Menu Simplification for Portfolio Selection Under Short-Sales Constraints

Farid AitSahlia[1] Thomas Doellman[2] Sabuhi Sardarli[3]

March 30, 2022

Abstract

Based on the empirical evidence of sparse and unstable optimized portfolios, and the fact that average investors prefer fewer choices, we introduce a procedure to identify a subset of low-volatility funds out of a larger set in such way that it does not adversely affect the risk-return spectrum available through the original menu. In fact, we show that the subset can span better than the full plan menu. In effect, our study introduces a new dimension to the debate regarding the so-called low-volatility anomaly, where low-volatility and low-beta have significantly outperformed high-volatility and high-beta portfolios over the last few decades.

JEL Classification Codes: G11, G20, G23

Keywords: asset allocation, mean-variance spanning, Wald test, mutual funds, short sales.

---

[1] Warrington College of Business, University of Florida, Gainesville, FL, 32611; email: farid1@ufl.edu ; Tel: +1 (352) 392-5058.

[2] John Cook School of Business, Saint Louis University, St. Louis, MO 63108; email: thomas.doellman@slu.edu, Tel: +1 (314) 977-3815.

[3] College of Business Administration, Kansas State University, Manhattan, KS 66506; email: ssardarli@ksu.edu ; Tel: +1 (785) 532-5771.

1. Introduction

The mean-variance tradeoff paradigm has been dominant for portfolio selection ever since its introduction in Markowitz (1952). Yet its practical implementation has encountered challenges that have hindered its effectiveness. They are primarily due to the difficulty in obtaining reliable estimates of fundamental parameters; namely, the expected values and the covariance matrix of the asset returns (cf. Brandt (2010) and Kan et al. (2022), among many others.) These econometric issues lead to allocation solutions that are unstable as the associated quadratic optimization is an ill-posed problem, where the optimal portfolio weights are mostly sparse and can switch from one extreme to another under minute perturbations in the input estimates (see, e.g., Brodie et al. (2009).) Over the years, a number of approaches have been proposed to either advance estimation techniques or devise robust optimization formulations (e.g., through regularization), or both (cf. Brodie et a. (2009), Dai and Wen (2018), Dai and Wang (2019), Kan et al. (2022).)

In this paper we take a different tack, where we want to reduce the number of investment options in some optimal mean-variance spanning sense without adversely affecting the risk-reward spectrum available through the original set. We should emphasize at the outset that our approach does not seek to identify the actual allocation strategy across the identified subset, for four reasons: 1) within the broad context of optimal portfolio selection, we want to focus on the preliminary step pursuant to optimally reducing the number of investment options prior to applying quadratic optimization so as to mitigate the propensity of the latter to generate sparse solutions; 2) alternatively, given the fierce debate that has taken hold regarding the lack of out-of-sample superiority of optimized strategies over the naïve even allocation[1] (cf. De Miguel et al. (2009), Kritzman et al. (2010), Scherer (2011), Pflug et al. 2012, De Carvallo et al. (2012), Kirby and Ostdiek (2012), and Zakamulin (2017)), one may apply the latter directly on our subset; 3) empirical evidence showing a preference for concentrated positions, as in the study of the portfolio holdings of the entire Swedish population in Calvet et al. (2008), for example,

---

[1] Allocate $1/n$ of total wealth in each of the n assets under consideration.

which has been rationalized with some theoretical support in Roche et al. (2013); and 4) within the narrower context of defined contributions plans to which employees make contributions to fund their own retirements, the need for simplifying options for investors with limited financial literacy has become critical as we elaborate in the next section. Therefore, optimally identifying a subset of assets out of a larger menu for portfolio selection is indeed justified on multiple grounds.

To support our simplifying proposal, we make use of mean-variance spanning tests under short-sales constraints, which are particularly prevalent with the average investor. A set of assets is said to be mean-variance spanning if their efficient frontier cannot be enhanced with one or more assets from a benchmark set of index funds (cf. Huberman and Kandel (1987).) It is then likely that the set will satisfy a variety of risk/return preferences. Given that the empirical validation of our selection approach rests on data for U.S. defined contributions plans, where positions cannot be held short, we resort to Wald tests, originally developed in Kodde and Palm (1986), as applied in De Roon et al. (2001).

Through an empirical illustration involving defined contribution (DC) plans for retirement in the United States, we show that our approach drastically reduces the number of funds from which investors may choose (from 18 to 6, on average), a significant benefit considering that individuals typically prefer to choose among fewer alternatives.[2] We also show that not only does the smaller subset identified by our method avoids adversely reducing the spectrum of risk preferences relative to the original set, it even spans better in many cases. We reiterate that our proposal does not mandate any specific asset allocation between funds, thus allowing, if they so desire, for naive investment strategies that unsophisticated investors are known to be comfortable with, particularly an even allocation (i.e. 1/n) across available investment options.[3] This simple strategy is easily explained and is intuitively attractive to the average person (Benartzi and Thaler (2001); Iyengar et al. (2004); Iyengar and Kamenica (2010).) While naive,

---

[2] See Tversky and Shafir (1992), Iyengar and Lepper (2000), and Boatwright and Nunes (2001).

[3] Huberman and Jiang (2006) find evidence that participants in U.S. DC plans indeed tend to allocate evenly across investment funds chosen.

this diversification strategy is also perfectly acceptable for many as recent work has shown that it is hard to persistently outperform (see, e.g., De Miguel et al., (2009), Pflug et al. (2012), and Zakamulin (2017).) Alternatively, plan participants may also consider managed accounts, which have grown in popularity, for help with allocation across the reduced set of funds.

