

The Adequacy of Investment Choices Offered
By 401(k) Plans

Edwin J. Elton*

Martin J. Gruber*

Christopher R. Blake**

December 2004

* Nomora Professors of Finance, New York University

** Professor of Finance, Fordham University

The Adequacy of Investment Choices Offered by 401(k) Plans

Abstract

Defined-contribution plans represent a major organizational form for investors' retirement savings. Today more than one third of all workers are enrolled in 401(k) plans. In a 401(k) plan, participants select assets from a set of choices designated by an employer. For over 60% of 401(k)-plan participants, retirement savings represent their sole financial asset. There have been a number of studies of the decisions made by 401(k) participants, but the choices made by 401(k) participants are the product of two different decisions: what is offered and what is chosen. There have been no studies of the choices offered to 401(k) participants. This paper analyzes the adequacy and characteristics of the choices offered to 401(k)-plan participants for over 400 plans. We find that, for 62% of the plans, the types of choices offered are inadequate, and that over a 20-year period this makes a difference in terminal wealth of over 300%. We find that funds included in the plans are riskier than the general population of funds in the same categories, but have a slightly higher rate of return. We also examine the performance of the specific mutual funds held by the plans, and we find that the performance is better than the average performance of similar randomly selected funds. However, we find that the performance difference is roughly equal to the difference in expenses between funds selected by plans and randomly selected funds. We study the characteristics of plans that are associated with adequate investment choices, including an analysis of the use of company stock, plan size, and the use of outside consultants.

Keywords: 401(k) plans, pension, spanning, portfolio, investment choices

JEL classification: G11, G12, G23, E21

“You Can’t Make a Silk Purse Out of A Sow’s Ear”

A major trend in pension plans offered by companies is a movement from defined-benefit plans to defined-contribution plans. The majority of defined-contribution plans offered by companies are 401(k) plans. More than one third of all workers are enrolled in 401(k) plans, and these plans have over one trillion dollars under management.

The value of any 401(k) pension plan to any participant is determined by two decisions: the set of investment choices offered to the participant in a plan, and what the participant invests in from among those choices. These are two very different decisions. The first decision is made by the plan administrator, the second is made by the participant in the plan. There is a large amount of research on the participant’s investment behavior, given the choices available to the participant.¹ However, none of this research examines the impact of the choices offered in a plan on the ability of plan participants to construct desirable portfolios.

Examining the choices offered to the participants in any plan is important, because an investor faced with an inappropriate set of choices cannot construct an efficient portfolio no matter what weights he places on the various offerings. There are two ways that the choices offered can be inappropriate: offering an insufficient number and type of choices to allow the construction of desirable portfolios, and offering poor-performing investment choices of any given type. In theory, the first type of suboptimality could be alleviated by holding funds outside the plan. However, for over 60% of plan participants, the 401(k) investments are their sole financial assets.² Even for those participants who hold other financial assets, the 401(k) assets are

¹ Examples of this extensive literature are Benartzi and Thaler (2001), Liang and Weisbenner (2002), Huberman and Sengmuller (2003), Agnew and Balduzzi (2003), Ameriks and Zeldes (2001), Nadrian and Shea (2001).

² Investment Company Institute (2000). This study reports that more than 60% of 401(k) plan participants have no other security investments (stocks, bonds, etc...) other than a bank account. Moreover, Choi, Laibson,

likely to be the bulk of their financial assets, so that plan offerings are likely to severely restrict the portfolios they can hold.

What choices should a corporation offer to plan participants? For those participants for whom 401(k) investments are their sole financial assets, the corporation should offer a sufficient set of investment alternatives so that the investor could construct the same efficient frontier that he or she would obtain if there were choices from a reasonable set of alternatives. Investors who have other financial assets would not be hurt by such a strategy, so this strategy is dominant for all investors.³

In this paper we find that only 38% of pension plans offer a set of choices that allows investors to construct an efficient frontier equivalent to one constructed from a reasonable set of alternative choices and that the loss because of this to investors is substantial. Second, we find that plans offer funds that have better performance than randomly selected funds but that the difference is about the same size as the expense difference between the funds they select and randomly selected funds. Third, we show that the portfolio of mutual funds offered by 401(k) plans is riskier than randomly selected mutual funds under realistic assumptions. Fourth, we show that the inclusion of company stock does not necessarily hurt investors.

This paper is divided into five sections. In the first section we discuss the data used. In the second section we explore issues of how well the fund offerings span the efficient set. In that section, we not only examine statistical tests, but we also examine economic significance (effect on participants' returns) of a failure to provide appropriate offerings. In the third section we explore the characteristics (risk-adjusted return and risk) of the specific funds selected relative to

Madrian, and Metrick (2004) estimate that for households with annual incomes between \$20,000 and \$70,000, the median household has less than one month's worth of income invested outside of their 401(k) plan.

the population of funds that could have been selected. In the fourth section we examine the effect of offering company stock on plan risk and the efficient frontier. In section five we examine whether other characteristics of the plans affect the appropriateness of the investment choices offered to plan participants. Finally, in section six we summarize our results.

