + \frac{exp\_ratio^{IF}_{j,\tau}}{12} \right), \tag{9}$$

which is the difference between a RF's gross return and the implied fund's gross return. *Active* measures the performance impact of active asset selection within the mutual funds held by a TDF, and will be negative if a TDF's managers pick mutual funds that outperform their matched passive ETFs. Importantly, even if *Active* is negative, it does not imply the active mutual funds outperformed the matched ETFs on a net return basis (since fees are added back to returns used in calculating *Active*).

Finally, *Fee Gap* is the incremental fees that TDFs (fund-of-fund fees plus fees on the underlying holdings) charge in excess of what RFs' underlying ETFs charge, and is calculated as,

$$Fee\ Gap = \frac{1}{12} \left(exp\_ratio^{TDF}_{j,\tau} + exp\_ratio^{IF}_{j,\tau} - exp\_ratio^{RF}_{j,\tau} \right). \tag{10}$$

Note that $Spread = Timing + Active + Fee\ Gap$.

Table 6 reports our *Spread* decomposition for the full sample and each year from 2008 to 2019. Panel A reports average *Spread*: over the full sample, the average spread between a RF's performance and a TDF's performance is 8.6 bps per month (1.03% per year). Examining spread across years shows two important facts. First, *Spread* has moderately decreased over the sample. *Spread* averages 1.16% per year in the first half of the sample, and 0.92% per year in the second half. Second, *Spread* is volatile across years, ranging from -0.29% in 2011 to 2.95% in 2010. Panel

[29] For some TDF holdings (and some matched ETFs included in the broad set discussed in the Internet Appendix), no expense ratios are reported in CRSP (or are reported as -99). In these rare occasions in which no expense ratio is reported for a fund, we set it equal to zero to be as conservative as possible.

B reports the decomposition of *Spread* and provides insight into these facts.

Beginning with *Timing*, it averages only 0.4 bps per month, which is not statistically significant. This amounts to a small average cost of 0.05% per year to investors. Importantly, this average cost embeds cash drag at the TDF level. As our sample is predominantly during an extended bull market, it is likely that cash held by TDFs themselves led to decreases in average performance (and higher *Timing*). Thus, TDF managers do not exhibit any significant ability to take advantage of intra-quarter rebalancing to enhance TDF returns (at least beyond the effects of cash drag at the TDF level). Across our sample, *Timing* has moderately increased and is not particularly volatile, suggesting that *Timing* does not play a significant role in the time-series properties of *Spread*.

Across our sample, *Active* averages 3.4 bps per month, which is a statistically and economically significant 0.41% per year. This suggests that active asset selection has actually hurt TDF investors, even before fees (recall that positive values indicate outperformance by RFs). Put differently, the funds selected by TDFs on average underperform their matched ETFs on a gross basis. Similar to *Timing*, *Active* embeds cash drag. In this case, cash drag comes from the cash held by the underlying funds. The presence of cash drag implies that the cost of *Active* underperformance is not strictly attributable to poor asset selection by active managers. Rather, it may be partially attributable to the additional cash mutual funds hold relative to passive ETFs. Nevertheless, as the TDFs could invest in the ETFs used in our Replicating Funds, *Active* reflects a cost to the investor. Across our sample, *Active* shows no clear time trend, but does show substantial volatility. The correlation between *Spread* and *Active* across years is 0.98, indicating that almost all of the volatility in *Spread* is due to *Active*.

The largest source of TDF underperformance is *Fee Gap*, which averages 4.8 bps per month, or 0.57% per year. *Fee Gap* is remarkably stable, albeit decreasing, over time. In 2008, *Fee Gap* averaged 0.87% per year, and in 2019 averaged 0.46% per year. Falling TDF fees are consistent with general trends in asset management fees over the last decades, and the trend in *Fee Gap* accounts for the moderate decrease in *Spread* over our sample. While decreases in *Fee Gap* and *Spread* are good for investors, the levels of *Fee Gap* and *Spread* remain high and reflect substantial underperformance in the TDF industry.

Returning to Figure 3, in addition to showing the cumulative monthly *Spread* over our sample, the figure also shows each of the cumulative monthly components. *Timing* is denoted by the dotted line, is essentially zero and exhibits little variation. *Acitve* is denoted by the short-dashed line, is positive valued, and exhibits considerable variation. Finally, *Fee Gap* is denoted by the long-dashed line, is positive valued, exhibits almost no volatility, and serves essentially as the trend line for the overall *Spread* measure.

Table 7 repeats our *Spread* decomposition at the fund-family level. As in Table 1, we focus on the six largest TDF providers: Vanguard, Fidelity, T. Rowe Price, American Funds, TIAA-CREF, and JP Morgan. The remaining 56 TDF families are aggregated together and labeled "Other." Table 7 shows *Spread* ranges from 0.01% per year for Vanguard to 1.31% per year for American Funds and 1.29% per year for Other families. The main component of *Spread*, *Fee Gap*, is positive and statistically significant for each of the fund families. In other words, RFs generate fee savings relative to their TDF counterparts, regardless of fund family affiliation. Moreover, there is significant heterogeneity in *Fee Gap*; the largest TDF provider, Vanguard, measures a *Fee Gap* of 1 bps per month (7 bps per year). Conversely, the next two largest TDF providers, Fidelity and T. Rowe Price, have *Fee Gaps* of 3 and 5 bps per month (41 and 61 bps per year). American Funds and Other families have the highest *Fee Gaps* of 6 bps per month (66-73 bps per year).

Only Vanguard and JP Morgan have *Spread* estimates that are not significantly greater than zero. For Vanguard, the estimated *Spread* of 0.01% per year is lower than the *Fee Gap* due to good *Timing* (i.e., Vanguard's TDFs outperform their implied funds), suggesting Vanguard managers add value through rebalancing and other within-quarter trading. For JP Morgan, the estimated *Spread* of 0.12% per year is significantly lower than the *Fee Gap* of 0.47% per year. The lower *Spread* is due to good *Timing* and *Active* measures. Good *Active* indicates that JP Morgan on average selects funds that outperform their matched ETFs (T. Rowe Price also selects funds that outperform).

Fidelity, T. Rowe Price, American Funds, TIAA-CREF, and Other families all significantly underperform their Replicating Fund benchmarks. For Fidelity, good *Timing* does not offset significantly poor fund selection (*Active*) and a large *Fee Gap*, resulting in average *Spread* of 0.82%. T.

Rowe Price and TIAA-CREF perform relatively better, with average *Spreads* of 0.33% and 0.46% per year. T. Rowe Price benefits from good *Active*, while TIAA-CREF is slightly hurt by poor *Timing*. American Funds has the largest single-family *Spread* of 1.31%, attributable to the highest *Fee Gap* and the poorest *Active* measure. Other families suffer from high *Fee Gap*, poor *Timing* and poor *Active*. Overall, our decomposition shows that the large dispersion in performance across fund families is attributable to several sources, further highlighting the heterogeneity in TDF offerings. Nevertheless, despite the heterogeneity and with few exceptions, TDFs underperform RFs.

Our benchmark decomposition highlights that using *Spread* is an improvement above and beyond simply looking at fees. Fees should be examined in concert with active management costs/benefits, and any timing costs/benefits from rebalancing and tactical asset allocation. For example, as previously discussed, JP Morgan carries a relatively high average *Fee Gap* at 4.7 bps per month, but those fees are also associated with superior *Timing* and *Active* measures such that average *Spread* is not significantly different from zero. Thus, *Spread* serves as holistic assessment of TDF performance and its decomposition allows investors to measure the sources of value-add that TDF managers provide (or fail to provide).

4.2 Performance Flow Relation

A growing literature suggests that if investors use a benchmark to evaluate returns, then asset flows should be positively related to abnormal returns relative to that benchmark (Barber, Huang, & Odean, 2016; Berk & Van Binsbergen, 2016; Ben-David, Li, Rossi, & Song, 2022). Balduzzi and Reuter (2019) finds that TDF flows are primarily sensitive to alphas from an asset-class-based five-factor model. While our main results suggest investors are not aware of, nor sensitive to, *Spread* (otherwise average *Spread* would not be as high), we formally test whether TDF flows are sensitive to *Spread* by following Del Guercio and Reuter (2014) and Balduzzi and Reuter (2019) and estimating the following flow-performance model:

$$\mathrm{Flow}_{i,j,t} = a_j + b_t + \beta Spread_{i,j,t-1} + \Gamma Z_{i,j,t-1} + \epsilon_{i,j,t} \tag{11}$$

in which $Flow_{i,j,t}$ is a measure of annual flows, $Spread_{i,j,t}$ is the *Spread* for TDF i of vintage j in year t, and $Z_{i,j,t}$ decomposes past returns into alphas and systematic returns following Barber et al. (2016).[30] In the flow-performance model, we use two specifications for $Flow_{i,j,t}$. First, we use annual market-share-change for each TDF within their vintage (e.g., 2030) following Spiegel and Zhang (2013). Market-share-change is a relatively new and alternative specification for flow that is more resilient to heterogeneity among TDFs. Second, we use the canonical measure, annual net percentage flow, which is calculated by aggregating each TDF's monthly flow (in dollars) and dividing by its beginning of year AUM. To focus on cross-sectional variation, we include year fixed effects (b_t), and because flows to TDFs depend strongly on investors' proximity to retirement, we include vintage fixed effects (a_j).[31] Standard errors are clustered by fund family and year.

Table 8 reports the results for both market-share-change and percentage flow. The table collectively shows that investors, at best, have only moderate sensitivity to *Spread*. Beginning with market-share-change, there is no relation between the prior year's *Spread* and flows, both with and without controls. Interestingly, market-share-change is also not related to funds' lagged systematic returns or lagged 5-factor alphas but is highly related to lagged market-share-change. In other words, market-share-change is highly persistent and independent of realized performance. This finding suggests that the market is becoming more highly concentrated for reasons unrelated to traditional annual performance measures.

For percentage flow, when controlling for only the prior year's flow, the prior year's *Spread* is negatively related to flows (t-statistic of -2.658). In economic terms, the coefficient implies that a 100 bps increase in *Spread* is associated with a 2.23% fund outflow. However, the relation between prior year's *Spread* and annual net flow percentage is attenuated in both economic and statistical

[30] Alphas are estimated using the five-factors in Balduzzi and Reuter (2019): the CRSP U.S. value-weighted market index, the MSCI World Index excluding the United States, the Barclays U.S. Aggregate Bond Index, the Barclays Global Aggregate excluding the United States, and the GSCI Commodity Index. Annual alphas and systematic factors returns are calculated by compounding monthly returns over twelve months. Monthly alphas and systematic returns are based on monthly realized factor returns and factor loading estimated using daily returns over the past year. We require at least 120 daily returns to estimate factor loadings. 5-factor data are from Bloomberg.

[31] Our sample includes TDF years with 12 months of data and beginning of year assets. We exclude TDFs with vintages before 2020 as TDFs often experience significant outflows after their target retirement date is reached. We also exclude the Fidelity Freedom series and Fidelity Freedom K series in 2017 and 2018. In 2017, the Fidelity Freedom K series was consolidated into the Fidelity Freedom series resulting in large inflows to the Fidelity Freedom series and correspondingly large outflows to the Fidelity Freedom K series.

significance when including the prior year's alpha and systematic return. In fact, while the estimated coefficient on prior year's *Spread* remains negative valued, it is statistically insignificant. However, consistent with Balduzzi and Reuter (2019), investors' percentage flows into TDFs are persistent and are sensitive to past alphas.

Investors may not be very sensitive to *Spread* for several reasons. First, as our benchmarking method is novel, investors have not been able to easily observe *Spread* in the past. Second, as previously mentioned, the Pension Protection Act of 2006 designated TDFs as safe harbor investments to auto-enroll plan participants into (i.e., the default choice). Consequently, DC participants often invest in TDFs and many investors only invest in TDFs. Third, it is well established that DC plan participants exhibit inertia and are unlikely to exit default options (Madrian & Shea, 2001). Consequently, there is good reason to think that TDF investors' capital may be sticky and less susceptible to outflows. Fourth, the design of DC plans can impede competition. In particular, plan service providers serve as an intermediary between participants and capital markets; service providers supply a pre-specified menu of investment choices (e.g., a list of selected funds). Recent research has shown that when a mutual fund family acts as the service provider, the mutual fund family has a propensity to favor its own funds and products in the menu of offerings.[32] Given this body of evidence, it is unsurprising that investors are not responding to the costs and inefficiencies of TDFs.

5 Implementing Replicating Target-Date Funds

In practice, investors may be restricted in their access to TDFs, as retirement plans offer a limited menu of choices. For example, 401(k) plans that include TDFs offer an average 9.5 TDFs (Investment Company Institute, 2021b). Nine to ten TDFs is a limited choice set as our data contain 62 different fund families, 14 different vintages, and, at times, multiple funds within a family-vintage category. Despite having a limited choice set, some investors have the ability to deviate from their plans' menus via individual retirement accounts (IRAs) or self-directed brokerage

[32] Pool, Sialm, and Stefanescu (2016) shows that mutual fund families acting as plan service providers display favoritism toward their own affiliated funds. Moreover, plan participants do not account for this bias and exhibit a tendency to select affiliated funds (even poor performing funds).

windows. For example, The Plan Sponsor Council of America 2019 reports that 21.5% of plans in its 2018 survey offered a self-directed brokerage option (Investment Company Institute, 2020a).[33] As such, given that some investors may move beyond the menus they are presented, what should investors do? In this section, we help answer this question by providing practical guidance for investors. Specifically, we provide boilerplate portfolios based on the aggregate holdings of TDFs. Our boilerplate portfolios synthesize the aggregate performance of TDFs within a particular vintage via the 50 Vanguard ETFs and via consolidated menus of Vanguard ETFs (the latter may be easier to manage and more practical to implement).

To begin, recall that TDFs differ along many dimensions. Investors also have their own investment horizons and risk preferences, making it difficult to identify the "best" glide path or to prescribe a one-size-fits-all TDF. Thus, any general recommendation comes with substantial caveats. That said, if investors are only using investment horizon (e.g., 2030) as the criteria for selecting TDFs, and investors have difficulty matching their own preferences with the proper TDF, a representative portfolio based on aggregate holdings may be an improvement over current options.