The funds we identify through our procedure either have low beta or they comprise the minimum-variance portfolio of the menu of investment options available in the plan, in addition to relatively riskless funds. At first glance, the fact that the subset of funds identified by our selection process spans at least as frequently as the complete set of funds in the plan seems counterintuitive if *all* parameters are known *with certainty*. But it is not surprising when accounting for parameter uncertainty. In fact, our approach is germane to the three-fund strategy of Kan and Zhou (2007) who advocate for the use of a riskless asset, the sample tangency portfolio and the minimum-variance risk portfolio. Their argument is based on the fact that theoretical mean-variance analysis, which relies on precisely known parameters, is significantly altered by the inaccuracy of estimated expected returns and covariance, with ex-post performance falling short of ex-ante expectation. Furthermore, we show that because the mean-variance test we apply relies on a Wald statistic, the better spanning results of the subset can be explained by a higher concentration of covariance entries among the smaller set, which ultimately leads to a smaller Frobenius norm of the matrix associated with that statistic. In essence, the empirical mean-variance spanning performance of our low-volatility strategy is also consistent with the superior out-of-sample return performance of others relative to high-volatility and high-beta strategies (see e.g., Baker et al., 2011; Frazzini and Pedersen, 2014). We therefore contribute to both the literature on mean-variance optimization and the literature on employee investment decision--making in DC plans.

The remainder of the paper is organized as follows. In Section 2 we relate the issue of optimal menu reduction to identifying the so-called Qualified Default Investment Alternatives designed to lessen the burden of retirement planning in U.S. DC plans, highlighting deficiencies regarding the popular current default QDIA option (i.e., target-date funds). Section 3 further details our proposal for plan menu simplification. Section 4 describes our data sources and contains sample details. Section 5 discusses our empirical results, including an illustration of how a subset based on our method can span better than the larger menu, and Section 6 concludes.

2. U.S. Defined Contribution Plans: The Fiduciary Default Allocation Issue

Defined contribution (DC) plans, such as those provided by employers under section 401(k) of the U.S. Internal Revenue Code, figure prominently in the retirement planning of tens of millions of working Americans despite their modest beginning in the early 1980s (Munnell and Sundén (2004).) In Europe, they are the subject of vigorous discussions regarding their implementation (cf. Hinrichs (2020), and Holzmann et al. (2021).) Though current views on defined contributions (i.e., nonfinancial DC) plans in Europe make them different than their U.S. versions, the experience on the latter over the past 40 years have much to offer the rest of the world.

DC plans place the burden of investment decisions on the shoulders of relatively unsophisticated investors and there has been a growing concern over a retirement crisis in the United States (Siegel, 2015) due partly to weak proactive participation and unwise strategies such as all equities or all bonds (Roche et al. (2013).) Many plan participants seem to exhibit limited effort, or even no effort, when making investment decisions in their retirement plan. For instance, Doellman et al. (2019) provide evidence that the average plan participant resorts to simply choosing funds from the top of the plan menu list. Many other employees follow the path of least resistance and rely on default decisions made by their employers (Choi et al., 2002; Carroll et al., 2009).

To help such investors in their decision-making, the push for widespread financial education, as advocated by certain academics and policy makers, has not improved the situation (see, e.g., Fernandes et al., 2014). Based on research suggesting that a bewildering plan menu may hamper participation (cf. Iyengar et. al. (2004) and Iyengar and Kamenica (2010)) and the previously cited literature regarding the preference for fewer choices in decision making, a potentially more effective remedy would be to simplify the investment decision process for employees while improving the mix of investment assets in their retirement portfolios.

Despite plan participants' limited ability to make informed investment decisions, DC plan fiduciaries, such as sponsoring employers and plan providers, had long been reluctant to make fund recommendations to employees due to possible liability exposure. This changed to a degree in 2006 when the U.S. Congress passed the Pension Protection Act. This Act included relief from liability for plan fiduciaries making fund choices on the behalf of plan participants if they comply with the Department of Labor's safe harbor rule. The latter was issued in October 2007 and provided guidance concerning Qualified Default Investment Alternative (QDIA) options in DC plans.

The question of how to implement a practical QDIA in DC plans remains open despite the emergence of target-date funds (TDFs) as a popular option. With a TDF, an investor picks a retirement age (target date) and the fund manager makes allocation decisions that are pre-determined and which change as the investor ages. The change in investment allocation over time is referred to as the "glide path," and it typically starts with a tilt towards stocks and ends with more bonds by the target date. The passage of the Pension Protection Act in 2006, along with the QDIA designation of target-date funds and the move to automatic enrollment in DC plans, has led a growing proportion of employees to allocate their entire plan balances to TDFs. In light of the limitations of the average DC investor, this trend has led to some favorable outcomes. For instance, the average DC investor is now less likely to employ extreme portfolio allocations (i.e., all-equity, all-bond or all-money-market) (cf. Huberman et al. (2006) and Roche et al. (2013).)