I. Data

Our data were provided by Moody's Investor Services. Moody's collects data by means of a survey of pension plans offered by both for-profit and non-profit entities (collected in 2002 with information through 2001). From this data set we selected all 401(k) plans that employed publicly available mutual funds for participant choices. However, we did not exclude plans that offered, in addition to mutual funds, non-public money market funds, GICs and stable value funds and/or company stock. We were able to identify 680 401(k) plans for which the CRSP mutual fund database contained at least some data on each of the mutual funds offered in the plan. Of the 680 plans, 417 had at least five years of monthly total returns data in the CRSP database for every mutual fund they offered.⁴ For each of these plans we collected data on the mutual funds offered, historical returns for each mutual fund, and the names and characteristics of the firms offering the plans.

Table 1 shows the number of distinct investment choices offered by the 680 plans mentioned above. The median number of 401(k) plan offerings is eight. Approximately 12% of the 401(k) plans offer four or fewer investment choices, and approximately 11% offer 13 or more investment alternatives. The median number of investment offerings we report is somewhat less

³ If a plan is administered by an external party, the administrator may charge additional fees if the company wishes to include funds outside the external party's normal offerings. However, there are so many plan administrators for a company to choose from that this is unlikely to have an effect on our findings.

⁴ When later we draw general samples of mutual funds for comparison purposes, we use the same selection procedure so that comparisons are not biased.

than that reported by Huberman and Sengmuller (2003). Huberman and Sengmuller's data sample came from 401(k) plans managed by Vanguard. Many plans restrict their offerings to one fund family. Vanguard is one of the largest mutual fund families in terms of number of funds offered. Thus it is not surprising, and is consistent with what we observed in our sample, that plans managed by Vanguard offer more choices than would be observed in the population.

Table 2 shows the percentage of plans that offer various types (using ICDI classifications) of investment choices to their participants. The most common investment choice (offered by 97.4% of plans) is a domestic equity fund. The next most common offering (86.8%) is an alternative such as a GIC or money market fund, where interest is intended to be the only source of return. Other common offerings fall in the following categories: domestic bond funds (71.5%), domestic mixed bond and stock funds (80.6%), and international bond and/or stock funds (75.1%). The high percentage of 401(k) plans that offer international funds is surprising, given the much lower percentage international funds constitute of mutual funds publicly available to investors. Finally, 22.9% of the 401(k) plans offer company stock as an alternative for their participants.

Forty-eight of the 680 plans offered pension participants at least one specialized fund as an alternative choice; there were a total of 56 such choices. Thirty-one of these specialized funds were science and technology funds, six were real estate funds, five were telecommunications funds, four were healthcare funds, four were natural resources funds, four were utilities funds, one was an e-commerce fund and one was a financial services fund. There is no relationship between type of specialized funds offered and the type of firm offering the 401(k) plan.⁵ The

⁵ We noted that 33 of the 56 specialized funds were T. Rowe Price funds, suggesting that recommending inclusion of a specialized fund is a strategy employed to market T. Rowe Price funds to 401(k) plan administrators.

large number of science and technology funds offered at the date our sample was constructed suggests that some plan administrators were including the then-current “hot” sector.

II. Adequacy of Investment Choices

In this section we examine the adequacy of the investment choices offered by 401(k) plans.

In order to determine if 401(k) plans offer their participants appropriate investment choices, we need to hypothesize an adequate set of alternative investment choices. There are two approaches that can be taken. The first approach draws on the field of financial economics, where extensive literature exists that discusses indexes that are necessary and sufficient to capture the relevant return characteristics for a range of investment choices. A second approach uses a set of classifications that the industry finds relevant for classifying investment portfolios. These classifications represent the investment industry’s attempt to separate mutual funds into groups of funds that behave similarly and define a complete set of relevant investment choices. In the body of the paper we use the literature of financial economics to define a relevant set of indexes for portfolio choice. In the appendix we motivate our use of these indexes by comparing their spanning properties and performance with the set of indexes commonly used by industry.⁶

The indexes employed will now be described in greater detail. For common stocks, we classify by value versus growth and by size as advocated by Fama and French (1994). We classified size into two groups: small-mid-cap and large-cap. Each of these two groups was then further divided into value and growth. All four indexes were taken from Wilshire. We chose Wilshire indexes because there exist tradable funds that attempt to match each of the indexes.

⁶ As discussed in the appendix, industry indexes were dominated by the research-based (RB) indexes in the sense that the industry indexes did not explain returns as well as the RB indexes and adding any industry index to a portfolio of RB indexes did not improve the efficient frontier.

For bonds, we combined a general bond index, including governments and corporates, and a mortgage-backed index. We also employed a high-yield index. This division is supported by Blake, Elton and Gruber (1993), who found this division was sufficient to capture differences in return across bond funds. We used the Lehman U.S. Government/Credit index for the general bond index, the Lehman Fixed-Rate Mortgage-Backed Securities index for the mortgage index, and the Credit Suisse First Boston High-Yield index for the high-yield bond index.⁷ We also included the Salomon Non-U.S.-Dollar World Government Bond index for international bonds and the MSCI EAFE index for international stocks.