To this end, we construct *Passive Replications of Funds (PROFs)* by aggregating, within each vintage, the asset-weighted replicated holdings of all TDFs. This approach has several advantages. First, it remains agnostic to which glide paths, asset allocations, and "to" versus "through" styles are better. Second, by the revealed preferences of investors, it places greater weight on larger TDFs.[34] Third, by asset-weighting our replicated holdings (i.e., the 50 Vanguard funds), it allows for a reasonable number of funds in the portfolio relative to the number of funds held by all TDFs. For example, PROFs hold at most 50 ETFs. In contrast, the asset-weighted holdings of all TDFs using actual holdings would require over 200 positions in different funds.

While holding 50 positions in a portfolio is not uncommon, it requires routine maintenance and rebalancing that many investors may not be willing to do. Thus, investors may prefer portfolios with even fewer positions. Accordingly, we analyze the holdings of each of our unconstrained

[33] The study also reports that few participants currently utilize the option: less than 1% of total plan assets in 2018 were invested via the brokerage window.

[34] Given that investors' choices are often limited, investors are often defaulted into TDF holdings, and investors may suffer from inertia, we recognize that fund size does not necessarily equate to revealed preference as it may in other settings.

PROFs and consolidate the holdings into a subset of ETFs. We term our initial, unconstrained fund PROF_{50}. We then consolidate the holdings across those 50 Vanguard ETFs into sets of 20, 13 and 6 ETFs which we term PROF_{20}, PROF_{13}, and PROF_{6}. The consolidation process replaces an ETF no longer in the choice set with the most highly correlated ETF in the choice set. For example, in consolidating from 50 to 20 ETFs, Mid Cap Growth and Mid Cap Value ETFs are removed. Naturally, these ETFs are replaced by the Mid Cap ETF. Moreover, the Mid Cap ETF is included in the choice sets for the PROF_{50}, PROF_{20} and PROF_{13} portfolios, but not the PROF_{6} portfolio. As a result, the Mid Cap ETF is replaced by the Total Stock Market ETF in PROF_{6}. Table 9 illustrates the consolidation process using 2030 vintage PROFs in September 2019.

We use several criteria to determine the consolidated sets of 20, 13 and 6 ETFs. We first determine which ETFs are most commonly used and which ETFs receive the largest allocations in the RFs described in Section 3.3. While each TDF's replicated holdings are unique, several common themes emerge. Almost all TDFs hold mutual funds that provide broad equity exposure, broad bond exposure, international exposure and real estate exposure. Most near-retirement vintages also hold significant positions in treasury inflation protected securities (TIPS). To represent the most common exposures, our consolidated set of 6 includes broad equity and bond ETFs, both for US and international markets, as well as the real estate ETF and the TIPS ETF offered by Vanguard. Beyond the broad asset categories represented by these 6 ETFs, many TDFs use more focused funds to get exposures to different sizes of equities, maturities of bonds, and regions of the world. Thus, we expand the consolidated set of 13 Vanguard ETFs to include small, medium and large cap ETFs, intermediate-term corporate bond and short-term treasury ETFs, and emerging and developed market ETFs. Our consolidated set of 20 ETFs incorporates TDFs' focused on different valuation metrics (i.e., value and growth), specific benchmarks (e.g., the S&P 500), long-term maturity bonds, and specific regions of international exposure (e.g., Europe). Table 9 details which ETFs are in the choice set for each PROF.

There are two natural benchmarks for analyzing the performance of our PROFs. As PROFs are formed based on the asset-weighted holdings of the individual RFs, the most direct comparison is to the asset-weighted returns of each vintage's TDFs. Formally, define

$$r_{v,t}^{Agg} = \frac{\sum\limits_{j \in X(v,t)} r_{j,t}^{TDF} AUM_{j,t}}{\sum\limits_{j \in X(v,t)} AUM_{j,t}}, \tag{12}$$

in which v indexes vintages, $X(v,t)$ is the set of TDFs in vintage v in month t, $r_{j,t}^{TDF}$ is the return in month t of TDF j and $AUM_{j,t}$ is the assets under management in TDF j at the beginning of month t. A second natural benchmark is how they perform relative to the individual TDFs in the same vintage. Thus, we perform pairwise comparisons of each vintage's PROFs relative to each TDF in that vintage.

We start by regressing PROF returns on aggregate TDF returns:

$$r_{v,t}^{PROF_z} = \alpha + \beta r_{v,t}^{Agg} + \epsilon_{v,t}, \tag{13}$$

in which $r_{v,t}^{PROF_z}$ is the PROF return of vintage v in month t with $z \in \{6, 13, 20, 50\}$, $r_{v,t}^{Agg}$ is the aggregate return of TDFs of vintage v in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{v,t}^{Agg}$, and $\epsilon_{v,t}$ is the error term. Panel A of Table 10 shows that the unconsolidated aggregate replicating portfolios (i.e., PROF_{50} portfolios) outperform the aggregate TDF portfolios with an α estimate of 2.5 bps per month. Moreover, the β estimate is significantly greater than one, consistent with cash drag lowering the returns of the aggregate TDF portfolio. The PROF_{20} and PROF_{13} portfolios perform similarly, although the α estimates are reduced to 1.9 and 2.0 bps per month. The PROF_{6} portfolio performs slightly worse from an α standpoint, earning an extra 1.6 bps per month. Notably, the PROFs maintain very high adjusted R^2 despite the limited choice set of ETFs, with the PROF_{6} portfolio having the lowest at 0.992.

Panel B of Table 10 shows that the PROF portfolios consistently outperform the aggregate TDF portfolios using traditional measures of performance. The annualized returns of the PROF portfolios exceed the aggregate TDF portfolio by between 0.34% and 0.50% per year. Consistent with cash drag dampening the returns of the aggregate TDF portfolios, the PROF portfolios take more risk, as measured by the standard deviation of annual returns. However, Sharpe ratios are still consistently higher for the PROF portfolios relative to the aggregate TDF portfolios.

As a final measure of performance, we compare the PROFs to each TDF in the same vintage. From an investor's perspective, the PROFs may be undesirable if they underperform a significant portion of existing TDFs. Accordingly, we compare the performance of each pair of PROF_z and TDF j by calculating the ratio of their cumulative performance in our sample period. Formally, we define

$$AnnualOutperformance_{z,v,j} = \left(\frac{\prod_{t \in T} (1 + r_{v,t}^{PROF_z})}{\prod_{t \in T, j \in X(v,t)} (1 + r_{j,t}^{TDF})} \right)^{12/T} - 1, \tag{14}$$

in which $z \in \{6, 13, 20, 50\}$ is the size of the choice set for the PROF, T is the set of months both TDF j and PROF_z have returns, v is the vintage, j is the comparable TDF, $r_{v,t}^{PROF_z}$ is the return of PROF_z in vintage v and in month t and $r_{j,t}^{TDF}$ is the return of TDF j in month t. For each PROF, we then average the annual outperformance across comparable TDFs on both an equal-weighted and an asset-weighted basis. We also calculate the percentage of positive outperformance measures on both an equal-weighted and an asset-weighted basis.

Panel C of Table 10 shows that PROFs perform better than most available TDFs over our sample period. On average, PROFs outperform other TDFs by between 0.80% and 0.92% on an equal-weighted basis and by between 0.30% and 0.43% on an asset-weighted basis. The large discrepancy is because Vanguard makes up a large portion of the aggregate TDF market on an asset-weighted basis, and in general, Vanguard TDFs perform well due to low expense ratios. This can also be seen in evaluating the aggregate TDF, which outperforms the average TDF by 0.43% (-0.07%) on an equal (asset) weighted basis. In addition to outperforming substantially on average, PROFs cumulatively outperform between 87% and 91% (70% and 89%) of available TDFs on an equal (asset) weighted basis. In contrast, the aggregate TDF outperforms only 68% (46%) of the individual TDFs.

Overall, the results show that PROFs significantly outperform TDFs in aggregate and compare favorably against individual TDFs. For investors who desire the simplicity of a TDF, PROFs provide a reliable means to earn low-fee returns and benefit from the wisdom of the TDF "crowd" in forming asset allocations. Importantly, an investor does not have to fully replicate the aggregate

TDF portfolio with ETFs. Rather, an investor can focus on a smaller set of ETFs, limiting the number of positions that must be monitored and rebalanced at quarterly intervals. Our analysis suggests that using a set of 6 ETFs produces results similar to portfolios comprised of up to 50 ETFs. Thus, our PROF_6 portfolios provide an easy to implement alternative to high-fee TDFs that exist in the current marketplace, allowing investors to directly compete with TDF providers.

6 Conclusion

Target-date funds offer investors the ability to take a set-it-and-forget-it approach to retirement investing. Investors can direct all of their retirement savings to one TDF that manages the underlying investments, providing diversification and adjusting asset allocations over the investment horizon. Unfortunately, most investors have a limited menu of TDFs and fund sponsors do not face the same types of competition as compared to other market settings. Moreover, TDFs are complex vehicles with little transparency relating to portfolio construction, fees, and properly benchmarked performance. Our benchmarking methodology demystifies TDFs by providing a clear performance metric that can be decomposed to highlight the high fees and poor fund selection that pollute the TDF industry.

While we show that TDF investors are incurring significant excess costs, our results do not imply that target date funds do not provide value to their investors. Rather, the initial portfolio allocations and dynamic adjustments (glide paths) are valuable services that many investors happily pay for. Our analysis takes these allocations and dynamics as given, as we utilize their publicly available holdings data to build our Replicating Funds. Assessing the value of TDFs' glide paths (and the value of deviations from planned glide paths) is an important area of research, but is outside the scope of our analysis.

We provide investors with easy-to-implement, low-cost target-date portfolios that outperform the majority of existing TDFs. Our *Passive Replications of Funds* (*PROFs*) give investors a simple, low-cost alternative to the wide range of TDFs on the market. Investors who desire particular glide paths or asset allocations can emulate specific Replicating Funds to save on fees. The fee savings from RFs can also be viewed as a substitute for financial advice. Financial advisors, or even roboadvisors, could lower clients' net costs by selecting the most appropriate TDFs, replicating their holdings, and covering their own fees with a portion of the Replicating Fund savings.

References

Alexander, G. J., Jones, J. D., & Nigro, P. J. (1998). Mutual fund shareholders: Characteristics, investor knowledge, and sources of information. *Financial Services Review*, *7*(4), 301–316.

Balduzzi, P., & Reuter, J. (2019). Heterogeneity in target date funds: Strategic risk-taking or risk matching? *The Review of Financial Studies*, *32*(1), 300–337.

Barber, B. M., Huang, X., & Odean, T. (2016). Which factors matter to investors? evidence from mutual fund flows. *The Review of Financial Studies*, *29*(10), 2600–2642.

Barber, B. M., Odean, T., & Zheng, L. (2005). Out of sight, out of mind: The effects of expenses on mutual fund flows. *The Journal of Business*, *78*(6), 2095–2120.

Ben-David, I., Li, J., Rossi, A., & Song, Y. (2022). What do mutual fund investors really care about? *The Review of Financial Studies*, *35*(4), 1723–1774.

Berk, J. B., & Van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, *118*(1), 1–20.

Berk, J. B., & Van Binsbergen, J. H. (2016). Assessing asset pricing models using revealed preference. *Journal of Financial Economics*, *119*(1), 1–23.

Brown, D. C., Cederburg, S., & Towner, M. (2021). Dominated ETFs. *Working Paper*.

Brown, D. C., Davies, S. W., & Ringgenberg, M. C. (2021). ETF arbitrage, non-fundamental demand, and return predictability. *Review of Finance*, *25*(4), 937–972.

Capon, N., Fitzsimons, G. J., & Prince, R. A. (1996). An individual level analysis of the mutual fund investment decision. *Journal of Financial Services Research*, *10*(1), 59–82.

Chernenko, S., & Sunderam, A. (2016). Liquidity transformation in asset management: Evidence from the cash holdings of mutual funds. *National Bureau of Economic Research Working Paper*.

Cooper, M., Halling, M., & Yang, W. (2021). The persistence of fee dispersion among mutual funds. *Review of Finance*, *25*(2), 365–402.

Del Guercio, D., & Reuter, J. (2014). Mutual fund performance and the incentive to generate alpha. *The Journal of Finance*, *69*(4), 1673–1704.

Elton, E. J., Gruber, M. J., de Souza, A., & Blake, C. R. (2015). Target date funds: Characteristics and performance. *Review of Asset Pricing Studies*, *5*(2), 254–272.

Investment Company Institute. (2020a). *The economics of providing 401(k) plans: Services, fees,* *and expenses, 2019*. Investment Company Institute.

Investment Company Institute. (2020b). *Investment company fact book*. Investment Company Institute.

Investment Company Institute. (2021a). *Investment company fact book*. Investment Company Institute.

Investment Company Institute. (2021b). *Ten important facts about 401(k) plans*. Investment Company Institute.

Kacperczyk, M., Sialm, C., & Zheng, L. (2008). Unobserved actions of mutual funds. *The Review of Financial Studies*, *21*(6), 2379–2416.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly Journal of Economics*, *116*(4), 1149–1187.

Malkiel, B. G. (1995). Returns from investing in equity mutual funds 1971 to 1991. *The Journal of Finance*, *50*(2), 549–572.

Mao, M. Q., & Wong, C. H. (2022). Why have target-date funds performed better in the COVID-19 selloff than the 2008 selloff? *Journal of Banking & Finance*, *135*, 106367.

Massa, M., Moussawi, R., & Simonov, A. (2022). The unintended consequences of investing for the long run: Evidence from target date funds. *Working Paper*.

Moraes, F., Cavalcante-Filho, E., & De-Losso, R. (2021). Unskilled fund managers: Replicating active fund performance with few ETFs. *International Review of Financial Analysis*, *78*, 101900.

Parker, J. A., Schoar, A., & Sun, Y. (2020). Retail financial innovation and stock market dynamics: The case of target date funds. *National Bureau of Economic Research Working Paper*.

Pool, V. K., Sialm, C., & Stefanescu, I. (2016). It pays to set the menu: Mutual fund investment options in 401 (k) plans. *The Journal of Finance*, *71*(4), 1779–1812.

Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management*, *18*(2), 7–19.

Shoven, J. B., & Walton, D. B. (2021). An analysis of the performance of target date funds. *The Journal of Retirement*, *8*(4), 43–65.

Spiegel, M., & Zhang, H. (2013). Mutual fund risk and market share-adjusted fund flows. *Journal of Financial Economics*, *108*(2), 506–528.

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy*, *112*(S1), S164–S187.