Certain characteristics of TDFs, however, have also made their increased use problematic. For example, Balduzzi and Reuter (2019) show that TDFs with similar target dates earn significantly different realized returns and follow glide paths with very heterogeneous ex-ante risks. Investors also largely share many misconceptions regarding these funds, such as whether TDFs offer guaranteed income. Surz and Israelsen (2007) discuss how TDFs should be assessed and find that their risk-adjusted performance falls short while Spitzer and Singh (2008) use simulation to show that TDFs have a higher shortfall risk than a constant equal-allocation between stocks and bonds. Scott et al. (2009) also show that glide-path strategies have higher shortfall risk compared to constant mix strategies. They observe that TDFs tend to lock in poor early returns, thereby decreasing the likelihood of portfolio recovery should returns improve. From a different perspective, Sandhya (2011) finds that TDFs are also subject to agency problems as some mutual fund companies use low quality funds to create TDFs. Additionally, as TDFs are funds of funds, their fees can be significant. Elton et al. (2015) provide evidence that these fees are somewhat offset by the lower fees of the funds into which the TDFs are invested. On the whole, however, they show that the resulting alphas are generally lower than alternatives for any given fund family.[4]

3. Plan Menu Simplification

The fund selection approach we prescribe is mainly driven by portfolio efficiency, specifically, mean-variance spanning. Our approach is also motivated by the strong empirical performances of low-volatility and low-beta portfolios (Karceski, 2002; Ang et al., 2009; Baker et. al., 2011; Frazzini and Pedersen, 2014) which was, in fact, anticipated as far back as Black (1972). This is especially true within the context of restricted borrowing that is characteristic of mutual funds.

To determine the QDIA subset of a given plan, we first partition the plan's options into two groups: equity-based (higher risk) funds and safer funds, which typically include stable value, money market, fixed income, target-date, and general conservative funds.[5] We then perform a variance minimization on the equity funds and keep only the funds that are part of the optimal solution. Next, we combine these with the safer funds, which were excluded from the variance minimization altogether. This approach has the same flavor as the classical two-fund theorem; however, it differs from it since the two funds (or sets of funds,

---

[4] Additional issues associated with TDFs are further presented in the review article of Spitzer and Singh (2012).

[5] Partitioning of funds is based on Morningstar categories. We first determine all unique Morningstar categories in our data and then identify categories that represent equity based funds.

with a cardinality typically not exceeding three) are not necessarily on the efficient frontier. It is, in fact, comparable to the three—fund approach of Kan and Zhou (2007).

We then perform a Wald spanning test (cf. De Roon et al. (2001)) on the combined set of funds to determine whether they can span the returns of index benchmarks. We show that the subset of funds tends to span at least as frequently as the full set of funds in a given plan. Thus, the manner with which we choose the subset of funds complies with the safe harbor provision that a QDIA must be "*diversified so as to minimize risk of large losses*" and "*designed to provide varying degrees of long-term appreciation through a mix of equity and fixed-income exposures*".[6]

4. Data

Our primary dataset is provided by Brightscope, Inc., an independent information provider of retirement plan ratings and investment analytics to plan participants, sponsors, asset managers, and advisors. Brightscope's proprietary dataset currently contains information on over 55,000 defined contribution (DC) plans, such as 401(k) and 403(b) plans. The specific dataset provided to us is a cross-sectional snapshot of plans at the end of 2007, and it contains over 25,000 DC plans. Items contained in the data that are important to our analysis include: plan menu investment fund options, plan size (net assets), individual fund balances, fund expense ratios, administrative costs, and plan sponsor and service provider information.

In this study, we focus on companies' primary DC plans as identified by using the Department of Labor codes, thus we eliminate any supplementary plans offered by the same plan sponsor.[7] This initial sample includes 17,386 DC plans. We further require full return data availability for every mutual fund within a plan in the 2004-2008 period from either CRSP Mutual Fund Database or Morningstar Direct

---

[6] Final rule in *Federal Register,* published on Oct 24, 2007 and available at: https://www.dol.gov/ebsa/regs/fedreg/final/07-5147.pdf

[7] We do this because we do not typically have complete data for all of a company's plans; thus, we simply analyze a company's primary, or largest, DC plan.

Database. Because of this return data constraint our final sample used in the analysis consists of 7,975 DC plans. Despite the loss of a large number of plans in the original sample, our final sample is significantly larger and richer than previous studies that have analyzed retirement plan menu efficiency. For instance, Elton et al. (2006) analyze a relatively small sample of 417 plans while Tang, et al. (2010) analyze a larger sample of 1,003, all of which are administered by one of the top mutual fund companies in the industry – Vanguard.[8] In addition, our data covers both publicly traded and private companies of all sizes which hire many types of TPAs.[9]

<< Insert Table 1 around here >>

Table 1, Panel A provides comparative descriptive statistics for the initial and the final sample. The average plan in our initial sample has $22.1 million in total net assets and offers 22.1 funds as investment options.[10] In contrast, the average plan in our final sample is larger than the average plan in the initial sample with $31.8 million in total assets and contains 18.2 fund options. For the majority of our sample we can also identify the third-party plan administrator (TPA), the financial institution in charge of designing and servicing the retirement plan. TPA categories used in our analysis can be found in Panel B of Table 1, along with plan size characteristics across these different categories. Mutual fund families represent a heavy majority of the subsample where we can identify the TPA. This is consistent with the overall retirement plan market where mutual fund companies hold a majority of the market share. In our sample, plans administered by investment banks, large commercial banks, asset management advisory firms, and mutual fund companies are considerably larger than plans administered by small/regional commercial banks, 401(k) services companies, and insurance firms. This is not particularly surprising since the clientele of the

---

[8] Elton et al. (2006) uses a sample provided by Moody's Investors Services that collects survey data from for-profit firms. The sample used in Tang et al. (2010) is supplied by Vanguard, a company that is well known to provide low cost and well diversified portfolios with heavy preference to index funds.