Since returns on all mutual funds are computed after expenses, we deducted expenses from each of our indexes. For each of our indexes, we used the expense charge of the index fund (including exchange-traded funds) that most closely matched the index. If there were multiple index funds matching the index, we used the expense charge of the lowest cost fund. In what follows we refer to these eight indexes as “Research-Based” indexes, or “RB” index

We now examine whether the choices given investors allow construction of an efficient frontier similar to that obtained by the 8 RB indexes. To do this we draw from the literature on spanning tests.

A. Methodology

The purpose of the intersection test is to examine whether, given a riskless rate, a particular set of benchmark assets is sufficient to generate the efficient frontier or whether including (long or possibly short) members of a second set of assets would improve the efficient frontier at a statistically significant level.

⁷ Originally the indexes for mortgage-backed securities and small- and mid-cap stocks were included separately, but empirical tests showed there was no improvement in explaining returns by including them separately. See appendix A.

As De Roon et al (2001) have shown, if the optimal (tangent) portfolio consists of K benchmark assets, then intersection is a test of the impact of restricting the intercept (α) in the following time-series model:

$$R_{it} - R_f = \alpha_i + \sum_{k=1}^K \beta_{ik} (R_{kt}^B - R_f) + \varepsilon_{it} \quad (1)$$

Where

R_{it} = the return on non-benchmark asset i ($i = 1, \dots, N$) in month t ;

R_f = the risk-free rate;

R_{kt}^B = the return on benchmark asset k in month t ;

ε_{it} = the error term for asset i in month t .

β_{ik} = the sensitivity of the non-benchmark asset to benchmark asset k in period t .

When short sales are allowed, intersection occurs if, for all of the N non-benchmark assets jointly, the α_i are not statistically significantly different from zero, i.e., the restrictions are

$$\alpha_i = 0 \quad \forall i \quad (1a)$$

When short sales are not allowed, the right-hand side of equation (1) includes returns on only those benchmark assets that are held long in the optimal portfolio of benchmark assets.⁸

Intersection occurs if, for all of the N non-benchmark assets jointly, the α_i are not statistically significantly positive, i.e., the restrictions are

$$\alpha_i \leq 0 \quad \forall i \quad (1b)$$

The logic behind the test can be easily understood. In the case of short sales allowed if an asset had a positive (or a negative) alpha, then including the asset long (or short) would improve

the efficient frontier. Without short sales, only the inclusion of an asset with a positive alpha would improve the efficient frontier.

To test whether, given a riskless rate, we have a set of benchmark assets that spans the relevant space; we simply have to test the unrestricted model (equation (1)) against the model with the restrictions on alpha. This involves employing equation (1) using the restrictions (1a) ($\alpha_i = 0 \forall i$) for the case where short sales are allowed and using the restrictions (1b) ($\alpha_i \leq 0 \forall i$) for the case where they are not allowed.

To test whether or not the restrictions hold, we use the likelihood ratio test statistic suggested by Gallant (1987) with small-sample adjustment. The likelihood ratio test is:

$$L = T \left(\ln |\tilde{\Sigma}| - \ln |\hat{\Sigma}| \right) \quad (2)$$

where T is the number of time-series observations, $\tilde{\Sigma}$ is the estimated variance-covariance matrix of the residual errors of the N non-benchmark assets from the restricted equation, and $\hat{\Sigma}$ is the estimated variance-covariance matrix of the residual errors of the N non-benchmark assets from the unrestricted equation. L is asymptotically distributed as chi-squared with q degrees of freedom, where q is the number of parametric restrictions. For small samples such as we have cross-sectionally, Gallant recommends the use of the F distribution with degree-of-freedom corrections instead of the chi-squared distribution. The small-sample adjustment is simply to compare L to $q \times F_x$, where F_x is the F statistic at significance level x with q numerator degrees of freedom and $T \times M - p$ denominator degrees of freedom, and where M is the number of equations estimated and p is the number of parameters. If L is greater than $q \times F_x$, then the null hypothesis that the restrictions hold is rejected.

⁸ These benchmark assets can be easily identified by solving a quadratic programming problem.

B. Results

We now employ the spanning tests described above to see if plans offer participant choices that allow a construction of an efficient frontier that is not statistically different from that obtained by the eight RB indexes.

Earlier we argued that an investor could be satisfied with a choice from among eight research-based indexes. The question is whether the choices offered by 401(k) plans span the space delineated by the eight RB indexes; if they do not, then optimal investment choices are not being offered. Since plan participants can not short sell assets in their 401(k) plans, we use the intersection test described earlier for the case where short sales are not allowed. In this application the benchmark assets are chosen from the mutual funds offered by any plan, while the non-benchmark assets are the eight RB indexes discussed earlier. The results of the intersection tests are shown in Table 3.⁹ Recall that the intersection tests would reject spanning (say that the choices offered are not adequate) if adding RB indexes to the optimal portfolio of funds offered by any plan would improve the efficient frontier at a statistically significant level. Plans holding four or fewer funds rarely offer a set of funds that span the eight RB indexes. For these plans there are more RB indexes than fund offerings. However, it is possible that a small set of funds spans the larger set of RB indexes, either because some of the RB indexes are not desirable investments or because some of the funds are combinations of two or more of the RB indexes. However, this does not happen for funds offering a small set of investment choices. For plans holding seven or more funds, we find that about 54% of the plans offer investment choices that span the relevant space investors are interested in.¹⁰ Of course, the glass is also half empty in

⁹ The results in Table 3 exclude the same 11 plans excluded in Table 4, leaving a total of 406 plans.

that 46% of the plans leave investors unsatisfied. Finally, it is not until plans offer 14 or more investment choices (4.2% of all plans) that virtually all plans offer investment choices that span the space investors should be interested in.