Table 1: Summary Statistics

This table summarizes our sample by year and fund family. A TDF is included in our sample if it reports quarterly holdings and if both the TDF and the TDF's fund holdings appear in the CRSP mutual fund data base with monthly returns and net asset values. See the Data Appendix for additional details. TDF Fee is the reported expense ratio of the TDF, and Holdings Fee is the fees collected on the funds held by the TDF. # Holdings reports the number of individual funds held by a TDF and % Index reports the fraction of held funds which are index funds. Number and Market capitalization are measured based on each fund's last calendar year observation. Annual aggregate TDF AUM is reported in billions. Fund family statistics in Panel B are based on each fund's last monthly observation in 2019.

Panel A: Summary Statistics by Year

				Equal-Weighted Average				Asset-Weighted Average			
Year	Funds	Classes	AUM	TDF Fee	Holdings Fee	# Holdings	% Index	TDF Fee	Holdings Fee	# Holdings	% Index
2008	148	663	$145	0.30%	0.68%	15.5	22.59%	0.07%	0.59%	17.3	32.26%
2009	167	747	$184	0.28%	0.74%	16.3	17.34%	0.08%	0.70%	17.2	19.06%
2010	351	1,312	$312	0.24%	0.57%	16.5	28.61%	0.08%	0.54%	18.4	36.80%
2011	369	1,324	$343	0.23%	0.52%	16.4	30.17%	0.08%	0.51%	16.1	37.52%
2012	408	1,513	$445	0.23%	0.49%	16.2	32.46%	0.08%	0.49%	15.8	39.54%
2013	446	1,710	$577	0.20%	0.52%	17.1	32.38%	0.09%	0.48%	17.3	39.42%
2014	484	1,807	$656	0.18%	0.50%	18.0	31.18%	0.09%	0.48%	17.0	37.48%
2015	521	1,952	$718	0.19%	0.47%	17.4	35.71%	0.08%	0.46%	16.4	39.26%
2016	572	2,190	$839	0.19%	0.44%	17.5	36.50%	0.08%	0.42%	15.9	42.72%
2017	588	2,191	$1,141	0.21%	0.37%	17.1	39.81%	0.17%	0.33%	16.6	42.94%
2018	572	2,218	$1,024	0.24%	0.30%	17.3	41.08%	0.19%	0.24%	15.1	47.04%
2019	569	2,270	$1,299	0.23%	0.29%	17.3	40.16%	0.18%	0.22%	14.8	49.65%

Panel B: 2019 Summary Statistics by Fund Family

				Equal-Weighted Average				Asset-Weighted Average			
Fund Family	Funds	Classes	AUM	TDF Fee	Holdings Fee	# Holdings	% Index	TDF Fee	Holdings Fee	# Holdings	% Index
Vanguard	22	22	$490	0.00%	0.11%	4.2	100.00%	0.00%	0.12%	4.2	100.00%
Fidelity	92	370	$266	0.31%	0.02%	22.4	26.65%	0.53%	0.02%	24.6	24.80%
T.RowePrice	37	83	$164	0.08%	0.54%	21.2	11.88%	0.05%	0.61%	21.5	12.23%
AmericanFunds	12	147	$154	0.26%	0.36%	18.1	0.00%	0.23%	0.36%	19.5	0.00%
TIAA-CREF	22	88	$60	0.23%	0.00%	11.4	48.72%	0.24%	0.00%	12.5	41.21%
JPMorgan	19	133	$50	0.15%	0.31%	19.4	37.73%	0.17%	0.41%	23.1	16.45%
Other	365	1,427	$115	0.25%	0.36%	16.7	43.76%	0.22%	0.38%	19.2	31.67%

Table 2: Fidelity Freedom 2030 Fund Replication Example

This table displays the Fidelity Freedom 2030 Fund's 29 mutual fund constituents according to their s12 reported holdings on September 30, 2019. Fund positions are reported in millions and correlation reports the monthly return correlation between the mutual fund holding and the matched Vanguard ETF.

Position	Fund Ticker	Fund Name	ETF Ticker	ETF Name	Correlation
$5,918	FSIGX	Investment Grade Bond Fund	BND	Total Bond Market Index Fund	0.92
$2,656	FEMSX	Emerging Markets Opportunities Fund	VWO	Emerging Markets Stock Index Fund	0.98
$2,471	FIGSX	International Growth Fund	VEA	Developed Markets Index Fund	0.96
$2,448	FINVX	International Value Fund	VEA	Developed Markets Index Fund	0.98
$2,285	FDMLX	Intrinsic Opportunities Fund	VT	Total World Stock Index Fund	0.92
$2,016	FGLGX	Large Cap Stock Fund	VTV	Value Index Fund	0.97
$1,858	FCGSX	Growth Company Fund	VUG	Growth Index Fund	0.95
$1,572	FBLEX	Stock Selector Large Cap Value Fund	VTV	Value Index Fund	0.98
$1,310	FSIPX	Inflation-Protected Bond Index Fund	VTIP	Short-Term Inflation-Protected Securities Index Fund	0.90
$1,145	FCSSX	Commodity Strategy Fund	VDE	Energy Index Fund	0.69
$1,114	FNKLX	Value Discovery Fund	VTV	Value Index Fund	0.98
$1,028	FVWSX	Opportunistic Insights Fund	VUG	Growth Index Fund	0.96
$920	FSBDX	Blue Chip Growth Fund	VUG	Growth Index Fund	0.95
$823	FSOPX	Small Cap Opportunities Fund	VB	Small-Cap Index Fund	0.99
$815	FTLTX	Long-Term Treasury Bond Fund	VGLT	Long-Term Treasury Index Fund	0.99
$766	FSOSX	Overseas Fund	VSS	FTSE All-World ex-US Small-Cap Index Fund	0.78
$581	FIOOX	Large Cap Value Index Fund	VTV	Value Index Fund	0.99
$540	FSAEX	All-Sector Equity Fund	VOO	500 Index Fund	0.98
$527	FSTSX	International Small Cap Fund	VSS	FTSE All-World ex-US Small-Cap Index Fund	0.96
$288	FCNSX	Canada Fund	VOE	Mid-Cap Value Index Fund	0.90
$287	FHKFX	Emerging Markets Fund	VWO	Emerging Markets Stock Index Fund	0.94
$275	FJACX	Small Cap Discovery Fund	VBR	Small-Cap Value Index Fund	0.96
$241	FSHNX	High Income Fund	VT	Total World Stock Index Fund	0.82
$216	FGNXX	Government Money Market Fund	VGSH	Short-Term Treasury Index Fund	0.45
$204	FEDCX	Emerging Markets Debt Fund	VWOB	Emerging Markets Government Bond Index Fund	0.91
$156	FSREX	Real Estate Income Fund	VNQ	Real Estate Index Fund	0.83
$90	FYBTX	Short-Term Credit Fund	VCSH	Short-Term Corporate Bond Index Fund	0.95
$40	FFHCX	Floating Rate High Income Fund	VT	Total World Stock Index Fund	0.66
$17	FCDSX	International Credit Fund	VCIT	Intermediate-Term Corporate Bond Index Fund	0.88

Table 3: Correlations of TDF Holdings Matched to Vanguard ETFs

This table summarizes the match quality of TDF holdings to Vanguard ETFs based on their return correlations. Correlations are based on all overlapping monthly returns between 2008 and 2019. We use the CRSP index fund flag to categorize TDF holdings.

	Correlation Range									
	≤0.5	0.50-0.75	0.75-0.80	0.80-0.85	0.85-0.90	0.90-0.925	0.925-0.95	0.95-0.975	0.975-0.99	>.99
All Funds										
Number of Funds	2,412									
Fund Distribution	113	168	125	123	138	115	206	523	569	332
Percent	4.68%	6.97%	5.18%	5.10%	5.72%	4.77%	8.54%	21.68%	23.59%	13.76%
Index Funds										
Number of Funds	511									
Fund Distribution	5	19	14	20	23	17	23	66	92	232
Percent	0.98%	3.72%	2.74%	3.91%	4.50%	3.33%	4.50%	12.92%	18.00%	45.40%
Non-Index Funds										
Number of Funds	1,901									
Fund Distribution	108	149	111	103	115	98	183	457	477	100
Percent	5.68%	7.84%	5.84%	5.42%	6.05%	5.16%	9.63%	24.04%	25.09%	5.26%

Table 4: Replicating Fund Performance

This table examines the performance of replicating funds that are formed by matching each TDF's holdings to the Vanguard sample ETF with the highest monthly return correlation. Each replicating fund's monthly performance is regressed on its corresponding TDF reported return:

$$r_{j,t}^{RF} = \alpha + \beta r_{j,t}^{TDF} + \epsilon_{j,t}$$

in which $r_{j,t}^{RF}$ is the replicating fund return of TDF j in month t, $r_{j,t}^{TDF}$ is the return of TDF j in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{j,t}^{TDF}$, and $\epsilon_{j,t}$ is the error term. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using White standard errors and t-statistics are reported below each estimated coefficient. For β, t-statistics are reported under the null that the coefficient is equal to 1. Each regression also reports the root mean squared error (RMSE) and the annualized appraisal ratio which is calculated as Appraisal Ratio = $\frac{\alpha\sqrt{12}}{RMSE}$.

	$r_{j,t}^{RF}$		
	(1) Full Sample	(2) 2008-2013	(3) 2014-2019
$r_{j,t}^{TDF}$	1.055***	1.069***	1.038***
	(14.082)	(11.339)	(8.788)
α	0.050***	0.052***	0.054***
	(6.356)	(3.127)	(6.906)
RMSE	0.613	0.844	0.443
Appraisal Ratio	0.284	0.212	0.423
Adjusted R^2	0.963	0.959	0.970
Observations	54,642	18,783	35,859

Table 5: Annual TDF Spreads by Year

This table summarizes *Spread* measures by year. Within each calendar year, each TDF's monthly spread is computed as $r_{j,t}^{RF} - r_{j,t}^{TDF}$, the monthly spread is averaged for the year, and is then annualized based on the number of months of data (per TDF) within the year. Within each calendar year, an equal-weighted average, an asset-weighted average, and the 10^{th}, 50^{th}, and 90^{th} percentile values are reported. Number and Market capitalization are measured based on each fund's last monthly observation within a year. Annual aggregate TDF AUM is reported in billions.

Year	Funds	AUM	EW Spread	AW Spread	10^{th} Perc.	50^{th} Perc.	90^{th} Perc.
2008	148	\$145	0.09%	1.01%	-4.62%	0.83%	5.57%
2009	167	\$184	0.82%	-1.34%	-4.44%	1.02%	6.04%
2010	351	\$312	3.34%	1.88%	0.18%	2.12%	7.28%
2011	369	\$343	-0.25%	0.20%	-2.10%	0.07%	2.44%
2012	408	\$445	1.76%	0.77%	-0.11%	1.22%	4.35%
2013	446	\$577	1.88%	0.83%	-0.50%	1.36%	5.43%
2014	484	\$656	1.28%	0.82%	0.05%	1.23%	2.90%
2015	521	\$718	0.42%	-0.11%	-0.95%	0.30%	1.88%
2016	572	\$839	1.02%	0.71%	-0.40%	0.91%	3.21%
2017	588	\$1,141	1.00%	0.16%	-0.64%	0.64%	2.68%
2018	572	\$1,024	0.07%	0.13%	-1.15%	0.10%	1.71%
2019	569	\$1,299	1.39%	0.87%	-0.34%	1.13%	3.60%
Full Sample	5,195	\$7,682	1.08%	0.53%	-0.90%	0.81%	3.38%

Table 6: Replicating Fund Performance: *Spread*

This table decomposes the difference in monthly performance of RFs and their corresponding TDFs. Panel A reports the average total monthly return *Spread* ($r^{RF}_{j,t} - r^{TDF}_{j,t}$) for the full sample and each year from 2008 to 2019. Panel B decomposes the average total return spread into three components. The first component is *Timing* which is calculated as the average of $r^{IF}_{j,t} - r^{TDF}_{j,t} - \left(exp\_ratio^{TDF}_{j,\tau}\right)/12$. The second component is *Active* which is calculated as the average of $r^{RF}_{j,t} - r^{IF}_{j,t} + \left(exp\_ratio^{RF}_{j,\tau}\right)/12 - \left(exp\_ratio^{IF}_{j,\tau}\right)/12$. The third component is *Fee Gap* which is calculated as the average of $\left(exp\_ratio^{TDF}_{j,\tau}\right)/12 + \left(exp\_ratio^{IF}_{j,\tau}\right)/12 - \left(exp\_ratio^{RF}_{j,\tau}\right)/12$. ***,**,* indicate statistical significance at the 1%, 5%, and 10% and t-statistics are reported below each reported average.