[9] We categorize TPAs in our sample into one of seven categories: mutual fund families, large/small (greater than or less than $50 billion in assets) commercial banks, insurance companies, asset management advisory companies, investment banks, and 401(k) services companies

[10] The average plan size in our initial sample is very comparable to the average plan size ($25.2 million) in EBRI's 2007 dataset, the largest provider of information on 401(k) plan.

latter groups are most likely to be smaller firms with fewer participants and lower retirement plan balances.[11]

<< Insert Table 2 around here >>

In Table 2, we provide more details on the types of funds offered in investment menus of plans in the sample. As summarized in Panel A, an average retirement plan in our sample offers 12.3 domestic equity funds, 1.8 domestic bond funds, 1.9 international funds and 0.6 low risk investments such as money market funds, stable value funds, guaranteed investment contracts, or annuities (MSGA).[12] Not surprisingly, almost all retirement plans include at least one domestic equity fund while 97% of plans offer at least one domestic bond fund and one international fund in their investment menus. Additionally, about 60% of plan menus contain at least one MSGA, while 4% of plans in the sample also include company stock as one of investment choices.[13] Plan participants in our sample, on average, direct 68% of their retirement wealth to domestic equity funds, 9% to domestic bond funds, 14% to international funds and 9% to MSGA options. When company stock is offered in the DC plan, the average plan participant also invests 13% of plan assets in the stock.[14] Further, Panel B reports the average number of unique Lipper Investment Objective categories covered by plans in the final sample. On average, 13 different objective categories are represented in plan menus.

<< Insert Table 3 around here >>

For purposes of spanning, we use two sets of benchmark funds. We obtain all return data for these benchmarks from DataStream. Table 3 lists both sets of benchmark indexes and provides descriptive statistics on monthly returns over the analysis period (2004-2008). Despite some changes in management

---

[11] One exception is the insurance company group. One drawback of our dataset is the fact that insurance firms are underrepresented due to the common use of proprietary funds that do not exist in our return data sources (CRSP or Morningstar).

[12] In our analysis we only focus on mutual funds in the plan and exclude company stock and all MSGA options.

[13] Interestingly, retirement participants of about 92% of plans in our sample have also borrowed against their retirement wealth, on average 8.3% of their plan balance.

[14] Although this figure seems quite high, often there are special incentives in investing in company stock for employees.

company affiliations, the first set of benchmarks are identical to sets used by the related spanning literature, including Elton et al. (2006) and Tang et al. (2010). In the first set, the Barclays Capital Aggregate Bond Index, Credit Suisse High Yield Bond Fund, and Citigroup World Government Bond Non-US$ Index capture returns of fixed income securities; the Russell 1000 Growth, Russell 1000 Value, Russell 2000 Growth, and Russell 2000 Value indices capture returns of large-, mid- and small-cap equities; and the MSCI EAFE Index provides international exposure. In the second set, each of the investment categories are represented with more benchmark indices. This set is widely used in the style analysis literature (cf. Sharpe (1992) and is further discussed in the next section. In this set, fixed income benchmarks are Barclays Government Intermediate, Barclays US Aggregate Long Government/Credit, Barclays Investment Grade: Corporates, Barclays US Agency Fixed Rate MBS, Citigroup World Government Bond Index World 5+ Year Non-USD; equity benchmarks are Standard and Poor's 500/ Citigroup - Value; Standard and Poor's Midcap 400/ Citigroup - Value; Standard and Poor's Midcap 400/ Citigroup - Growth; Standard and Poor's Smallcap 600/ Citigroup- Value; Standard and Poor's Smallcap 600/ Citigroup - Growth, and finally, the international benchmarks are MSCI Europe, MSCI Pacific, and S&P IFCI Emerging Market Index. This set then differs from that used by Elton et al. (2006) and Tang et al. (2010) by disentangling the small and mid-cap groups and by differentiating between the international regions. As summarized in Table 3, the average monthly returns on most benchmark indices were negative in the 2004-2008 period due to the financial crisis of 2008.[15]

5. Mean—Variance Testing Implementation and Results

Given that our proposal reduces the number of funds to consider in a given plan, it is natural to ask whether the resulting subset would span less than the original in the mean-variance sense, potentially

---

[15] Average monthly returns are positive for all indices in both sets if year 2008 returns are removed from the descriptive statistics. Note that a positive arithmetic average monthly return does not imply that the corresponding average compounded return is positive.

reducing the spectrum of risk-return trade-offs that would otherwise be available to participants. We resort to employing Wald tests, developed originally by Kodde and Palm (1986), as advocated in DeRoon et al. (2001) in the context of regression-based mean-variance spanning when short-sales constraints such as those in DC plans are present. Specifically, denote by r the N-dimensional vector of fund returns (e.g., the full plan or its subset) the mean-variance spanning of which are to be assessed against a K-dimensional vector R of benchmark returns. Through the multivariate linear regression specification (using notation from DeRoon et al. (2001))

$$r = a + BR + \varepsilon, \tag{1}$$

where a and B are of dimensions N and $N \times K$, respectively, we test the hypothesis (cf. (15) on p. 727 of DeRoon et al.(2001))

$$va + (Bi_K - i_N) \leq \mathbf{0}, \tag{2}$$

where v is a mean discount factor, set to $\frac{1}{1+r_f}$ given the presence of a risk-free rate r_f, i_K and i_N are unit vectors of size K and N, respectively, and $\mathbf{0}$ is a zero-vector of size N[16]. The rejection of (2) is then interpreted as the N funds not spanning. In keeping with the current literature labelling (e.g., Elton et al. (2006) and Tang et al. (2010)) that reflects the dichotomy spanning vs. non-spanning, we will similarly refer to failure to reject as spanning, despite it not necessarily being the case strictly speaking[17].