Of the 406 plans, only 38% span the space obtainable from the eight RB indexes.¹¹ While some 401(k) plans offer participants a rich enough selection of investment choices to satisfy their needs, clearly a number of 401(k) plans do not do so.¹²

Before leaving this section, it is worthwhile examining the loss in return to 401(k) plan holders due to plans not spanning the relevant space. For the 406 plans in our sample to have the same Sharpe ratio as the optimum portfolio comprised of the 8 RB indexes, the average return on the plans' optimum portfolios, holding risk constant, would have to increase by 1.81% per year. For the 249 plans that do not span the space, average return would have to increase by 3.16% per year to match the Sharpe ratio on the 8 RB indexes. The 3.16% increase in return is equal to 42% of the return on the 8-RB-index portfolio. Thus, investors in 401(k) plans are sacrificing significant return because plan administrators are offering an incomplete set of investment alternatives.¹³

¹⁰ The sample of 417 plans was constructed to include only those 401(k) plans where all offerings had five years of history. The distribution of the number of offerings with that restriction differs from the distribution of the number of offerings by 401(k) plans in general. If we apply the distribution of "yes" and "no" shown in Table 3 to the distribution of investment choices shown in Table 1 and assume that all plans with 17 or more investment choices span, the percentage rises to 58%.

¹¹ For the reasons discussed in the prior footnote, we apply the distribution of "yes" and "no" shown in Table 3 to the distribution of investment choices shown in Table 1, counting each plan offering one investment choice as a "no" and each plan offering 17 or more investment choices as a "yes." Applying these rules, the percentage of plans that span is 40%.

¹² As a further check on plans spanning, we considered whether plans spanned the space of the simplest set of choices we could think of: a broad stock market index (the Wilshire 5000 index), a bond market index (the Lehman U.S. Government/Credit index), and an international index (the MSCI EAFE index). We adjusted the returns of the 3 indexes to reflect normal management fees (just as we did for the 8 RB indexes). With this limited set of 3 indexes, more plans offered choices that spanned the indexes' space. However, 42 of the 406 plans still did not span, and over half of those plans offered 6 or more choices.

We will now examine the characteristics of mutual funds selected by pension plans to see how that impacts plan characteristics.

III. Characteristics of the Specific Mutual Funds Selected

In this section we will explore the characteristics of the funds selected by plan administrators given the number and types they select. We will initially explore whether the funds they selected have positive alphas. We will then explore the risk characteristics of the funds selected.

A. The Risk-Adjusted Performance of Plan Funds

The analysis up to this point has been concerned with whether pension plans offer participants adequate types of choices. A second and very interesting question is: given the types of choices offered to participants, is management selecting individual mutual funds that outperform random selection from those types? Even if management is offering the plan participant enough choices, the investor may be forced to choose from among mutual funds that are dominated by other funds of the same types not offered by the plan. To determine this, we need to construct a model to measure performance.

For each mutual fund i ($i = 1, \dots, n$), we will use the alpha (α_i) estimated from a multi-index model to measure the fund's performance. Alpha is the intercept of the following time-series regression:

$$R_{it} - R_{Ft} = \alpha_i + \sum_{k=1}^K \beta_{ik} I_{kt} + \varepsilon_{it} \quad (4)$$

where

1. R_{it} is the return on mutual fund i in month t ;

¹³ These differences are much larger than any possible differences due to expense ratios between index funds and active portfolios. See Elton, Gruber and Blake (1996) for estimates of expense ratios.

2. R_{Ft} is the riskless rate in month t ;
3. I_{kt} is either the return on index k in month t if the index is the difference between two return series, or the excess return on index k in month t above R_{Ft} if the index is based on a single return series (the actual indexes are described below);
4. β_{ik} is the sensitivity of fund i to index k ;
5. ε_{it} is the random error term for fund i in month t ;

Mutual funds were divided into three types: stock funds, bond funds and international funds. For stock funds we used a five-index model: the excess return (over the riskless rate) on the S&P 500 index, the return on the Fama-French small minus big (SMB) factor, the Fama-French high minus low (HML) book to market value factor, the excess return on the Lehman Government/Credit index, and the excess return on MSCI EAFE index. This is similar to the model used by Elton, Gruber and Blake (1996). Two indexes require some comment. First, the bond index is needed both because the stock category includes many funds that are combinations of bonds and stock, such as balanced and income funds, and because funds in common stock categories such as aggressive growth or long-term growth often hold part of their portfolios in long-term bonds. Failure to include a bond index imputes to alpha any return on long bonds different from the riskless rate. The other non-standard index is the international index. During the period of this study many domestic stock funds included international stocks in their portfolios, usually in the form of ADRs. Again, failure to include an international index could cause an alpha to include the effects of a fund's international holdings.