Panel A: Monthly Return Spread

	(1) Full Sample	(2) 2008	(3) 2009	(4) 2010	(5) 2011	(6) 2012	(7) 2013	(8) 2014	(9) 2015	(10) 2016	(11) 2017	(12) 2018	(13) 2019
Spread	0.086***	0.013	0.052	0.246***	-0.024	0.143***	0.149***	0.110***	0.033**	0.095***	0.082***	0.018	0.119***
	(11.555)	(0.202)	(0.687)	(5.370)	(-0.526)	(5.544)	(7.539)	(7.474)	(2.005)	(4.563)	(5.755)	(0.991)	(6.526)
Observations	54,642	1,455	1,617	2,765	4,060	4,242	4,644	5,248	5,705	6,066	6,391	6,336	6,113

Panel B: Monthly Return Spread Decomposition

	(1) Full Sample	(2) 2008	(3) 2009	(4) 2010	(5) 2011	(6) 2012	(7) 2013	(8) 2014	(9) 2015	(10) 2016	(11) 2017	(12) 2018	(13) 2019
Timing	0.004	-0.062***	-0.010	0.004	-0.016	0.011	0.016	0.015***	0.001	0.010**	0.009***	0.012	-0.002
	(1.393)	(-4.121)	(-0.213)	(0.304)	(-1.368)	(1.494)	(1.404)	(2.864)	(0.155)	(2.124)	(3.104)	(0.958)	(-0.281)
Active	0.034***	0.002	-0.010	0.182***	-0.061	0.081***	0.085***	0.047***	-0.015	0.040**	0.031**	-0.034*	0.083***
	(4.776)	(0.038)	(-0.154)	(4.171)	(-1.383)	(3.228)	(4.042)	(3.286)	(-0.970)	(1.997)	(2.294)	(-1.737)	(4.564)
Fee Gap	0.048***	0.072***	0.073***	0.060***	0.053***	0.051***	0.048***	0.049***	0.048***	0.045***	0.042***	0.041***	0.038***
	(69.157)	(23.898)	(36.611)	(16.286)	(17.011)	(17.800)	(18.619)	(23.854)	(25.779)	(24.912)	(24.078)	(22.968)	(21.609)
Observations	54,642	1,455	1,617	2,765	4,060	4,242	4,644	5,248	5,705	6,066	6,391	6,336	6,113

Table 7: Replicating Fund Performance: *Spread* by Family

This table decomposes the difference in monthly performance of RFs and their corresponding TDFs by fund family. Panel A reports the average total monthly return *Spread* ($r_{j,t}^{RF} - r_{j,t}^{TDF}$) for Vanguard, Fidelity, T. Rowe Price, American Funds, TIAA-CREF, JP Morgan, and 56 other sponsors included in Other. Panel B decomposes the average total return spread into three components. The first component is *Timing* which is calculated as the average of $r_{j,t}^{IF} - r_{j,t}^{TDF} - \left(exp\_ratio_{j,\tau}^{TDF}\right)/12$. The second component is *Active* which is calculated as the average of $r_{j,t}^{RF} - r_{j,t}^{IF} + \left(exp\_ratio_{j,\tau}^{RF}\right)/12 - \left(exp\_ratio_{j,\tau}^{IF}\right)/12$. The third component is *Fee Gap* which is calculated as the average of $\left(exp\_ratio_{j,\tau}^{TDF}\right)/12 + \left(exp\_ratio_{j,\tau}^{IF}\right)/12 - \left(exp\_ratio_{j,\tau}^{RF}\right)/12$. ***,**,* indicate statistical significance at the 1%, 5%, and 10% and t-statistics are reported below each reported average.

Panel A: Monthly Return Spread

	(1) Vanguard	(2) Fidelity	(3) T. Rowe Price	(4) American Funds	(5) TIAA-CREF	(6) JP Morgan	(7) Other
Spread	0.001	0.069***	0.028***	0.109***	0.038***	0.010	0.107***
	(0.167)	(17.893)	(4.604)	(8.027)	(7.850)	(1.116)	(26.565)
Observations	1,888	9,492	3,032	1,292	2,586	1,892	34,460

Panel B: Monthly Return Spread Decomposition

	(1) Vanguard	(2) Fidelity	(3) T. Rowe Price	(4) American Funds	(5) TIAA-CREF	(6) JP Morgan	(7) Other
Timing	-0.009**	-0.002**	0.002	-0.001	0.012***	-0.021***	0.008***
	(-2.000)	(-2.324)	(0.914)	(-0.710)	(4.009)	(-4.196)	(3.898)
Active	0.004***	0.037***	-0.025***	0.049***	0.001	-0.016*	0.045***
	(2.785)	(9.682)	(-4.355)	(3.642)	(0.130)	(-1.761)	(12.051)
Fee Gap	0.006***	0.034***	0.051***	0.061***	0.026***	0.047***	0.055***
	(104.275)	(114.882)	(284.472)	(215.855)	(88.484)	(131.470)	(352.666)
Observations	1,888	9,492	3,032	1,292	2,586	1,892	34,460

Table 8: Performance Flow Regressions

This table examines the relation between TDF performance and future flows. We estimate:

$$\text{Flow}_{i,j,t} = a_j + b_t + \beta Spread_{i,j,t-1} + \Gamma Z_{i,j,t-1} + \epsilon_{i,j,t}$$

in which $\text{Flow}_{i,j,t}$ is a measure of annual flows, a_j and b_t are vintage and year fixed effects, $Spread_{i,j,t}$ is *Spread* for TDF i of vintage j in year t, and $Z_{i,j,t}$ decomposes returns into alphas and systematic returns (following Barber et al. (2016)). $Flow_{i,j,t}$ is calculated using two different specifications. First, we use annual market-share-change for each TDF within their vintage (e.g., 2030) following Spiegel and Zhang (2013). Second, we use annual net percentage flow, which is calculated by aggregating each TDF's monthly flow (in dollars) and dividing by its beginning of year AUM. Alphas are based on the 5-factor model of Balduzzi and Reuter (2019) which uses the CRSP U.S. value-weighted market index, the MSCIWorld Index excluding the United States, the Barclays U.S. Aggregate Bond Index, the Barclays Global Aggregate excluding the United States, and the GSCI Commodity Index. Annual alphas and systematic factors returns are calculated by compounding monthly returns over twelve months. Monthly alphas and systematic returns are based on monthly realized factor returns and factor loading estimated using daily returns over the past year. We require at least 120 daily returns to estimate factor loadings. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using standard errors clustered by fund family and year and t-statistics are reported in parentheses.

	MarketShareChange, year t		PercentageFlow, year t	
	(1)	(2)	(3)	(4)
Spread, year $t-1$	0.005	0.007	-2.228**	-1.023
	(0.316)	(0.792)	(-2.658)	(-1.167)
MarketShareChange, year $t-1$	0.510***	0.510***		
	(4.794)	(4.802)		
PercentageFlow, year $t-1$			0.120***	0.119***
			(4.401)	(4.282)
Systematic return, year $t-1$		0.010		0.818
		(0.753)		(1.698)
5-factor alpha, year $t-1$		-0.007		3.399**
		(-0.222)		(2.594)
Year Fixed Effects	Yes	Yes	Yes	Yes
Target Date Fixed Effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.415	0.415	0.129	0.132
Observations	2,252	2,252	2,252	2,252

Table 9: PROF Portfolio Example

This table illustrates the construction of our $PROF_{50}$, $PROF_{20}$, $PROF_{13}$ and $PROF_{6}$ portfolios within the 2030 vintage. In this example, PROF holdings are determined using September 30, 2019 reported holdings of all 2030 vintage RFs.

			Consolidated Funds		
ETF Name	Ticker	$PROF_{50}$	$PROF_{20}$	$PROF_{13}$	$PROF_{6}$
Total Stock Market	VTI	19.16%	19.24%	24.49%	44.46%
Large Cap	VV	6.60%	8.39%	14.35%	
Dividend Appreciation	VIG	0.06%			
Mid Cap	VO	0.29%	3.47%	4.01%	
Mid Cap Growth	VOT	1.13%			
Mid Cap Value	VOE	0.82%			
Small Cap	VB	1.22%	2.12%	2.12%	
Small Cap Growth	VBK	0.41%			
Small Cap Value	VBR	0.36%			
Extended Market	VXF	0.13%			
Growth	VUG	5.28%	6.21%		
Value	VTV	3.94%	4.16%		
Financials	VFH	0.01%			
High Dividend Yield	VYM	1.27%			
S&P 500	VOO	1.32%	1.37%		
Health Care	VHT	0.01%			
Total Bond Market	BND	13.67%	14.65%	18.22%	20.31%
Mortgage Backed Securities	VMBS	0.63%			
Intermediate Term Corporate Bond	VCIT	2.41%	3.33%	3.38%	
Short-Term Corporate Bond	VCSH	0.19%			
Short-Term Treasuries	VGSH	0.39%	0.40%	0.53%	
Long-Term Bond	BLV	0.07%	1.15%		
Long-Term Corporate Bond	VCLT	0.01%			
Long Term Treasuries	VGLT	1.09%			
Intermediate Term Bond	BIV	1.29%	1.94%		
Intermediate Term Treasuries	VGIT	1.17%			
Short-Term Bond	BSV	0.56%	0.67%		
FTSE All-World Ex-US	VEU	11.28%	14.25%	14.28%	27.50%
Total International Stock	VXUS	2.38%			
FTSE All-World ex-US Small-Cap	VSS	0.82%			
Global ex-U.S. Real Estate	VNQI	0.01%			
Intl Dividend Appreciation	VIGI	0.29%			
FTSE Developed Markets	VEA	5.62%	7.64%	7.97%	
FTSE Pacific	VPL	0.00%			
FTSE Emerging Markets	VWO	2.83%	4.79%	4.79%	
Emerging Govt Bonds	VWOB	1.69%			
Total World Stock	VT	4.56%			
Energy	VDE	0.70%			
Intl High Dividend Yield	VYMI	0.02%			
FTSE Europe	VGK	0.33%	0.35%		
Total Intl Bond	BNDX	3.82%	3.82%	3.82%	4.61%
TIPS	VTIP	1.70%	1.70%	1.71%	2.79%
Real Estate	VNQ	0.34%	0.34%	0.34%	0.34%

Table 10: PROF Performance

This table examines the performance of PROFs that are formed by matching each TDF's holdings to the Vanguard sample ETF with the highest monthly return correlation. In Panel A, each PROF's monthly performance is regressed on the aggregate TDF's monthly return:

$$r_{v,t}^{PROF_z} = \alpha + \beta r_{v,t}^{Agg} + \epsilon_{v,t},$$

in which $r_{v,t}^{PROF_z}$ is the PROF return of vintage v in month t with $z \in \{6, 13, 20, 50\}$, $r_{v,t}^{Agg}$ is the aggregate return of TDFs of vintage v in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{v,t}^{Agg}$, and $\epsilon_{v,t}$ is the error term. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using White standard errors and t-statistics are reported below each estimated coefficient. For β, t-statistics are reported under the null that the coefficient is equal to 1. Each regression also reports the root mean squared error (RMSE) and the annualized appraisal ratio which is calculated as Appraisal Ratio $= \frac{\alpha\sqrt{12}}{RMSE}$ In Panel B, PROF's and aggregate TDFs' monthly returns are annualized and then used to calculate Sharpe ratios. In Panel C, we compare the PROFs to each TDF in the same vintage. We calculate

$$AnnualOutperformance_{z,v,j} = \left(\frac{\prod_{t \in T} (1 + r_{v,t}^{PROF_z})}{\prod_{t \in T, j \in X(v,t)} (1 + r_{j,t}^{TDF})} \right)^{12/T} - 1,$$

in which $z \in \{6, 13, 20, 50\}$ is the size of the choice set for the PROF, T is the set of months both TDF j and PROF$_z$ have returns, v is the vintage, j is the comparable TDF, $r_{v,t}^{PROF_z}$ is the return of PROF$_z$ in vintage v and in month t and $r_{j,t}^{TDF}$ is the return of TDF j in month t. For each PROF, we average annual outperformance across comparable TDFs on both an equal-weighted and a asset-weighted basis. We also calculate the percentage of positive outperformance measures on both an equal-weighted and a asset-weighted basis.

Panel A: Performance Relative to Aggregate TDFs

	$r_{j,t}^{PROF}$			
	PROF$_{50}$	PROF$_{20}$	PROF$_{13}$	PROF$_{6}$
$r_{j,t}^{Agg}$	1.036***	1.039***	1.040***	1.028***
	(10.676)	(11.552)	(10.770)	(6.789)
α	0.025***	0.019**	0.020**	0.016*
	(3.187)	(2.388)	(2.274)	(1.744)
RMSE	0.288	0.298	0.329	0.332
Appraisal Ratio	0.300	0.223	0.212	0.169
Adjusted R^2	0.994	0.993	0.992	0.992
Observations	1,737	1,737	1,737	1,737

Panel B: Annual Performance Statistics

	PROF$_{50}$	PROF$_{20}$	PROF$_{13}$	PROF$_{6}$	Aggregate TDF
Average Annualized Return	7.04%	6.99%	7.01%	6.88%	6.54%
Annualized Std. Dev. Of Monthly Returns	12.66%	12.7%	12.72%	12.58%	12.19%
Sharpe Ratio	0.556	0.55	0.551	0.547	0.537

Panel C: Lifetime Performance Comparison to Individual TDFs

	PROF$_{50}$	PROF$_{20}$	PROF$_{13}$	PROF$_{6}$	Aggregate TDF
Annual Outperformance (equal weight)	0.88%	0.92%	0.84%	0.80%	0.43%
Annual Outperformance (asset weight)	0.43%	0.39%	0.34%	0.30%	-0.07%
% Outperformance (equal weight)	91%	90%	88%	87%	68%
% Outperformance (asset weight)	89%	80%	76%	70%	37%

Off Target: On the Underperformance of Target-Date Funds

Internet Appendix

The Internet Appendix includes several robustness checks and additional results to support our findings in the main prose.

IA1 Vanguard ETFs and Index Funds

In this section, we detail the Vanguard ETFs and Vanguard Index Funds used throughout the paper. Table IA1 provides details of the Vanguard funds used in our analysis, providing the fund name (listed in alphabetical order), the ETF ticker, the date at which the ETF first surpassed $50 million in AUM (and is therefore available for matching in our analysis), the Admiral and Investor Share Class tickers, and the first dates on which the Admiral Share Class and Investor Share Class reported monthly returns to CRSP.

In our analysis of constructing RFs from Vanguard Index Funds (see Section IA2), we utilize the Admiral Share Class where possible. However, if necessary, we use the Investor Share Class until Vanguard introduced an Admiral Share Class for the fund (Admiral Share Class shares have lower fees but require a higher minimum investment). Using the Investor Share Class until an Admiral Share Class is available is consistent with existing work (see, for example, Berk and Van Binsbergen (2015)).

IA2 Replicating Funds Formed From Index Funds

In this section, we construct RFs using Vanguard Index Funds. There are reasons that some investors may prefer a RF constructed from open-end index funds. For example, using open-end index funds ensures that the investor purchases and redeems at net asset value (NAV). Conversely, using ETFs rather than open-end index funds involves an additional source of uncertainty: the ETF premium/discount to NAV. Specifically, purchasing or selling ETF shares occurs at a secondary market price, not at NAV. While ETF premiums and discounts (relative to NAV) are generally small, they are not zero.

We construct RFs using the Vanguard Index Funds in a similar fashion to how we construct the RFs using ETFs. Each quarter, we match each TDF reported holding to the Vanguard Index Fund with the highest monthly return correlation. Additionally, we use TDFs' quarterly holdings reports to establish portfolio weights. Table IA2 reports results for regressions similar to those in Section 4 using RFs formed from Vanguard index funds. It is most natural to compare the results in Table IA2 to the results in Table 4. As can be seen by comparing the two tables, the results are qualitatively the same.