As a reminder, we determine the QDIA--compliant subset of funds by first partitioning a plan's options into equity-based funds and all other funds. Next, we perform a variance minimization on the equity funds and keep only the funds that are part of the optimal solution. Finally, we combine the funds from this optimal solution with the other, safer funds that were excluded from the variance minimization to constitute our proposed QDIA—compliant subset of funds. Furthermore, in order to address the arbitrariness of the benchmark choice and potential collinearity issues with the benchmark indexes, we perform spanning tests

---

[16] DeRoon et al. (2001) use an additional superscript for the regression parameters a and B as they derive their expression (15) based on running the regression (1) above using only a subset of R that is associated with non-binding short-sales constraints. However, in their Appendix they argue that for implementation purposes, the correct identification of such a sub-vector is not practical and that even if done incorrectly will be asymptotically negligible.

[17] The proper label should be "fail to reject spanning" at the given (5%) significance level.

where their returns are re-expressed in terms of those of all their principal components (see, e.g., Lai and Xing, 2008, pp. 41--44, or Connor and Korajczyk, 2010, pp. 401—418.)

<< Insert Table 4 around here >>

<< Insert Table 5 around here >>

Table 4, Panel A, shows that relative to the same benchmarks used by Elton et al. (2006) and Tang et al. (2010), we find that 46% of plans span when considering all fund options in the plan ("All Funds"). On the other hand, with QDIA-compliant subsets of funds ("QDIA Funds"), spanning occurs around 49% of the time. Should our alternative benchmark of 13 indices be used instead, spanning occurs at much lower rates: 28% for all funds and 31% for QDIA- compliant subsets of funds, as reported in Panel B of Table 4.

The first important finding from these results is that the spanning rates markedly differ according to the two benchmark sets. However, as shown in Table 5, the proportion of spanning increases dramatically for both sets of funds (All Funds and QDIA--compliant subsets of Funds) and both sets of benchmarks when the principal components of the funds are used as regressors. These results clearly highlight the significance of the correlation among fund returns. We also note that our selection approach tends to reduce the number of choices by almost two-thirds, with a reduction from a mean (median) of 18.2 (17) funds to a mean (median) of 6.2 (5) in a typical plan.

The second, more important takeaway from Tables 4 and 5 is that limiting the plan menu to QDIA-compliant subsets of funds does not impair the spanning opportunities offered by the complete set of menu options in the plan. However, given the above--mentioned issues with offering the full menu to participants, the suggestion of QDIA-compliant subsets of funds could greatly alleviate the burden of the investment decision for the participants. If the TPA and/or the plan sponsor would choose to offer, through managed accounts, for example, guidance on the set of funds in the plan menu that are part of the minimum-variance optimized portfolio, together with the conservative funds, then a plan participant is more likely to create an investment portfolio that spans a set of benchmark indices. In this fashion, fiduciaries have an objective basis for their fund recommendations, which may otherwise be viewed under a cloud of suspicion.[18]

Our approach is in contrast to Tang et al. (2010) who further assess the impact of the deviation of the risk-adjusted performance across all participants relative to the mean-variance efficient portfolios and conclude that while plans may be spanning, individuals' portfolio constructions are overwhelmingly inefficient. Their recommendation is then to support strategies targeting behavioral change including improved default strategies and educational programs. However, behavioral change is difficult to achieve (see, e.g., Iyengar and Kamenica, 2010) and financial education has a very short "shelf life", thus severely limiting its efficacy (see, e.g., Fernandes et al., 2014). In addition, the estimation issues that arise in the course of mean-variance optimization have led some to question the practicality of mean-variance efficient portfolios, leading some to further wonder whether the "1/N" strategy could be a viable strategy (see De Miguel et al., 2009 and Zakamulin (2017), among others.)

We close this section by providing an illustration explaining how a subset selected through our procedure can span better than the original menu of options. The rejection of the null hypothesis captured by (2) is predicated on large values of the Wald statistic

$$\xi = \min_{\gamma \geq 0} (\tilde{\gamma} - \gamma)' \tilde{\Sigma}^{-1} (\tilde{\gamma} - \gamma), \tag{3}$$

where

$$\tilde{\gamma} = -\frac{1}{1+r_f}\widehat{\alpha} - \hat{\beta} \times i_K + i_N,$$

with $\widehat{\alpha}$ and $\hat{\beta}$ being the respective estimates for the regression (1), and

$$\tilde{\Sigma} = \left(\frac{1}{1+r_f} I_N \ - A\right) \Omega \left(\frac{1}{1+r_f} I_N \ - A\right)' \tag{4}$$

---

[18] Highlighted in a recent Government Accountability Office (GAO) report –"401(k) Plans: Improved Regulation Could Better Protect Participants from Conflicts of Interest." available at http://www.gao.gov/new.items/d11119.pdf.

where I_N is the $N \times N$ identity matrix, A is the Kronecker product $I_N \otimes i_K{}'$, and Ω is the $(N + NK) \times (N + NK)$ covariance matrix between the multivariate intercept α and the loading matrix β in the regression (1).

<< Insert Table 6 around here >>

Table 6 shows the matrices that result from the estimates $\widehat{\alpha}$ and $\hat{\beta}$. Observe that the covariance between these terms is stronger for the QDIA subset, leading to a more robust matrix $\tilde{\Sigma}$, based on both determinant and Frobenius norm values, and ultimately yielding a distance-defining $\tilde{\Sigma}^{-1}$ that clearly tilts the Wald statistic toward smaller values for the QDIA subset as captured by its Frobenius norm. One can explain this mechanism by the way that the subset is chosen. It consists of assets with low risk or which minimize the risk, thus close to the efficient frontier of the original set. Given that the benchmark indexes are efficient, it is not surprising to see that the estimates of the regression of the subset on the benchmark indexes result in a covariance matrix with larger terms than those of the original set as assessed via the Frobenius norm of Ω and the min, max, and median values.