For bond funds we used the excess return on the following four indexes: the Lehman Government/Credit index, the Lehman Fixed Rate Mortgage-Backed Securities index, the Credit

Suisse/First Boston High-Yield Index, and the Salomon Brothers Non-Dollar World Government Bond Index. The first three indexes are supported by the work of Blake, Elton and Gruber (1993). The addition of the international index is needed to capture the tendency of some bond funds to include international bonds in their portfolio over this period.

Finally, for international funds we used the excess return on the following five indexes: The S&P 500 index (since world funds invest in part in the U.S.), three MSCI indexes (Europe, Pacific and Emerging Market), and the Salomon Brothers Non-Dollar World Government Bond Index.

We computed alphas for 27 months following the date of our sample. (27 months is the longest period for which we have data.) Using subsequent return data to evaluate performance eliminates the bias that would result if 401(k) plans added new funds with superior past performance. In this case evaluation using the period prior to the sample year would pick up return not available to investors (if a fund was newly added) and would bias 401(k) plans' relative alphas upwards. A few of the funds did not have 27 months of data. In these cases we used the 27 months of returns up to the time a fund disappeared to estimate the coefficients of the regression. We then computed an adjusted alpha by taking a weighted average of the alpha on the fund over the relevant period (overall alpha plus residuals over the period subsequent to the sample year and before the fund disappeared) and the average alpha of the remaining funds in the plan (overall alpha plus residuals subsequent to disappearance date). This assumes the plan administrator did not add a new fund when the old ceased to exist and that investors transferred their investment to the remaining funds.

Alpha from a multi-index model is widely accepted as a measure of both relative and absolute performance. We made one added adjustment to obtain our performance measure. The

overwhelming evidence is that alpha is on average negative for mutual funds. Thus a negative alpha for the mutual funds offered by 401(k) plans would not indicate that the manager of these plans selected funds that were poorer than random selection. To ascertain whether management is doing a good job of selecting funds, we subtract from the alpha for each fund selected by a pension plan the average alpha from a randomly selected sample of funds from the same category (stock, bond or international). We call this difference the “differential” alpha. The population of funds we use as a comparison consists of all mutual funds that exist as of the end of 2001 and have five years of history. These are the same criteria we used when selecting 401(k) plans to include in our total returns sample.

Table 4 shows the average alphas by category for the funds selected by the 401(k) plan administrators as well as the differential alphas. Although 401(k) plan administrators selected funds with poor performance (negative alpha for each category) they selected funds that had smaller negative alphas than random selection. However, only for one category, bond funds, was the average differential alpha significantly different from zero.

What can account for the superior performance of 401(k) pension fund administrators? From Table 4 it is clear that the difference in performance is related to the difference in expense ratios for the three categories of funds they buy. Differential expenses account for 150% of the differential performance for common stock funds, 65% of the differential performance of bond funds, 59% of the differential performance of international funds, and over 100% of the differential performance on average across all types. We don’t know the decision-making process of plan administrators. Thus, we cannot tell whether the lower expenses are because of a choice of the plan administrators or are a characteristic of the mutual funds that are sold to 401(k) plans.

In summary, plan administrators have a slightly higher excess return than a randomly selected set of funds of the same type but the majority of the difference is due to plan administrators holding funds with lower expense ratios.

B. Risk Characteristics

In this section we examine the risk characteristics of funds selected by 401(k) plan administrators. The average individual monthly return variance of the mutual funds held by 401(k) plans is 26.76. If 401(k) plan sponsors selected mutual funds randomly but maintained the same percentage in each ICDI category as the aggregate of all plans, the average fund variance would have been 30.49. If instead we simply computed an average fund variance across all mutual funds, weighting each fund equally, the variance would be 31.26.¹⁴ Thus 401(k) plan administrators select mutual funds with a lower fund variance both relative to what it would be if they randomly selected funds while maintaining the aggregate plan proportions in ICDI categories and relative to what it would be if they simply randomly selected across all available funds.

The other element that affects portfolio variance is correlation. The average pairwise correlation of monthly returns among funds selected by 401(k) plans is 0.60, while for random selection, maintaining ICDI proportions, it is 0.55. The difference is statistically different at the 1% level. Thus plan administrators select funds that are more highly correlated than the average correlation between pairs of funds. One explanation for the higher correlation is that plans tend to restrict their choice to one fund family. Elton, Gruber and Green (2004) have shown that intra-

¹⁴ The random selection leads to more small funds being selected than 401(k) plans actually hold. If we control for this by eliminating funds less than \$50 million in size, the variance of randomly selected fund is reduced to 29.592 if we maintain the same percentage in each ICDI category as funds selected, or 30.726 using equal probability of selection for all funds.

family correlation is much higher than inter-family correlation, and thus limiting choices to one family increases correlation.