IA3 Replicating Funds Using All Available ETFs

Our main analysis focuses on RFs that are constructed using a set of 50 Vanguard ETFs. In this extension, our sample starts with the set of all ETF that have monthly returns and quarterly characteristics reported by CRSP. To mitigate the impact of illiquidity and possible non-synchronous prices due to infrequent trading, we limit our sample to ETFs that have achieved at least $50 million in assets (i.e., an ETF is included in our sample after the date on which it first exceeds $50 million). We also exclude ETFs that are backed by derivative securities rather than physical assets or are flagged as using leverage in order to avoid leveraged ETFs and inverse ETFs. We also filter out ETFs backed by derivative securities as a practical consideration as some investors may have an aversion to investing in synthetic products. We obtain data from Bloomberg to determine the dates that ETFs surpass the $50 million assets under management (AUM) threshold, ETFs' uses of derivative securities, and ETFs' uses of leverage. Our resulting broad ETF sample includes 1,320 ETFs.

To parallel our example from the main analysis, Table IA3 shows the matches of the Fidelity Freedom 2030 Fund to our broad sample of ETFs. Using the 50 Vanguard ETFs, the Large Cap Value Index Fund (FIOOX) matches to the Vanguard Value Index Fund ETF (VTV) with a correlation of 0.99, whereas when using the broader set of ETFs, FIOOX matches to the iShares Russell 1000 Value ETF (IWD) with a correlation of 1.00. By using a broader set of ETFs, the correlations between the underlying holdings and their matched ETFs are consistently higher than those reported in Table 2. For example, using our broad ETF sample, the average correlation is 0.96 (versus 0.90) and the lowest correlation is 0.78 (versus 0.30).

Table IA4 confirms that using our broad ETF sample results in higher correlations between ETFs and their matched underlying TDF holdings. Of the 2,809 unique ETF-holding-ETF matches, 466 (17% of the matches) had a monthly return correlation greater than 0.99. Correlations between TDF-holdings' monthly returns and their matched ETFs' monthly returns are generally larger for index funds. This is intuitive as index funds are typically a type of "pure play" strategy (e.g., S&P

Table IA1: Vanguard ETF and Index Fund Universe

This table lists the Vanguard ETFs and index funds used in our analysis. For each Vanguard fund, the ETF share class ticker appears along with the first date at which the ETF reached $50 million in assets under management. For each Vanguard fund, the Admiral Share Class also appears with the first date of reported monthly returns. For some Vanguard funds, the Investor Share Class also appears with the first date of reported monthly returns. In any analysis using the Vanguard index funds, we use the Investor Share Class until Vanguard introduced an Admiral Share Class for the fund (Admiral Share Class shares have lower fees but require a higher minimum investment).

Vanguard Fund Name	ETF Ticker	ETF $50 Million Date	Admiral Ticker	Admiral Start Date	Investor Ticker	Investor Start Date
Communication Services Index Fund	VOX	03/28/2006	VTCAX	04/29/2005	No Fund	
Consumer Discretionary Index Fund	VCR	05/05/2006	VCDAX	08/31/2005	No Fund	
Consumer Staples Index Fund	VDC	05/03/2005	VCSAX	02/27/2004	No Fund	
Developed Markets Index Fund	VEA	08/01/2007	VTMGX	09/30/1999	VDVIX	01/31/2014
Dividend Appreciation Index Fund	VIG	04/27/2006	VDADX	01/31/2014	VDAIX	05/31/2006
Emerging Markets Govt Bd Idx	VWOB	06/04/2013	VGAVX	07/31/2013	VGOVX	07/31/2013
Emerging Markets Stock Index Fund	VWO	04/25/2005	VEMAX	07/31/2006	VEIEX	06/30/1994
Energy Index Fund	VDE	02/09/2005	VENAX	11/30/2004	No Fund	
European Stock Index Fund	VGK	03/22/2005	VEUSX	09/28/2001	VEURX	07/31/1990
Extended Market Index Fund	VXF	08/19/2003	VEXAX	12/29/2000	VEXMX	01/29/1988
Financials Index Fund	VFH	08/23/2005	VFAIX	03/31/2004	No Fund	
FTSE All-World ex US Index Fund	VEU	03/27/2007	VFWAX	10/31/2011	VFWIX	04/30/2007
FTSE All-World ex-US Small-Cap Index	VSS	05/08/2009	VFSAX	03/29/2019	VFSVX	05/29/2009
FTSE Pacific Fund	VPL	05/02/2005	VPADX	09/28/2001	VPACX	07/31/1990
Global ex-US Real Estate Index Fd	VNQI	12/17/2010	VGRLX	03/31/2011	VGXRX	12/31/2010
Growth Index Fund	VUG	02/06/2004	VIGAX	12/29/2000	VIGRX	12/31/1992
Health Care Index Fund	VHT	12/28/2004	VHCIX	03/31/2004	No Fund	
High Dividend Yield Index Fund	VYM	02/21/2007	VHYAX	03/29/2019	VHDYX	12/29/2006
Industrials Index Fund	VIS	02/02/2006	VINAX	06/30/2006	No Fund	
Information Technology Index Fund	VGT	08/23/2005	VITAX	04/30/2004	No Fund	
Intermediate-Term Bond Index Fund	BIV	07/02/2007	VBILX	12/31/2001	VBIIX	03/31/1994
Intermediate-Term Corp Bond Idx Fund	VCIT	01/25/2010	VICSX	04/30/2010	No Fund	
Intermediate-Term Treasury Index Fd	VGIT	07/06/2011	VSIGX	09/30/2010	No Fund	
Internatl High Div Yield Index Fund	VYMI	06/23/2016	VIHAX	04/29/2016	VIHIX	04/29/2016
Intl Dividend Appreciation Index Fund	VIGI	06/08/2016	VIAAX	04/29/2016	VIAIX	04/29/2016
Large-Cap Index Fund	VV	01/30/2004	VLCAX	03/31/2004	VLACX	02/27/2004
Long-Term Bond Index Fund	BLV	07/20/2007	VBLAX	03/29/2019	VBLTX	03/31/1994
Long-Term Corporate Bond Idx Fund	VCLT	08/24/2010	VLTCX	05/28/2010	No Fund	
Long-Term Treasury Index Fund	VGLT	07/29/2010	VLGSX	04/30/2010	No Fund	
Materials Index Fund	VAW	08/05/2005	VMIAX	03/31/2004	No Fund	
Mid-Cap Growth Index Fund	VOT	03/13/2007	VMGMX	10/31/2011	VMGIX	09/29/2006
Mid-Cap Index Fund	VO	12/03/2004	VIMAX	12/31/2001	VIMSX	06/30/1998
Mid-Cap Value Index Fund	VOE	11/29/2006	VMVAX	10/31/2011	VMVIX	09/29/2006
Mortgage-Backed Secs Idx Fund	VMBS	03/15/2011	VMBSX	01/29/2010	No Fund	
Real Estate Index Fund	VNQ	10/22/2004	VGSLX	12/31/2001	VGSIX	06/28/1996
S&P 500 Index Fund	VOO	10/07/2010	VFIAX	12/29/2000	VFINX	09/30/1976
Short-Term Bond Index Fund	BSV	04/24/2007	VBIRX	12/31/2001	VBISX	03/31/1994
Short-Term Corporate Bond Idx Fd	VCSH	12/23/2009	VSCSX	12/31/2010	No Fund	
Short-Term Treasury Index Fund	VGSH	05/10/2010	VSBSX	02/26/2010	No Fund	
Sht-Term Inflation-Protected Sec Idx	VTIP	11/28/2012	VTAPX	11/30/2012	VTIPX	11/30/2012
Small-Cap Growth Index Fund	VBK	04/22/2004	VSGAX	10/31/2011	VISGX	06/30/1998
Small-Cap Index Fund	VB	02/10/2004	VSMAX	12/29/2000	NAESX	12/31/1963
Small-Cap Value Index Fund	VBR	01/21/2005	VSIAX	10/31/2011	VISVX	07/31/1998
Total Bond Market Index Fund	BND	04/26/2007	VBTLX	12/31/2001	VBMFX	01/30/1987
Total International Bond Index Fund	BNDX	06/25/2013	VTABX	06/28/2013	VTIBX	06/28/2013
Total International Stock Index Fund	VXUS	02/24/2011	VTIAX	01/31/2011	VGTSX	05/31/1996
Total Stock Market Index Fund	VTI	06/06/2001	VTSAX	12/29/2000	VTSMX	05/29/1992
Total World Stock Index Fund	VT	09/18/2008	VTWAX	03/29/2019	VTWSX	08/29/2008
Utilities Index Fund	VPU	12/17/2004	VUIAX	05/28/2004	No Fund	
Value Index Fund	VTV	02/11/2004	VVIAX	12/29/2000	VIVAX	12/31/1992

Table IA2: Replicating Fund Performance: Vanguard Index Funds

This table exams the performance of replicating funds that are formed by matching each TDF's holdings to the Vanguard Index Fund with the highest monthly return correlation. Each replicating fund's monthly performance is regressed on its corresponding TDF reported return:

$$r_{j,t}^{RF} = \alpha + \beta r_{j,t}^{TDF} + \epsilon_{j,t}$$

in which $r_{j,t}^{RF}$ is the replicating fund return of TDF j in month t, $r_{j,t}^{TDF}$ is the return of TDF j in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{j,t}^{TDF}$, and $\epsilon_{j,t}$ is the error term. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using White standard errors and t-statistics are reported below each estimated coefficient. For β, t-statistics are reported under the null that the coefficient is equal to 1. Each regression also reports the root mean squared error (RMSE) and the annualized appraisal ratio which is calculated as Appraisal Ratio = $\frac{\alpha\sqrt{12}}{RMSE}$.

	$r_{j,t}^{RF}$		
	(1) Full Sample	(2) 2008-2013	(3) 2014-2019
$r_{j,t}^{TDF}$	1.054***	1.069***	1.034***
	(13.637)	(11.308)	(7.910)
α	0.055***	0.047***	0.064***
	(6.598)	(2.751)	(7.587)
RMSE	0.626	0.857	0.457
Appraisal Ratio	0.303	0.188	0.487
Adjusted R^2	0.962	0.958	0.968
Observations	54,642	18,783	35,859

Table IA3: Fidelity Freedom 2030 Fund Example: Broad ETF Sample

This table displays the Fidelity Freedom 2030 Fund's 29 mutual fund constituents according to their s12 reported holdings on September 30, 2019. Fund positions are reported in millions and correlation reports the monthly return correlation between the mutual fund holding and the matched Vanguard ETF.

Position	Fund	Fund Name	ETF	ETF Name	Correlation
$5,918	FSIGX	Investment Grade Bond Fund	NUAG	Nuveen Enhanced Yield US Aggregate Bond	0.98
$2,656	FEMSX	Emerging Markets Opportunities Fund	VWO	Vanguard Emerging Markets Stock Index Fund	0.98
$2,471	FIGSX	International Growth Fund	EFG	iShares MSCI EAFE Growth	0.97
$2,448	FINVX	International Value Fund	VEA	Vanguard Developed Markets Index Fund	0.98
$2,285	FDMLX	Intrinsic Opportunities Fund	CRBN	iShares MSCI ACWI Low Carbon Target	0.92
$2,016	FGLGX	Large Cap Stock Fund	EXT	WisdomTree US Total Market Fund	0.98
$1,858	FCGSX	Growth Company Fund	ONEQ	Fidelity NASDAQ Composite Index Tracking Stock	0.97
$1,572	FBLEX	Stock Selector Large Cap Value Fund	IWD	iShares Russell 1000 Value	0.99
$1,310	FSIPX	Inflation-Protected Bond Index Fund	TDTF	FlexShares iBoxx 5-Year Target Duration TIPS Index Fund	0.98
$1,145	FCSSX	Commodity Strategy Fund	GCC	WisdomTree Continuous Commodity Index Fund	0.94
$1,114	FNKLX	Value Discovery Fund	IWD	iShares Russell 1000 Value	0.98
$1,028	FVWSX	Opportunistic Insights Fund	JKE	iShares Morningstar Large-Cap Growth	0.97
$920	FSBDX	Blue Chip Growth Fund	ONEQ	Fidelity NASDAQ Composite Index Tracking Stock	0.96
$823	FSOPX	Small Cap Opportunities Fund	VB	Vanguard Small-Cap Index Fund	0.99
$815	FTLTX	Long-Term Treasury Bond Fund	VGLT	Vanguard Long-Term Treasury Index Fund	0.99
$766	FSOSX	Overseas Fund	DFE	WisdomTree Europe SmallCap Dividend Fund	0.92
$581	FIOOX	Large Cap Value Index Fund	IWD	iShares Russell 1000 Value	1.00
$540	FSAEX	All-Sector Equity Fund	SCHX	Schwab US Large-Cap	0.99
$527	FSTSX	International Small Cap Fund	SCZ	iShares MSCI EAFE Small-Cap	0.97
$288	FCNSX	Canada Fund	EWC	iShares MSCI Canada	0.98
$287	FHKFX	Emerging Markets Fund	EEMS	iShares MSCI Emerging Markets Small-Cap	0.97
$275	FJACX	Small Cap Discovery Fund	FNDA	Schwab Fundamental US Small Company Index	0.96
$241	FSHNX	High Income Fund	JNK	SPDR Bloomberg Barclays High Yield Bond	0.98
$216	FGNXX	Government Money Market Fund	BIL	SPDR Bloomberg Barclays 1-3 Month T-Bill	0.99
$204	FEDCX	Emerging Markets Debt Fund	EMHY	iShares Emerging Markets High Yield Bond	0.98
$156	FSREX	Real Estate Income Fund	FREL	Fidelity MSCI Real Estate Index	0.89
$90	FYBTX	Short-Term Credit Fund	SPSB	SPDR Portfolio Short Term Corporate Bond	0.96
$40	FFHCX	Floating Rate High Income Fund	BLHY	Virtus Newfleet Dynamic Credit	0.93
$17	FCDSX	International Credit Fund	CORP	PIMCO Investment Grade Corporate Bond Index	0.90

500 value) and do not deviate meaningfully from benchmark weights (unlike actively managed mutual funds). As such, there is typically an ETF offered with the same benchmark and nearly identical returns. While the match quality based on monthly return correlations is not as strong for non-index funds, 81% of matches exhibit monthly return correlations greater than 0.90 and 61% of matches exhibit return correlations greater than 0.95.