6. Conclusion

Through either direct retail channels or indirect institutional channels such as retirement plans, average or unsophisticated investors constitute a very significant segment of the financial universe. Yet much remains to be done to help them with their asset selections. In view of the empirical evidence pointing out that individuals prefer fewer choices when making complex decisions such as investments, we propose a systematic procedure that not only reduces the number of options without limiting the risk-return opportunities relative to the original offering but can also capture better risk-return tradeoffs thanks to its reliance on risk minimization.

References

Ang, A., Hodrick, R. J., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics* 91, 1-23.

Baker, M., Bradley, B., Wurgler, J., 2011. Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal* 67 (1), 40-54.

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.

Benartzi, S., Thaler, R. H., 2001. Naive diversification strategies in defined contribution savings plans. *American Economic Review* 91(1), 79–98.

Black, F., 1972. Capital market equilibrium with restricted borrowing. *Journal of Business* 45 (3), 444-455.

Boatwright, P., Nunes, J. C., 2001. Reducing assortment: an attribute-based approach. *Journal of Marketing* 65(3), 50-63.

Brandt, M. W. 2010. Portfolio choice problems. In Ait-Sahalia, Y. and Hansen, L. P., eds. *Handbook of Financial Econometrics*. Elsevier. 269—336.

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., and Loris, I. 2009. Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences* 106 (30), 12267—122272.

Calvet, L.E., Campbell, J.Y., Sodini, P., 2008. Down or out: assessing the welfare costs of household investment mistakes. *Journal of Political Economy* 115, 707—747.

Carroll, G.D., Choi, J.J., Laibson, D., Madrian, B.C., Metrick, A., 2009. Optimal defaults and active decisions. *The Quarterly Journal of Economics* 124 (4), 1639—1674.

Choi, J., Laibson, D., Madrian, B., Metrick, A., 2002. Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. *Tax Policy and the Economy* 16, 67- 113.

Connor, G., Korajczyk, R., 2010. Factor models in portfolio and asset pricing theory, in Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques (J.B. Guerard, ed.) . Springer, New York, NY.

Dai, Z., Wang, F., 2019. Sparse and robust mean-variance portfolio optimization. *Physica A* 523, 1371—1378.

Dai, Z., Wen, F. 2018. Some improved sparse and stable portfolio optimization problems. *Finance Research Letters* 27, 46—52.

De Carvallo, R. L., Lu, X., Moulin, P., 2012. Demystifying equity risk-based strategies: a simple alpha plus beta description. *Journal of Portfolio Management* 38 (3), 56—70.

De Miguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies* 22(5), 1915-1953.

De Roon, F., Nijman, T., Werker, B., 2001, Testing for the mean-variance spanning with short sales constraints and transaction costs: The case of emerging markets. *Journal of Finance* 56(2), 721-742.

Doellman, T., Itzkowitz, J., Itzkowitz, J., Sardarli, S., 2019. Alphabeticity bias in 401(k) investing. *Financial Review*, 54(4), 643-677.

Doellman, T., Sardarli, S., 2016. Investment fees, net returns, and conflict of interest in 401 (k) plans. *Journal of Financial Research* 39(1), 5-33

Elton, E., Gruber, M., Blake, C., 2006. The adequacy of investment choices offered by 401(k) plans. *Journal of Public Economics* 90, 1299–1314.

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

Fernandes, D., Lynch, J.G., Netemeyer, R. G., 2014. Financial literacy, financial education, and downstream financial behaviors. *Management Science* 60(8):1861-1883

Frazzini, A. and Pedersen, L. H., 2014. Betting against beta. *Journal of Financial Economics* 111,1-25.

Hinrichs, K. 2020. Recent pension reforms in Europe: More challenges, new directions. An overview. *Social Policy Administration* 55, 409—422.

Holzmann, R., Palmer, E., Palacios, R., and Sacchi, S. 2020. *Progress and Challenges of Nonfinancial Defined Contribution Pension Schemes*. Vol. 1. The World Bank, Washington, D.C.

Huberman, G., Jiang, W., 2006. Offering versus choice in 401(k) Plans: Equity exposure and the number of funds. *Journal of Finance* 61, 763– 801.

Huberman, G., Kandel, S., 1987. Mean-variance spanning. *Journal of Finance* 42, 873– 888.

Iyengar, S., Huberman, G., Jiang, W., 2004. How much choice is too much: determinants of individual contributions in 401K retirement plans? In: Mitchell, O.S., Utkus, S. (Eds.), Pension Design and Structure: New Lessons from Behavioral Finance. Oxford University Press, Oxford, pg. 83 – 95.

Iyengar, S., Kamenica, E., 2010. Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics* 94(1), 530–539.

Iyengar, S., Lepper, M., 2000. When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology* 79(6), 995-1006.

Kan, R. and G. Zhou, 2007. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis* 37, 559–594.

Kan, R., Wang, X., and G. Zhou, 2022. Optimal portfolio choice with estimation risk: No risk-free asset case. *Management Science* 68 (3), 2047–2068.

Karceski, J., 2002. Returns-chasing behavior, mutual funds, and beta's death. *Journal of Financial and Quantitative Analysis* 37 (4), 559–594.

Kirby, C. Ostdiek, B., 2012. It's all in the timing: simple active portfolios that outperform naïve diversification. *Journal of Financial and Quantitative Analysis* 47, 437—467.

Kodde, D., Palm, F., 1986. Wald criteria for jointly testing equality and inequality restrictions. *Econometrica* 56, 1243–1248.