What does the higher correlation and lower variance imply for the riskiness of portfolios of 401(k) plan mutual funds? To examine this we compare the risk of the funds held by 401(k) plans to the risk of “synthetic” 401(k) plans constructed by using random selection of publicly available mutual funds. We have several ways to implement this random selection. The simplest, and most direct, method is to make the odds of selecting a fund from any ICDI category equal to the proportion of that category held by our sample of 401(k) plans. Within a category (e.g., aggressive growth) the odds of choosing any single fund are made equal.

This is an extremely naïve selection rule that ignores completely the correlation between ICDI categories. A slightly less naïve strategy would force all synthetic plans to hold at least one randomly selected bond fund and one randomly selected stock fund. Examining the holdings of the actual 401(k) plans shows that this is a strategy followed by almost all plans. Thus our second random-selection strategy, called “constrained random selection,” follows the same random-selection rules describe above except that all synthetic plans are forced to hold at least one bond and one stock fund.

To calculate portfolio variances for both the actual 401(k) plans and the synthetic 401(k) plans, we need to formulate a rule to represent the investment weighting for a hypothetical plan participant. Many investors equally weight their 401(k) plan offerings (see Benartzi and Thaler (2002) and Liang and Weisbenner (2002)), so we initially use one divided by the number of a plan’s investment choices to represent a participant’s chosen investment weight in each of the plan’s mutual fund offerings.¹⁵

¹⁵ We exclude from the investment choice sets company stock, GICs, stable value funds and money market funds.

Table 5 presents the average values (by number of funds offered) of the variances for the monthly returns of the actual 401(k) plans as well as the variances that would occur if plan sponsors selected funds at random using either of the selection rules described above.¹⁶ The first thing to note from Table 5 is that while on average the variance of return on actual 401(k) plans is lower than the variance would have been if plan sponsors had randomly selected a set of mutual fund, it is higher once we make the realistic assumption that the synthetic plans have at least one bond fund and one stock fund. Plans on average have a variance that is 2.29 lower than that using random selection of funds but 2.087 higher than that using constrained random selection.¹⁷ Both differences are statistically significant at better than the 0.01 level. It is also interesting to note that as plans offer more investment choices (beyond three); the overall risk is reasonably flat.

There are an infinite number of alternative weighting schemes that can be examined, and none have any special empirical support. While we will not explicitly examine alternatives to equal weighting, we can gain insight into these results with the following logic. Assume a plan offered five funds and that the first two were held in a proportion different from that of the other three. This is equivalent to constructing a portfolio of the results shown for five and two funds shown in Table 5. For example, if the investor's proportions in five funds were 35%, 35%, 10%, 10%, and 10%, these proportions can be obtained by combining an equally weighted two-fund portfolio, with 50% in each of the first two funds, with an equally weighted five-fund portfolio with 20% in each of the 5 funds, where the investment is 50% in each portfolio. The variance of this combined portfolio depends on the variance of the equally weighted two- and five-fund

¹⁶ Only one of our 417-sample 401(k) plans offered 17 funds; therefore we do not report average values for 17-fund plans in Table 4.

portfolios and the correlation between them. For constrained random selection, the variances of equally weighted portfolios are generally smaller than those for actual plan portfolios. Further, we know correlations are smaller for random portfolios than for actual plan portfolios. Thus the variance of unequal weighted plan portfolios should be higher because both the variances and correlations are higher than for portfolios from constrained random funds. When we compare plan funds to unrestricted random funds, unequal weighting should make the differences between plan portfolios and unrestricted random portfolios smaller, for while the variance is lower for the plan sample, the correlations are higher than those of the random funds.

If we assume at least one bond fund and one stock fund are included in the offering, which is consistent with most plans and hence the most realistic case, plan participants' portfolios are likely to be more risky than if the plans had added funds randomly, regardless of whether the funds are combined using equal or unequal weights.

IV. Company Stock

The analysis to this point has ignored company stock as an asset in 401(k) plan offerings. We know that there is a tendency for plan participants to place a disproportionate fraction of their plan assets in company stock. In this section we explore whether including a firm's own stock as one of the investment choices in the 401(k) plan is harmful *per se*. We examine the impact of including company stock by analyzing spanning, plan risk, and the Sharpe ratios.

On average, companies offering company stock as an investment choice offer the same number of mutual fund choices as those that do not offer stock; therefore, companies offering company stock do not offer plan participants fewer fund choices as a mechanism to encourage participants to hold more company stock.

¹⁷ Because average variances cannot be constructed using constrained random selection for plans with only one fund, the reported averages and significance tests exclude the ten 401(k) plans in our sample that offered only one

The most important test of the impact of company stock is the spanning test. Does including company stock increase the number of 401(k) plans that have offerings that span the space of our eight RB indexes? Analysis shows that whether company stock is included in the choice set or not, there is no change in the number or identity of the plans for which spanning takes place.