Table IA5 replicates our main analysis in Table 4 using our broad set of ETFs to form replicating portfolios. In general, the results are very similar, though subtle differences exist. Compared to our set of 50 Vanguard ETFs, our broad ETF sample has the benefit that the match quality between a TDF's holdings and its matched ETFs will be stronger. However, it also has the cost that non-Vanguard ETFs are characterized by relatively higher fees. The benefit from better match quality can be seen in higher R^2 values, which are about 1% higher than those reported in Table 4. The cost of higher fees can be seen in the α estimates, which are about 1 to 2 bps lower on average.

IA4 Replicating Fund Performance By Vintage and Fund Family

In this section, we examine the performance of our RFs based on their vintages (e.g., 2030 target date retirement) and their fund families (e.g., Fidelity). The analysis provides additional evidence that RFs achieve superior performance to their TDF counterparts. Before reporting the regression results, Table IA6 shows summary statistics by vintage (using end of 2019 values), highlighting that the 2030 is the largest vintage and most assets are concentrated among the 2020–2040 vintages.

We run regressions similar to those reported in Table 4, that is, we regress the monthly return of our RFs formed from Vanguard ETFs on the respective TDF's monthly return from the same month. First, Table IA7 reports regressions based on fund vintage. The results show positive values of α for each vintage. The α estimates are not statistically significant for the earliest vintages (2000, 2005, and 2010) nor the latest vintage (2065), which is partially attributable to smaller samples and limited power. Nevertheless, the results are consistent with those reported in Section 4: RFs outperform their TDF counterparts and exhibit low tracking error which yields sizeable appraisal ratios.

Table IA8 reports the results by fund family. Specifically, we focus on the six largest TDF

Table IA4: Correlations of TDF Holdings Matched to Broad Sample ETFs

This table summarizes the match quality of TDF holdings to our broad sample of ETFs based on their return correlations. Correlations are based on all overlapping monthly returns between 2008 and 2019. We use the CRSP index fund flag to categorize TDF holdings.

	Correlation Range									
	≤0.5	0.50-0.75	0.75-0.80	0.80-0.85	0.85-0.90	0.90-0.925	0.925-0.95	0.95-0.975	0.975-0.99	>.99
All Funds										
Number of Funds	2,412									
Fund Distribution	13	29	31	43	78	115	193	520	738	652
Percent	0.54%	1.20%	1.29%	1.78%	3.23%	4.77%	8.00%	21.56%	30.60%	27.03%
Index Funds										
Number of Funds	511									
Fund Distribution		2				2	3	16	59	429
Percent	0.00%	0.39%	0.00%	0.00%	0.00%	0.39%	0.59%	3.13%	11.55%	83.95%
Non-Index Funds										
Number of Funds	1,901									
Fund Distribution	13	27	31	43	78	113	190	504	679	223
Percent	0.68%	1.42%	1.63%	2.26%	4.10%	5.94%	9.99%	26.51%	35.72%	11.73%

Table IA5: Replicating Fund Performance: Broad ETF Sample

This table examines the performance of replicating funds that are formed by matching each TDF's holdings to the broad sample ETF with the highest monthly return correlation. Each replicating fund's monthly performance is regressed on its corresponding TDF reported return:

$$r_{j,t}^{RF} = \alpha + \beta r_{j,t}^{TDF} + \epsilon_{j,t}$$

in which $r_{j,t}^{RF}$ is the replicating fund return of TDF j in month t, $r_{j,t}^{TDF}$ is the return of TDF j in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{j,t}^{TDF}$, and $\epsilon_{j,t}$ is the error term. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using White standard errors and t-statistics are reported below each estimated coefficient. For β, t-statistics are reported under the null that the coefficient is equal to 1. Each regression also reports the root mean squared error (RMSE) and the annualized appraisal ratio which is calculated as Appraisal Ratio = $\frac{\alpha\sqrt{12}}{RMSE}$.

	$r_{j,t}^{RF}$		
	(1) Full Sample	(2) 2008-2013	(3) 2014-2019
$r_{j,t}^{TDF}$	1.012***	1.020***	1.001***
	(3.432)	(3.764)	(0.156)
α	0.041***	0.033**	0.048***
	(6.569)	(2.453)	(7.829)
RMSE	0.492	0.681	0.353
Appraisal Ratio	0.287	0.168	0.467
Adjusted R^2	0.974	0.970	0.979
Observations	54,642	18,783	35,859

Table IA6: Vintage TDF Sample as of 2019

This table summarizes our sample by vintage as of 2019. A TDF is included in our sample if it reports quarterly holdings and if both the TDF and the TDF's fund holdings appear in the CRSP mutual fund data base with monthly returns and net asset values. TDF Fee is the reported expense ratio of the TDF, and Holdings Fee is the fees collected on the funds held by the TDF. # Holdings reports the number of individual funds held by a TDF and % Index reports the fraction of held funds which are index funds. Reported values are based on each fund's last monthly observation in 2019. Aggregated TDF AUM numbers are reported in billions.

				Equal-Weighted Average				Asset-Weighted Average			
Vintage	Funds	Classes	AUM	TDF Fee	Holdings Fee	# Holdings	% Index	TDF Fee	Holdings Fee	# Holdings	% Index
2005	13	45	\$3	0.21%	0.12%	21.3	25.49%	0.21%	0.28%	22.5	21.27%
2010	29	116	\$19	0.17%	0.23%	20.2	31.81%	0.25%	0.23%	21.6	17.93%
2015	41	153	\$57	0.22%	0.23%	17.6	42.34%	0.16%	0.18%	14.5	55.85%
2020	61	245	\$164	0.24%	0.29%	18.8	38.61%	0.17%	0.20%	16.1	48.92%
2025	52	205	\$204	0.24%	0.26%	18.3	40.32%	0.17%	0.20%	15.0	52.14%
2030	60	242	\$220	0.24%	0.33%	17.7	37.66%	0.19%	0.23%	15.8	46.38%
2035	51	205	\$177	0.25%	0.30%	16.5	40.29%	0.19%	0.21%	14.0	51.26%
2040	56	225	\$166	0.23%	0.33%	17.3	40.63%	0.19%	0.24%	14.7	47.10%
2045	51	206	\$123	0.25%	0.30%	16.1	44.05%	0.18%	0.22%	13.3	53.34%
2050	55	220	\$99	0.23%	0.33%	16.5	43.20%	0.18%	0.23%	13.6	50.47%
2055	49	203	\$50	0.23%	0.31%	15.7	41.48%	0.18%	0.23%	13.3	51.44%
2060	42	176	\$16	0.21%	0.32%	16.4	40.73%	0.16%	0.22%	12.0	57.78%
2065	9	29	\$1	0.21%	0.20%	14.6	50.38%	0.02%	0.14%	4.9	94.62%

Table IA7: Replicating Fund Performance: By Vintage

This table exams the performance of replicating funds, by vintage (e.g., 2030), that are formed by matching each TDF's holdings to the Vanguard sample ETF with the highest monthly return correlation. Each replicating fund's monthly performance is regressed on its corresponding TDF reported return:

$$r_{j,t}^{RF} = \alpha + \beta r_{j,t}^{TDF} + \epsilon_{j,t}$$

in which $r_{j,t}^{RF}$ is the replicating fund return of TDF j in month t, $r_{j,t}^{TDF}$ is the return of TDF j in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{j,t}^{TDF}$, and $\epsilon_{j,t}$ is the error term. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using White standard errors and t-statistics are reported below each estimated coefficient. For β, t-statistics are reported under the null that the coefficient is equal to 1. Each regression also reports the root mean squared error (RMSE) and the annualized appraisal ratio which is calculated as Appraisal Ratio = $\frac{\alpha\sqrt{12}}{RMSE}$.

	$r_{j,t}^{RF}$													
	(1) 2000	(2) 2005	(3) 2010	(4) 2015	(5) 2020	(6) 2025	(7) 2030	(8) 2035	(9) 2040	(10) 2045	(11) 2050	(12) 2055	(13) 2060	(14) 2065
$r_{j,t}^{TDF}$	1.163***	1.105***	1.171***	1.135***	1.097***	1.078***	1.066***	1.046***	1.037***	1.027***	1.021***	1.020***	1.007***	0.942***
	(6.385)	(7.964)	(10.954)	(12.915)	(13.392)	(14.385)	(12.013)	(12.503)	(8.850)	(6.285)	(2.180)	(5.281)	(1.345)	(-1.765)
α	0.010	0.016	0.013	0.033**	0.052***	0.052***	0.056***	0.053***	0.058***	0.054***	0.055***	0.045***	0.037***	0.111
	(0.382)	(0.911)	(0.767)	(2.450)	(4.478)	(5.451)	(5.034)	(6.523)	(5.828)	(6.649)	(3.488)	(5.513)	(3.324)	(1.387)
RMSE	0.187	0.364	0.811	0.670	0.688	0.484	0.668	0.441	0.647	0.446	0.731	0.370	0.354	0.371
Appraisal Ratio	0.183	0.152	0.057	0.169	0.261	0.370	0.292	0.420	0.312	0.418	0.260	0.426	0.363	1.032
Adjusted R^2	0.974	0.968	0.911	0.938	0.941	0.974	0.962	0.984	0.969	0.985	0.957	0.987	0.986	0.986
Observations	144	1,463	3,896	5,145	6,226	5,369	5,949	5,151	5,553	5,030	5,102	3,628	1,899	87

providers: Vanguard, Fidelity, T. Rowe Price, American Funds, TIAA-CREF, and JP Morgan. These six TDF providers collectively controlled 91% of 2019 TDF assets. The remaining 56 TDF sponsors are aggregated together and labeled "Other." The results in Table IA8 echo the main insight from Section 4: RFs exhibit outperformance and low-measures of tracking error. Table IA8 shows that α is generally positive valued and statistically significant. For example, the estimate of α for Fidelity funds is 4.2 bps a month (50 bps a year). There are notable exceptions, however. First, the estimate of α for Vanguard is -0.1 bps per month (-1 bps per year) and the value is not statistically significant. As such, while Vanguard has a tendency to hold relatively more expensive share classes in their TDFs, doing so does not yield statically significant underperformance in their TDFs as compared to our RFs based on regression analysis. Second, JP Morgan TDFs *outperform* our RFs in this sample (-1.6 bps per month or -19 bps per year), although the difference is not significant different from zero.

IA5 Main Results Using Rolling Correlations

In this section, we reproduce our main results from Tables 4 and 6. Rather than using monthly in-sample data to compute return correlations, we measure return correlations using daily returns over the prior year. We require at least 120 overlapping daily returns to calculate return correlations. Tables IA9 and IA10 show that our main results and conclusions are unchanged using the alternative matching procedure.

Table IA8: Replicating Fund Performance: By Family

This table exams the performance of replicating funds, by fund family (e.g., Fidelity), that are formed by matching each TDF's holdings to the Vanguard sample ETF with the highest monthly return correlation. Each replicating fund's monthly performance is regressed on its corresponding TDF reported return:

$$r_{j,t}^{RF} = \alpha + \beta r_{j,t}^{TDF} + \epsilon_{j,t}$$

in which $r_{j,t}^{RF}$ is the replicating fund return of TDF j in month t, $r_{j,t}^{TDF}$ is the return of TDF j in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{j,t}^{TDF}$, and $\epsilon_{j,t}$ is the error term. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using White standard errors and t-statistics are reported below each estimated coefficient. For β, t-statistics are reported under the null that the coefficient is equal to 1. Each regression also reports the root mean squared error (RMSE) and the annualized appraisal ratio which is calculated as Appraisal Ratio = $\frac{\alpha\sqrt{12}}{RMSE}$.

	$r_{j,t}^{RF,UL}$						
	(1) Vanguard	(2) Fidelity	(3) T. Rowe Price	(4) American Funds	(5) TIAA-CREF	(6) JP Morgan	(7) Other
$r_{j,t}^{TDF}$	1.001***	0.964***	0.997***	1.113***	1.012***	0.981***	1.040***
	(0.485)	(-15.243)	(-0.720)	(22.071)	(6.082)	(-4.579)	(11.702)
α	-0.003	0.031***	0.025***	0.022*	0.051***	-0.009	0.064***
	(-0.562)	(7.530)	(3.532)	(1.692)	(9.407)	(-0.984)	(8.575)
RMSE	0.214	0.358	0.321	0.428	0.256	0.359	1.271
Appraisal Ratio	-0.050	0.305	0.274	0.179	0.684	-0.089	0.175
Adjusted R^2	0.995	0.983	0.990	0.979	0.994	0.985	0.858
Observations	1,888	9,492	3,032	1,292	2,586	1,892	34,460

Table IA9: Replicating Fund Performance

This table examines the performance of replicating funds that are formed by matching each TDF's holdings to the Vanguard sample ETF with the highest monthly return correlation. Each replicating fund's monthly performance is regressed on its corresponding TDF reported return:

$$r_{j,t}^{RF} = \alpha + \beta r_{j,t}^{TDF} + \epsilon_{j,t}$$

in which $r_{j,t}^{RF}$ is the replicating fund return of TDF j in month t, $r_{j,t}^{TDF}$ is the return of TDF j in month t, α is the intercept (i.e., the alpha), β is the coefficient on $r_{j,t}^{TDF}$, and $\epsilon_{j,t}$ is the error term. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels using White standard errors and t-statistics are reported below each estimated coefficient. For β, t-statistics are reported under the null that the coefficient is equal to 1. Each regression also reports the root mean squared error (RMSE) and the annualized appraisal ratio which is calculated as Appraisal Ratio = $\frac{\alpha\sqrt{12}}{RMSE}$.

	$r_{j,t}^{RF}$		
	(1) Full Sample	(2) 2008-2013	(3) 2014-2019
$r_{j,t}^{TDF}$	1.043***	1.053***	1.030***
	(11.404)	(9.104)	(7.014)
α	0.054***	0.058***	0.056***
	(6.343)	(3.354)	(6.129)
RMSE	0.616	0.804	0.487
Appraisal Ratio	0.303	0.248	0.396
Adjusted R^2	0.962	0.962	0.963
Observations	54,415	18,780	35,635

Table IA10: Replicating Fund Performance: *Spread*

This table decomposes the difference in monthly performance of RFs and their corresponding TDFs. Panel A reports the average total monthly return *Spread* ($r_{j,t}^{RF} - r_{j,t}^{TDF}$) for the full sample and each year from 2008 to 2019. Panel B decomposes the average total return spread into three components. The first component is *Timing* which is calculated as the average of $r_{j,t}^{IF} - r_{j,t}^{TDF} - \left(exp\_ratio_{j,\tau}^{TDF}\right)/12$. The second component is *Active* which is calculated as the average of $r_{j,t}^{RF} - r_{j,t}^{IF} + \left(exp\_ratio_{j,\tau}^{RF}\right)/12 - \left(exp\_ratio_{j,\tau}^{IF}\right)/12$. The third component is *Fee Gap* which is calculated as the average of $\left(exp\_ratio_{j,\tau}^{TDF}\right)/12 + \left(exp\_ratio_{j,\tau}^{IF}\right)/12 - \left(exp\_ratio_{j,\tau}^{RF}\right)/12$. ***,**,* indicate statistical significance at the 1%, 5%, and 10% and t-statistics are reported below each reported average.