Kritzman, M., Page, S., Turkington, D., 2010. In defense of optimization: the fallacy of 1/N. *Financial Analysts Journal* 66 (2), 31—39.

Lai, T. L., Xing, H., 2008. Statistical models and methods for financial markets. Springer, New York, NY.

Markowitz, H., 1952. Portfolio selection. *Journal of Finance* 7, 77--91.

Munnell, A. and Sundén, A. 2004. *Coming Up Short: The Challenge of 401(k) Plans*. The Brookings Institution, Washington, D.C.

Pflug, G., Alois Pichler, C., Wozabal, D., 2012. The 1/N investment strategy is optimal under high model ambiguity. *Journal of Banking and Finance* 36, 410–417.

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

Roche, H., Tompaidis, S., Yang, C., 2013. Why does junior put all his eggs in one basket? A potential rational explanation for holding concentrated portfolios. *Journal of Financial Economics* 109, 775-796.

Sandhya, V., 2011. Agency problems in target-date funds. Discussion Paper 03/2011-035. Network for Studies on Pensions, Aging and Retirement.

Scherer, B. 2011. A note on the returns from the minimum variance investing. *Journal of Empirical Finance* 18 (4), 652—660.

Scott, J., Sharpe, W., Watson, J., 2009. The 4% rule – At what price? *Journal of Investment Management* 7(3), 31–48.

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

Siegel, 2015. After 70 years of fruitful research, why is there still a retirement crisis? *Financial Analyst Journal* 71, 6-15.

Spitzer, J., Singh, S., 2008. Shortfall risk of target-date funds during retirement. *Financial Services Review* 17, 143-153.

Spitzer, J., Singh, S., 2012. Target-date mutual funds, in consumer knowledge and financial decisions: Lifespan perspectives, D.J. Lamdin (ed.). Springer (DOI 10.1007/978-1-4614-0475-0_17)

Surz, R., Israelsen, C., 2007. Evaluating target date lifecycle funds. *Journal of Performance Measurement* 12, 62-70.

Tang, N., Mitchell, O., Mottola, G., Utkus, S., 2010. The efficiency of sponsor and participant portfolio choices in 401(k) plans. *Journal of Public Economics* 94, 1073-1085.

Tversky, A., Shafir, E., 1992. Choice under conflict: the dynamics of deferred decision. *Psychological Science* 3(6), 358-361.

Table 1. Plan-Level Descriptive Statistics

Panel A reports descriptive statistics for plan size and fund balances (both in U.S. dollars) for 401(k) plan data provided by Brightscope, Inc. Based on data filters such full return availability, the original sample of 17,386 plans is reduced to 7,975. *Plan Size* is the sum of all fund balances held within the plan while *Average Fund Balance* is the balance within each mutual fund. *Number of Fund Options* provides the size of plan menu for participants.

Panel B reports descriptive statistics on plan size by the type of company administering the 401(k) plan, i.e. third-party administrator (TPA).

Panel A. Plan and Fund Balance Size

Original Sample **N = 17,386**	Mean	Standard Deviation	10th percentile	25th percentile	Median	75th percentile	90th percentile
Plan Size (000's)	22,100	111,000	1,157	2,534	5,894	13,700	34,200
Average Fund Balance (000's)	1,228	8,456	63	133	313	734	1,865
Number of Fund Options	22.13	15.00	11	15	19	26	32

Final Sample **N = 7,975**	Mean	Standard Deviation	10th percentile	25th percentile	Median	75th percentile	90th percentile
Plan Size (000's)	31,827	136,058	1,141	2,684	6,697	17,172	52,676
Average Fund Balance (000's)	1,891	8,564	77	180	426	1,059	3,110
Number of Fund Options	18.21	6.78	10	13	17	23	28

Panel B. Plan Size by Category

Plan Size (000's)	Number of Plans	Mean	Standard Deviation	Min	Median	Max
Mutual Fund Families	2,340	60,089	272,790	37	13,448	9,641,714
Asset Management Advisory	519	40,030	204,915	49	6,557	2,602,408
Investment Banks	183	91,400	355,811	14	14,628	4,040,556
Large Commercial Banks	725	41,197	139,039	70	8,508	2,261,397
Small/Regional Comm. Banks	203	12,821	33,027	10	4,756	349,804
Insurance Firms	435	15,002	32,758	147	5,826	470,023
401(k) Services Companies	362	18,860	62,782	15	4,091	5,825,942
TPA Unknown	3,224	13,183	52,622	3	4,090	1,146,085

Table 2. Plan Menu Options Descriptive Statistics

Panel A provides frequency data on the different types of funds available in the final sample, across all plans (unconditional) and across only the plans that offer those types (conditional). The last three columns refer to the proportion of plan assets directed to the types of funds listed.

Panel B reports descriptive statistics on the number of different Lipper Investment Objective categories available in the sample plans.