To examine the effect of company stock on overall risk, we took all plans that offered company stock as an investment choice for which stock returns existed over our five-year period. For these plans we computed the variance using data for the last five years of an equally weighted portfolio of all offerings, with and without the company stock. For the companies offering company stock, when the company stock was included, the variance of the portfolio of 401(k) offerings using the $1/n$ Rule went up by 3.17. Of the 55 plans for which we have data, 36 have a higher variance when company stock is included in the portfolio. The 3.17 increase in variance associated with including company stock is a percentage increase of about 19%, and using a one-tailed pairwise t -test, this increase is statistically significant at the 1% level ($t = 3.6$). Empirical evidence supports over-investment in company stock (see Huberman and Sengmuller (2003)). Increasing the proportion invested in company stock should increase the risk differences.

Although the inclusion of company stock leads to risk increasing, Sharpe ratios might also increase. To examine this we examined the Sharpe ratios for optimal portfolios with no short sales. When company stock was not allowed to enter the optimal portfolio, the average Sharpe ratio was 0.240. When company stock was allowed to enter, the Sharpe ratio increased slightly to 0.255. Remember that including more securities in the population will in general increase the Sharpe ratio. If we control for this by comparing the increase from including company stock with

the increase from including a randomly selected mutual fund, the difference is close to zero and is neither statistically significant nor economically significant. This is true despite the fact that company stock enters the optimal portfolio in 26 out of 55 cases.

In summary, the inclusion of company stock doesn't change the set of plans that span the space of the RB indexes. Including company stock increases risk. This increase in risk is more than offset by an increase in return, resulting in a very slight improvement in the Sharpe ratio. However, the increase in the Sharpe ratio is about the same as it would be if we randomly included an additional mutual fund rather than the common stock. Considering the 401(k) plan as the participant's sole financial asset, the inclusion of company stock in a plan seems to neither improve nor harm the investor making intelligent 401(k) plan choices. However, since a plan participant's labor income may be highly correlated with the performance of the company stock, a portfolio including labor income, 401(k) mutual funds and the company stock may be significantly more risky than a portfolio excluding the company stock. In addition, since research has shown that many individuals hold too large a fraction of their plans assets in the stock of their own company, the inclusion of company stock can unduly increase the risk of a participant's portfolio.

V. Plan Characteristics

In this section of the paper we examine the relationship between plan characteristics and performance. Before we turn to performance *per se*, we want to examine one characteristic of plans that seems to have a major impact on how management behaves and which serves as a parameter that might affect performance.

In Table 6 we divide all plans by the size of assets invested in each plan into 10 deciles.¹⁸ The average size of the plan in each decile is shown in the second column. There is a wide variation in plan size, with the average plan in the tenth decile over 300 times as large as the average plan in the first decile. The first question we examine is whether plans with more assets under management offer participants more investment choices. As shown in Table 6, there is a clear and statistically significant relationship (at the 1% level) between plan size and the number of investment choices offered. Since from our spanning tests we know that more investment choices are generally better for investors, this suggests that large plans *ceteris paribus* offer an advantage to the 401(k) participants.

Are companies that manage large plans more sophisticated than companies that manage small plans? In particular, are companies with large 401(k) plans more likely to hire outside consultants and use sophisticated strategies such as utilizing futures and options, hedging strategies and quantitative methods? As shown in Table 6, a higher percentage of larger plans hire outside consultants and engage in more sophisticated investment strategies. The relationship of both with size is statistically significant at the 1% level in both cases.

We next examine the relationship between size and whether a plan votes proxies in the companies it owns. Proxy voting can be interpreted as either another measure of sophistication or as a measure of social consciousness. We find at best a weak positive relationship, one that is not statistically significant.

Finally, we examine the relationship between the size of plan assets and the probability of a company including its own stock in its 401(k) plan. Not surprisingly, large plans show a

¹⁸ We were unable to obtain plan size data for 28 of the 417 tracked plans; the size deciles were formed using the remaining 389 plans.

stronger tendency to include company stock in the plan than do small plans, and this relationship is significant at the 1% level.

We now turn to an examination of whether the use of outside consultants or sophisticated strategies improves the position of plan participants. To do so we examine their impact on number of plan investment choices, spanning, and average risk-adjusted return.

It is clear from Table 6 that there is an association between average plan size and both the employment of outside consultants and the use of sophisticated investment tools. It is also clear that larger sized plans have more investment choices. Therefore, if we want to discover whether employing outside consultants or using sophisticated strategies leads to more investment choices *per se* we need to control for plan size. We divided all plans into two groups based on whether or not they employed outside consultants and two groups based on whether or not they used sophisticated strategies. For each plan in the group we calculated the difference between the number of investment choices the plan actually offered and the number of investment choices we would expect given the plan's size. We then computed the average difference for the group employing outside consultants (or sophisticated strategies) and the group that did not. The significance of this difference was then tested using a standard *t*-test. Although the sign was as expected, the relationship between the number of investment choices and the use of outside consultants or sophisticated strategies was not statistically significant at meaningful levels of significance.