Panel A: Monthly Return Spread

	(1) Full Sample	(2) 2008	(3) 2009	(4) 2010	(5) 2011	(6) 2012	(7) 2013	(8) 2014	(9) 2015	(10) 2016	(11) 2017	(12) 2018	(13) 2019
Spread	0.082***	0.125**	0.047	0.244***	-0.004	0.094***	0.119***	0.129***	0.011	0.132***	0.079***	-0.005	0.098***
	(10.784)	(2.528)	(0.841)	(4.926)	(-0.120)	(3.293)	(4.115)	(8.656)	(0.406)	(6.063)	(5.332)	(-0.294)	(5.589)
Observations	54,415	1,455	1,617	2,768	4,060	4,242	4,638	5,248	5,705	6,034	6,391	6,276	5,981

Panel B: Monthly Return Spread Decomposition

	(1) Full Sample	(2) 2008	(3) 2009	(4) 2010	(5) 2011	(6) 2012	(7) 2013	(8) 2014	(9) 2015	(10) 2016	(11) 2017	(12) 2018	(13) 2019
Timing	0.003	-0.062***	-0.010	0.004	-0.016	0.011	0.016	0.015***	0.001	0.010**	0.009***	0.008	-0.003
	(1.194)	(-4.121)	(-0.213)	(0.302)	(-1.368)	(1.494)	(1.398)	(2.864)	(0.155)	(2.115)	(3.104)	(0.653)	(-0.440)
Active	0.031***	0.114**	-0.016	0.181***	-0.041	0.032	0.055**	0.066***	-0.038	0.077***	0.028**	-0.054***	0.062***
	(4.098)	(2.345)	(-0.210)	(3.813)	(-1.141)	(1.160)	(2.076)	(4.808)	(-1.492)	(3.689)	(1.980)	(-2.667)	(3.857)
FeeGap	0.048***	0.072***	0.073***	0.060***	0.053***	0.051***	0.049***	0.049***	0.048***	0.045***	0.042***	0.041***	0.039***
	(68.671)	(23.822)	(37.452)	(16.404)	(16.880)	(17.450)	(18.518)	(23.715)	(25.894)	(24.586)	(23.857)	(22.761)	(21.303)
Observations	54,415	1,455	1,617	2,768	4,060	4,242	4,638	5,248	5,705	6,034	6,391	6,276	5,981

Off Target: On the Underperformance of Target-Date Funds

Data Appendix

The Data Appendix includes details regarding our data collection and choices made to form the TDF data set used in our empirical tests.

DA1 Data Appendix

DA1.1 Data Sources and Target-Date Fund Identification

We construct our Target-Date Fund (TDF) sample by combining data from several sources acquired via Wharton Research Data Services (WRDS). From the CRSP Mutual Funds data, we use *Monthly Returns and Net Asset Values*, *Daily Returns and Net Asset Values*, *Holdings*, and *Summary* data.[1] We link CRSP data to *Thomson Reuters Mutual Fund Holdings (s12)* data using *MFLINKS* data. Our sample includes data from 2006 to 2019.[2]

We identify 4,042 unique fundnos (crsp_fundno) in the CRSP *Summary* data with fund names that include target retirement dates 2000, 2005, ..., 2065.[3] Of those 4,042 unique fundnos, we identify 34 fundnos corresponding to 26 unique bond funds (e.g., the IShares IBonds Sep 2020 Term Muni Bond ETF). We identify the bond mutual funds by searching for the terms "bond", "municipal", "muni", or "coupon" in the fund name, and we remove them from our TDF sample. The remaining 4,008 unique fundnos are linked to 1,306 unique portnos (crsp_portno), and because of multiple share classes, result in 5,353 fundno-portno pairs. Because portnos sometimes change over time, we map portnos into unique TDFIds to determine the total number of TDFs in our sample (we note that TDFIds are unique to our study). We join two (or three) portnos together when they are linked by fundnos that change portnos over time. Because different share classes of the same TDF may relate to only one or multiple portnos, we ensure that each fundno (share class) is mapped to the same unique TDF identifier. Merging portnos together results in 935 unique TDFs. There are 586 TDFs matched to a unique portno, 341 TDFs that are matched to two portnos, and eight TDFs that are matched to three portnos.[4] When aggregating to the TDF level, we use our unique TDF identifiers linked to fundnos, as fundnos are not always mapped to portnos in CRSP data.

[1] Italics are used to denote names of data sets accessed via WRDS.

[2] We use fund characteristics data from 2004 in order to forward fill missing data and the holdings data include several months in 2020. The 2020 holdings data is sparsely populated, so we end our sample in 2019.

[3] We manually remove funds referring to the Russell 2000 index and related products.

[4] For example, the Fidelity Advisor Freedom 2030 TDF is represented by portnos 1002396 and 1021235. Portno 1002396 has data until September 2010 and portno 1021235 has data from September 2010 through the end of our sample.

To determine the size of our sample TDFs, we merge TDF level data to monthly returns and NAVs (net asset values) from CRSP *Mutual Funds - Monthly Returns and Net Asset Values.* When monthly returns are missing, we aggregate daily return data from CRSP *Mutual Funds - Daily Returns and Net Asset Values* to the monthly return level. To prevent discrepancies that originate due to how the data is downloaded from WRDS, we round returns to six decimal places.[5]

Table DA1 displays the number of TDFs, the number of underlying shares classes and the total AUM (in billions) of our sample TDFs by year. For comparison, the table also shows the number of TDFs and total AUM reported by the 2021 Investment Company Institute (ICI) Fact Book. The two samples are well aligned, as our annual sample AUM represents between 95% and 106% of the total assets reported by ICI.

DA1.2 Fund Characteristics

Data on fees and fund family are from CRSP *Mutual Funds - Summary* data. We forward fill fees (exp_ratio) and the percentage of cash held in the fund (per_cash) when those values are missing.[6] Many fund family names (mgmt_name) appear in the CRSP data using several variations, and several TDF series are acquired throughout our sample period. We combine fund family names if they are associated with the same fundno, portno or our self-identified TDFId. We also combine several fund family names based on manually identified mergers. Table DA2 details our consolidation of fund family names.

As fund characteristics are reported at the share class level, we aggregate those characteristics to the TDF level based on share class weights. To determine weights, we first merge quarterly fund characteristics data with monthly total net assets (mtna) from CRSP *Mutual Funds - Monthly Returns and Net Asset Values* data. Weights for each share class are formed based on average monthly total net assets, which we calculate by averaging the end of the current month's and prior month's values.[7]

[5] Data downloaded from WRDS as a CSV file have six decimal places, while data downloaded as a STATA data file have twelve decimal places.

[6] To forward fill fee and cash holdings data, we look back as far as 2004, two years before our sample begins.

[7] When either the current month's value or prior month's values is missing, we use the remaining available value to calculate average monthly total net assets.

Table DA1: TDF Sample Size

This table reports the number of TDFs, number of underlying share classes and the total AUM (in billions) of our sample TDFs by year. The last two columns present the number of TDFs and the total AUM (in billions) from Table 56 of the 2021 Investment Company Institute Fact Book.

	Our Sample			**ICI 2021 Fact Book**	
Year	TDFs	Share Classes	AUM	TDFs	AUM
2006	168	690	$112	181	$114
2007	252	1,086	$178	246	$183
2008	385	1,525	$154	339	$160
2009	446	1,684	$247	380	$256
2010	444	1,670	$329	378	$340
2011	466	1,739	$363	413	$376
2012	517	1,974	$467	429	$481
2013	527	2,106	$602	492	$618
2014	579	2,252	$683	542	$703
2015	630	2,458	$746	599	$763
2016	652	2,570	$848	641	$887
2017	680	2,603	$1,185	632	$1,116
2018	677	2,721	$1,060	685	$1,101
2019	684	2,760	$1,350	679	$1,396

After consolidating fund family names, our sample of 935 TDFs come from 62 fund families (42 fund families have active TDFs in 2019). Table DA3 provides summary statistics for each fund family. The summary statistics show that the majority of our sample is concentrated among relatively few fund families. The top 4 families hold 80% of 2019 TDF assets and the top 10 families hold 94% of 2019 TDF assets.

DA1.3 Holdings Data

We use CUSIPs provided in the CRSP and Thomson Reuters holdings data to map individual holdings to mutual funds and ETFs. In most cases, CUSIPs uniquely map to CRSP fundnos and are stable throughout our sample. In a small set of cases, funds change CUSIP (for example, after an acquisition), resulting in multiple CUSIPs mapping to the same CRSP fundno. In those cases, Bloomberg data are used to verify mappings on a case-by-case basis. For mutual fund and

Table DA2: Consolidated Fund Family Names

This table lists the alternative fund family names associated with each TDF fund family. The Master Family Name is one of the alternative names and is chosen arbitrarily. Only TDF fund families with multiple names are included in the table.

Master Family Name	Alternative Family Names
American Independence Financial Svcs Llc	Intrust Financial Services Inc Intrust Financial Services
BlackRock Inc	BlackRock Fund Advisors Barclays Global Fund Advisors State Farm Inv Mgmt State Farm Investment Mgmt Corporation
Charles Schwab Investment Management Inc	Charles Schwab Invstment Management Inc
Columbia Funds	Riversource Investments Llc Columbia Management Inv Advisers Llc Ameriprise Financial Inc Ridgeworth Funds Sti Classic Funds
Dbx Strategic Advisors Llc	Xshares Advisers Llc
Dws Investment Management Americas Inc	Deutsche Asset & Wealth Management Dws Investments Dws Scudder Deutsche Investment Mgmt Americas Inc Scudder Investments
Fidelity Management & Research Company	Fidelity Mgmt & Research Co
First Investors Management Co Inc	First Investors Management Company Inc
Goldman Sachs & Co/Gsam	Goldman Sachs & Co Madison Asset Management Llc
Invesco Funds	Aim Investments Aim Investments Invesco Aim Invesco Capital Management Llc Guggenheim Investments Guggenheim Funds Investment Advisors Llc Claymore Advisors Llc
Legg Mason	Legg Mason/Western Asset Management
Manning & Napier Advisors Inc	Manning & Napier Advisors Llc
Mml Investment Advisers Llc	Massmutual Life Insurance Company
Natixis Advisors Lp	Ngam Advisors Lp
Principal Global Investors Llc	Principal Management Corporation Principal Mgmt Corp
Sunamerica Asset Management Corp	Aig Sunamerica Asset Management Corp
T Rowe Price Associates Inc	T. Rowe Price Associates T. Rowe Price Associates Inc
Usaa Asset Management Company	Usaa Investment Manaagement Company Usaa Mutual Funds Victory Capital Management Inc Usaa Investment Management Company
Voya Investments Llc	Directed Services Llc Ing Investments Llc
Wells Fargo Fund Management Llc	Wells Fargo Funds Mgmt Wells Fargo Funds Management Llc
Xshares Advisers Llc	Xshares Advisors Llc

ETF holdings, we replace missing prices with CRSP net asset values (mnav), in both Thomson Reuters and CRSP holdings, when the price is missing and the mapping exists. Given our focus on replicating TDFs using ETFs, we do not map holdings to individual stock or bond data. For most TDFs, this does not significantly affect their holdings. However, a small set of TDFs hold a significant fraction of their assets in individual securities. To determine what fraction of a TDF's assets is held in other funds (mutual funds and ETFs), we aggregate the value of all reported holdings prior to dropping the individual securities holdings.

Thomson Reuters reports holdings on a quarterly basis (at quarter end) for 326 TDFs in our sample (based on MFLINKs) amounting to 10,664 holdings quarters. We drop 550 holdings quarters that do not match to any CRSP fund data. In almost all cases, the reported holdings are either very small or do not match to CRSP data due to fund retirement.

CRSP reports holdings on a quarterly basis (at quarter end) for most funds, but occasionally includes holdings reports for additional or non-quarter-end months. As a result, there are cases in which several holdings reports exist within the same quarter. To avoid double counting holdings, we only use the most recent effective date for each combination of portno, CUSIP, and report date.

Because many TDFs have several portnos, there are a number of cases in which two CRSP holdings reports exist for the same TDF in the same month.[8] To determine which holdings report to use, we determine, for each portno, the latest report date. We elect to use the holdings report for the portno with the latest overall report date, indicating that it is the most up-to-date portno in the CRSP holdings data.[9]

As not all funds regularly report holdings, we construct balanced panels of holdings by forward filling the most recent holdings reports. For Thomson Reuters, all reports are at quarter end and missing holdings are filled (within Thomson Reuters FundNo) with the most recently available quarter's data. For CRSP, given that reporting is not always at quarter end, we forward fill (within TDFId) on a monthly basis with the most recently available month's data. After filling at the monthly level, we only retain quarterly observations to align with Thomson Reuters holdings

[8]The overlapping holdings reports typically take place as TDFs are transitioning from one portno to another portno and the overlap can occur over several years.

[9]One exception occurs for the TDF represented by portnos 1021485 and 1004885. Due to apparent duplication of fundnos, we drop holdings for portno 1004885 and begin using holdings for portno 1021485 in September 2010.

Table DA3: Fund Family Summary Statistics

This table shows summary statistics for each fund family. The first set of columns shows the number of TDFs and the number of underlying share classes for our full sample. The second set of columns shows the number of TDFs and AUM (in millions) for 2019. The final set of columns shows the number of TDFs and AUM (in millions) in 2019 for our filtered sample, which only includes TDFs with well-populated holdings data (see Section DA1.3 for details.)