Panel A. Fund Type Coverage

	Number of Fund Options (unconditional)		**Number of Fund Options (conditional)**		**% of Plans Assets Held in (unconditional)**		**% of Plans Assets Held in (conditional)**		**% of Plans Containing at least one**
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	
Domestic Equity Funds	12.36	6.01	12.37	6.00	68.18%	13.50%	68.24%	13.35%	99.92%
Domestic Bond Funds	1.82	1.16	1.88	1.13	8.97%	19.72%	9.25%	19.96%	96.98%
International Funds	1.86	1.16	1.92	1.13	14.36%	9.63%	14.82%	9.43%	96.90%
Money Market/ Stable Value/ GIC/ Annuity (MSGA)	0.63	0.57	1.06	0.29	8.80%	11.77%	14.88%	11.99%	59.12%
Company Stock	0.04	0.21	1.06	0.23	0.51%	3.95%	12.97%	15.36%	3.93%
Participant Loans					7.67%	9.32%	8.32%	9.43%	92.17%

Panel B. Lipper Objective Category Coverage

Unique Lipper Category Coverage by Plan

Mean	Standard Deviation	10th percentile	25th percentile	Median	75th percentile	90th percentile
13.21	6.03	6	9	12	17	22

Table 3. Benchmark Indices

Listed below are descriptive statistics on the performance of the two sets of benchmark sets used in the mean-variance spanning test. They cover the period 2004-2008. The performance is measured as monthly returns net of fees. Benchmark Index Set 1 is consistent with Elton et al. (2006), Tang et al. (2010), while Benchmark Index Set 2 is aligned with the indexes used for style analysis in Sharpe (1992).

	Monthly Returns	
	Mean	Standard Deviation
Benchmark Index Set 1		
MSCI EAFE Index	0.28%	4.76%
Barclays Capital Aggregate Bond Index	0.37%	1.34%
Credit Suisse High Yield Bond Fund	-0.04%	2.96%
Citigroup World Gov't Bond Non-US$ Index	0.50%	2.36%
Russell 1000 Growth	-0.22%	4.09%
Russell 1000 Value	0.01%	3.78%
Russell 2000 Growth	-0.05%	5.64%
Russell 2000 Value	0.14%	4.99%
Benchmark Index Set 2		
Barclays Government Intermediate	0.06%	0.93%
Barclays US Aggregate Long Government	-0.02%	2.89%
Barclays Investment Grade - Corporates	-0.32%	2.10%
Barclays US Agency Fixed Rate MBS	0.44%	0.91%
Citigroup World Gov't Bond Index World 5+Yr Non-US$	0.57%	2.57%
S&P 500 / Citigroup - Value	-0.28%	3.90%
S&P Mid-cap 400 / Citigroup - Value	-0.03%	4.75%
S&P Mid-cap 400 / Citigroup - Growth	0.01%	4.90%
S&P Small-cap 600 / Citigroup - Value	0.05%	4.95%
S&P Small-cap 600 / Citigroup - Growth	0.14%	5.01%
MSCI Europe Index	0.03%	4.99%
MSCI Pacific Index	0.13%	4.86%
S&P IFCI Emerging Market Index	0.77%	7.35%

Table 4. Spanning Test Results: Effect of Choice of Benchmark Indexes

This table summarizes spanning test results at 5% significance level. Panel A reports results with Benchmark Set 1 indices while Panel B reports results with Benchmark Set 2, as described in Table 3. "All Funds" refers to the number of plans that either span or do not span using the all funds within a given plan. "QDIA Funds" the number of plans that either span or do not span when the plans are restricted to the subset of funds picked by our QDIA procedure.

Panel A: Results Based on Benchmark Index Set 1

	Span	No Span	Total
All Funds	3,630	4,345	7,975
	45.52%	54.48%	100%
QDIA Funds	3,875	4,100	7,975
	48.59%	51.41%	100%

Panel B: Results Based on Benchmark Index Set 2

	Span	No Span	Total
All Funds	2,239	5,736	7,975
	28.08%	71.92%	100%
QDIA Funds	2,468	5,507	7,975
	30.95%	69.05%	100%

Table 5. Spanning Tests Based on Principal Component Analysis

This table summarizes spanning test results at 5% significance level. Principal component returns are constructed from returns of funds in plans. In "All Funds", principal component returns are based on all the funds in a given plan. In "QDIA", only the funds selected via the QDIA procedure are involved. Panel A and Panel B refer, respectively, to the index benchmark sets described in Table 3.

Panel A: Results for Benchmark Index Set 1

	Span	No Span	Total
All Funds	7,966	9	7,975
	99.89%	.11%	100%
QDIA Funds	7,970	5	7,975
	99.94%	0.06%	100%

Panel B: Results for Benchmark Index Set 2

	Span	No Span	Total
All Funds	7,967	8	7,975
	99.90%	0.10%	100%
QDIA Funds	7,975	0	7,975
	100%	0%	100%

Table 6. Illustration of a Subset Spanning Better Than the Original Set of Assets

This table compares the three matrices that affect the Wald statistic used to assess spanning relative to Benchmark Index Set 1 given in Table 3. In "QDIA", only the funds selected via the QDIA procedure are involved. Ω is the covariance matrix between all the coefficients in the multivariate regression of a set on the benchmark indices. $\tilde{\Sigma}$ is the matrix given in (5) and $\tilde{\Sigma}^{-1}$ defines the metric associated with the Wald statistic. In this example, the original set consists of $N = 30$ funds and its QDIA subset consists of $N = 11$ funds. Ω is of size 270 x 270 for the original set and 99 x 99 for the QDIA subset. $\tilde{\Sigma}$ (and thus its inverse) is of size 30 x 30, for the original set, and 11 x 11, for the QDIA subset.

	Ω		**$\tilde{\Sigma}$**		**$\tilde{\Sigma}^{-1}$.**	
	Original	**QDIA**	**Original**	**QDIA**	**Original**	**QDIA**
Min	-14.3914	-24.7539	-0.2842	-12.1019	-3.6234	-0.1804
Max	27.2468	43.7524	3.2838	22.6728	13.0576	3.2352
Median	0	-2.00×10^{-6}	0.0238	0.0244	-0.0205	0.1043
Frobenius norm	56.6224	136	3.5996	51.4421	26.0142	3.4296
Determinant			0.00148	1.31×10^{6}		