	Full Sample		**2019 Sample**		**Filtered Sample**	
Family Name	TDFs	Share Classes	TDFs	AUM	TDFs	AUM
Vanguard Group Inc	25	25	22	$489,513	22	$489,513
Fidelity Management & Research Company	123	451	95	$266,288	92	$266,280
T Rowe Price Associates Inc	37	83	37	$164,045	37	$164,045
American Funds	20	183	12	$154,365	12	$154,365
Tiaa-Cref	22	88	22	$59,717	22	$59,717
Jpmorgan Funds	22	166	19	$49,748	19	$49,748
BlackRock Inc	44	186	31	$36,120	10	$143
Principal Global Investors Llc	29	153	29	$23,679	29	$23,678
John Hancock Group	34	240	33	$16,631	32	$16,606
American Century Investment Mgmt Inc	11	62	9	$14,406	9	$14,406
Callan Associates Inc	11	11	11	$9,945	11	$9,945
Voya Investments Llc	29	144	27	$8,188	18	$2,960
State Street Bank And Trust Company	10	30	10	$7,798	10	$7,798
Mml Investment Advisers Llc	23	135	21	$7,752	5	$1,284
Gw Capital Management Llc	35	182	30	$7,127	26	$7,116
Charles Schwab Investment Management Inc	22	33	22	$6,195	22	$6,195
Mutual Of America Cap Mgmt Corporation	11	11	9	$4,901	9	$4,901
Usaa Asset Management Company	5	5	5	$4,511	5	$4,511
Guidestone Funds Trust	6	11	5	$3,927	5	$3,927
Mfs Investment Management	11	101	9	$2,823	9	$2,823
Wells Fargo Fund Management Llc	21	125	21	$2,645	0	$0
Nationwide Fund Advisors	12	69	11	$1,866	7	$1,228
Putnam Investment Management Llc	19	83	9	$1,112	9	$1,112
Dimensional Fund Advisors Lp	13	13	12	$1,007	12	$1,007
Alliancebernstein Lp	23	161	11	$873	10	$872
Pimco	19	97	17	$822	14	$807
Franklin Templeton Investments	10	47	10	$620	9	$595
Axa Enterprise Funds	5	14	5	$546	5	$546
Lincoln Investment Advisors Corporation	10	25	5	$495	5	$495
Transamerica Asset Management Inc	10	30	10	$409	10	$409
Manning & Napier Advisors Inc	11	53	10	$405	2	$86
Mainstay Funds	6	36	6	$336	6	$336
Prudential Investments Llc	12	72	12	$291	11	$291
Goldman Sachs & Co/Gsam	19	88	13	$230	13	$230
Invesco Funds	16	115	15	$207	9	$127
Dws Investment Management Americas Inc	5	18	3	$197	3	$197
Harbor Capital Advisors Inc	11	31	10	$183	9	$183
Allianz Global Investors	9	55	8	$170	6	$147
1290 Funds	9	9	9	$68	9	$68
Natixis Advisors Lp	10	10	10	$53	0	$0
Columbia Threadneedle Investments	9	18	9	$21	6	$18
Bmo Asset Management Corp	10	40	10	$0	10	$0
Columbia Funds	19	96	0	$0	0	$0
Old Mutual Capital Inc	12	24	0	$0	0	$0
Russell Investment Group	10	60	0	$0	0	$0
Virtus Retirement Investment Adv Llc	10	30	0	$0	0	$0
Vantagepoint Investment Advisers Llc	10	20	0	$0	0	$0
Hartford Mutual Funds	9	45	0	$0	0	$0
Van Kampen Asset Management	9	36	0	$0	0	$0
Legg Mason	8	48	0	$0	0	$0
Oppenheimerfunds Inc	7	35	0	$0	0	$0
Allstate Institutional Investors Llc	7	30	0	$0	0	$0
American Independence Financial Svcs Llc	6	16	0	$0	0	$0
Seligman Funds	4	20	0	$0	0	$0
Pnc Funds	4	8	0	$0	0	$0
Dbx Strategic Advisors Llc	4	4	0	$0	0	$0
Payden & Rygel	4	4	0	$0	0	$0
Symetra Investment Management Inc	4	4	0	$0	0	$0
Federated Investors	3	9	0	$0	0	$0
Sunamerica Asset Management Corp	2	6	0	$0	0	$0
Citigroup First Inv Management Am Llc	2	2	0	$0	0	$0
First Investors Management Co Inc	2	2	0	$0	0	$0

reports. For both Thomson Reuters and CRSP holdings data, we record whether the quarterly holdings data is based on a current holdings report or is filled using a past holdings report.[10]

Table DA4 reports the coverage of our filled Thomson Reuters and CRSP holdings data. The table shows that Thomson Reuters holdings represent about one-third of TDFs (but a majority of TDF assets), which is due to limited MFLINKS mappings between Thomson Reuters and CRSP. CRSP holdings cover nearly our full sample of TDFs, in both number and in total assets. In fact, for most of our sample, CRSP reports holdings on every TDF in the Thomson Reuters data. This suggests that Thomson Reuters holdings data is not utilized except in rare cases. However, note that Table DA4 does include filled holdings. As such, there are cases in which Thomson Reuters data is more current than CRSP data.

To create our final holdings data, we combine Thomson Reuters and CRSP holdings at the quarterly level. As holdings often exist in both data sets, we select which holdings to use based on the following priority structure. Given the more extensive coverage by CRSP, we first use CRSP holdings if they are available and if they are current holdings. Second, when current CRSP holdings data are missing, we use Thomson Reuters holdings if they are available and are current holdings (this occurs most often in the early years of our sample). If no current holdings are available, we next use CRSP filled holdings, and lastly, we use Thomson Reuters filled holdings.[11] Table DA5 shows the number of quarterly reports from CRSP and Thomson Reuters by year. 68% of quarterly holdings reports are current and from CRSP, 8% of reports are current and from Thomson Reuters, and the remaining 24% of reports use past CRSP holdings. Returning to Table DA4, the last two columns show the portion of our sample covered by the combination of Thomson Reuters and CRSP holdings data. Nearly 100% of our TDF sample has holdings data, particularly after 2010.

It is worth noting that there are numerous cases in which a parent TDF reports holding another TDF within its portfolio.[12] In some cases the held TDF represents the only holding in the portfolio, while in other cases it is among dozens of other holdings. When TDFs are held by other TDFs, we

[10] In cases that CRSP holdings are reported within the current quarter but not at quarter end, we consider these filled using past holdings.

[11] Our data never utilizes Thomson Reuters filled holdings as there are no instances in which there are missing Thomson Reuters holdings reports when prior CRSP holdings reports do not exist.

[12] While possible, there are no cases in our holdings data in which the held TDF also holds TDFs.

Table DA4: Holdings Report Sample Coverage

This table reports the number of TDFs with holdings reports and the total AUM (in billions) of those TDFs by year. Missing holdings reports are filled using the most recently available holdings report. The first set of columns report full sample statistics (without regard for holdings reports), the second set of columns report statistics for TDFs with Thomson Reuters holdings reports, the third set of columns report statistics for TDFs with CRSP holdings reports and the fourth set of columns report statistics for TDFs with either Thomson Reuters or CRSP holdings reports.

	Full Sample		**Thomson Reuters**		**CRSP**		**Combined**	
Year	Num TDFs	AUM	Num TDFs	AUM	Num TDFs	AUM	Num TDFs	AUM
2006	168	$112	46	$88	105	$106	110	$107
2007	252	$178	127	$161	160	$171	170	$172
2008	385	$154	130	$136	179	$148	186	$148
2009	446	$247	132	$209	246	$234	246	$234
2010	444	$329	132	$257	422	$327	422	$327
2011	466	$363	164	$266	450	$363	450	$363
2012	517	$467	163	$321	491	$467	491	$467
2013	527	$602	179	$405	515	$602	515	$602
2014	579	$683	224	$454	562	$682	562	$682
2015	630	$746	253	$495	595	$746	595	$746
2016	652	$848	278	$552	635	$848	641	$848
2017	680	$1,185	280	$803	669	$1,185	669	$1,185
2018	677	$1,060	255	$756	667	$1,060	667	$1,060
2019	684	$1,350	244	$924	673	$1,350	673	$1,350

attempt to replace the held TDF with its underlying holdings. Replacing the single TDF holding with that TDF's underlying holdings allows for better replicating portfolios (as no single ETF is highly correlated with most TDFs). When the held TDF's underlying holdings are available, we replace the single TDF holding with the held TDF's underlying holdings. We scale the held TDF's holdings' position sizes based on the held TDF's position value in the parent TDF and we adjust portfolio weights by multiplying the held TDF's portfolio weight in the parent TDF by the portfolio weights in the held TDF's portfolio. When the held TDF's underlying holdings are not available, we keep the held TDF in the parent TDF's holdings.[13]

Finally, given that we focus on only mutual fund and ETF holdings, we assess how much of each

[13]In March 2010, Mass Mutual's Select Destination Retirement 2050 fund reports holding about 2% of its assets in itself. We drop this holding and re-scale the other portfolio holdings to 100%.

Table DA5: Holdings Report Sources

This table reports the number of quarterly holdings reports use in our sample, by year, from CRSP and Thomson Reuters. It also reports whether the quarterly reports are current or based on prior reports (filled).

	CRSP		**Thomson Reuters**	
Year	Current	Filled	Current	Filled
2006	190	51	88	0
2007	328	115	191	0
2008	314	141	257	0
2009	395	222	312	0
2010	1,146	254	140	0
2011	1,464	283	140	0
2012	1,639	338	89	0
2013	1,820	360	108	0
2014	1,919	475	112	0
2015	1,874	699	195	0
2016	2,037	786	177	0
2017	2,124	916	209	0
2018	2,275	935	224	0
2019	2,314	1,038	192	0
Total	21,002	7,266	2,622	0

TDF's assets are represented by those mutual funds and ETFs. For the majority of TDFs, their fund holdings' reflect almost all of their assets (when comparing the aggregate fund holdings to the reported TDF assets in CRSP). However, for a fraction of TDFs, they hold many individual stocks and bonds, or have sparsely populated holdings data. We calculate portfolio values in two ways to measure the fraction of holdings represented by mutual funds and ETFs. First, we calculate holdings-based portfolio values by summing the values of all positions in each holdings report. Position values are calculated by multiplying the shares reported by the price reported. When prices are missing, those holdings are not incorporated into the portfolio values.[14] Second, we calculate share-class-based portfolio values by summing the CRSP end-of-quarter total net assets (mtna) across shares classes.

We filter our sample by comparing the total value of each TDF's mutual fund and ETF holdings to the holdings-based and share-class-based portfolio values. We include reported holdings in our sample when we believe the mutual fund and ETF holdings make up at least 90% and no more than 110% of the portfolio values.[15] First, we include the holdings if the mutual fund and ETF holdings make up between 90% and 110% of the share-class-based portfolio values. In these cases, the holdings data align well with the CRSP Mutual Funds - Summary data and we have confidence in the holdings and share-class-based portfolio values. Because there are cases when the CRSP Mutual Funds - Summary data appear erroneous, we also include the holdings data if the mutual fund and ETF holdings make up between 90% and 110% of the holdings-based portfolio values for *both* Thomson Reuters and CRSP holdings.[16]

Table DA6 displays summary statistics for our filtered and unfiltered data. We require TDFs to have monthly return, fee and cash holdings data to be included in our unfiltered and filtered samples. As a result, our unfiltered data has fewer TDFs and less AUM than the values reported

[14] For Thomson Reuters and CRSP, we fill in missing prices with CRSP prices for mutual fund and ETF holdings, but do not fill in missing prices for individual stocks or bonds as we lack the securities' mappings. For CRSP, in some cases we can fill in missing prices by dividing the reported position value (market_val) by the shares held (nbr_shares). Also note that we round shares held to the nearest integer. Discrepancies can arise as data downloaded from WRDS to a CSV file are all integers, while data downloaded from WRDS to STATA data files can have additional precision.

[15] Because of duplicated holdings and differences in report timing, there are cases in which mutual fund and ETF holdings exceed 100%.

[16] Our secondary criteria only applies to months in which a TDF has holdings reports from both Thomson Reuters and CRSP.

in the last columns of Table DA4. Filtering our sample based on the percentage of holdings made up by mutual funds and ETFs has a bigger impact on the number of TDFs than the AUM. The number of TDFs reduces by between 8% and 29%, with the median year losing 14% of TDFs. The total AUM reduces by between 4% and 21%, with the median year losing 4% of AUM. Overall, restricting ourselves to holdings that represent between 90% and 110% of total portfolio value does not significantly reduce our sample size.

Interestingly, the losses in TDFs and AUM predominantly come from a few fund families. The last two columns of Table DA3 show the number of TDFs and total AUM of our filtered sample for each fund family. Several fund families stand out. BlackRock has 31 TDFs with $36 billion in total assets in 2019, but only 10 TDFs with $143 million in asset are included in our filtered sample. While not as severe, Voya Investments, MML Investment Advisors and Wells Fargo Fund Management also have significant portions of their assets excluded. These fund families' portfolios have significant holdings in securities that are not mutual funds or ETFs, or their data is sparsely populated. In the case of Wells Fargo, their TDFs hold many individual securities and thus the mutual fund and ETF holdings are relatively small. For BlackRock, many of their holdings do not report CUSIPs, so we cannot link those holdings to existing mutual funds or ETFs.

Overall, our filtered sample provides significant coverage of the TDF universe. Comparing Table DA6 to Table DA1 shows that our annual filtered sample assets average 91% of annual reported TDF AUM by ICI. Moreover, our coverage is nearly 100% for the major TDF providers (with the exception of BlackRock).

Table DA6: Filtered Sample

This table reports the annual numbers of TDFs and AUM (in billions) in our unfiltered and filtered samples. The unfiltered sample includes observations using all holdings data, regardless of the percentage of the portfolio holdings represented by mutual fund or ETF holdings. The filtered sample includes observations whose mutual fund and ETFs holdings make up between 90% and 110% of total portfolio holdings.

	Unfiltered Sample		**Filtered Sample**	
Year	TDFs	AUM	TDFs	AUM
2006	90	\$106	66	\$93
2007	166	\$171	143	\$166
2008	176	\$147	148	\$145
2009	234	\$232	167	\$184
2010	411	\$327	351	\$312
2011	432	\$363	369	\$343
2012	470	\$467	408	\$445
2013	508	\$602	446	\$577
2014	546	\$682	484	\$653
2015	581	\$745	521	\$716
2016	620	\$848	572	\$821
2017	660	\$1,185	588	\$1,139
2018	666	\$1,060	572	\$1,020
2019	669	\$1,350	569	\$1,299