A second issue we examine is whether employing outside consultants or sophisticated strategies leads to a greater likelihood of the investment choices offered spanning the investment space. From portfolio theory we know that the greater the number of investment choices offered *ceteris paribus*, the more likely the offerings will span the space. Thus, to examine this question

we need to control for number of investment choices. We divided the plans into two groups based on whether or not they employed consultants, and two groups based on whether they used sophisticated strategies. Within each group, given the number of investment choices offered, we compute differences in actual proportions that span and expected proportions that span. We then compared these differences between the group that employed outside consultants and the group that did not and the differences between the group that employed sophisticated strategies and the group that did not. For each case, the difference, while in the expected direction, was not statistically significant.

Finally, we examine average differential risk-adjusted returns between plans that employ consultants or employ sophisticated strategies and those that do not. Differential risk-adjusted returns are the risk-adjusted return of the plan fund minus the average risk-adjusted return for the population of funds within the same ICDI category. Risk-adjusted returns are then averaged by plan and then across plans. The differential return on mutual funds for plans that employ outside consultants is lower than on those that don't, although the difference is not significant. Those plans which use sophisticated strategies have higher differential alpha than those that don't at the 1% significance level.

IX. Conclusion

In this paper we examine the reasonableness of the investment choices offered by 401(k) plans. This is an important subject. The payoff of a pension plan to any investor is the product of two different decisions: what the investor is offered, and what he or she chooses from what is being offered. While a lot of attention has been paid to participant choice in the literature of financial economics, no attention has been paid to the relevancy of the choices offered to

participants. If investors are given an inferior set of choices in their plan, the effectiveness of their choice is severely constrained.

The major findings of this paper concern the adequacy of plan offerings. We use spanning tests to see if the plan offerings span the space offered by the eight RB indexes. Only 38% of 417 plans span the space defined by the eight RB indexes. This means that, for 62% of the plans, the plan participants would be better off with additional investment choices. In fact, if these plans spanned the 8 RB indexes, participants' average return would improve by 3.2% per year, which is 42% of the return on an 8-index portfolio with the same level of risk. While significant on a 1-year basis, over a 20-year period (a reasonable investment horizon for a plan participant), the cost of not offering sufficient choices makes a difference in terminal wealth of over 300%.¹⁹ Since, for more than one half of plan participants, a 401(k) plan represents the participant's sole financial asset, the consequences are serious.

We next examine the return and risk characteristics of the mutual funds held by plans. We find that the plans select funds that show negative performance, but the performance is still better than the average performance of similar randomly selected funds. The differential performance between funds selected by plans and randomly selected funds is roughly equal to the difference in expenses between funds selected by plans and the universe of all funds.

We next examine risk. We find that 401(k) plans have slightly less risk than randomly selecting funds. However, if a plan sponsor used a common-sense rule of insisting that the plan include at least one stock and one bond fund, then plan risk from random selection would be smaller than the actual risk of the 401(k) plans. Although the individual funds selected by 401(k)

¹⁹ This is the difference in terminal wealth between compounding for 20 years with the improved return on the portfolio that spans and the return on the actual portfolio.

plans have lower variance of monthly return than randomly selected funds, the correlation between them is higher.

We then examine plan characteristics to see if they can add insight into the adequacy of plan investment choices. There is a strong correlation between the number of investment choices a plan offers and size. This is a strong indication that participants in larger plans are better off than participants in smaller plans. In addition, larger plans are more likely to use outside consultants and to include more sophisticated strategies in the plan. This raises the question of whether the use of consultants or sophisticated strategies improves results for investors. We find that, controlling for plan size, the use of outside consultants or sophisticated investment strategies increases with the number of investment choices and increases the probability of spanning. However, none of these increases is statistically significant. In addition, plans that employ consultants offer funds with lower differential returns, while plans that use sophisticated strategies offer funds with higher differential returns. Thus we have at best weak evidence that the use of consultants or sophisticated strategies leads to better results.

Finally, we examine the effect of offering company stock as an investment choice. We find that plans that offer company stock on average provide the same number of mutual fund choices as plans that do not offer company stock. The inclusion of company stock in a plan increases the variance of the plan and also leads to a slight increase in the Sharpe ratio. Including company stock results in no increase in the number of plans that span the relevant space. The overall evidence is that including company stock does not have a major positive or negative effect on the desirability of the investment choices offered to participants.

Appendix A

In the body of this paper we employ eight RB indexes to represent the appropriate classification of funds (influences) that an investor might wish to diversity across. Another approach would be to use publicly accepted classification of mutual funds to represent influences that an investor might be interested in. For this purpose we employed the 14 ICDI classifications shown in Table 2. Monthly return indexes were computed as an equally weighted monthly return index on the mutual funds that existed in any group over the period January 1992 to December 2001.

Two tests were performed to determine whether the research-based indexes (RB) or the industry practice indexes are better alternative choices for benchmarks against which to judge adequacy of fund offerings. These were individual Sharpe ratios for optimal portfolios and intersection tests as described in the text.

Sharpe Ratios

Sharpe ratios for optimal portfolios when short sales are not allowed are constructed from the RB indexes and from the ICDI indexes. In each five-year subperiod of the 1992-2001 period, the Sharpe ratio for the optimal portfolio was larger when RB indexes were used than when the larger set of ICDI-based indexes was used. Furthermore, the difference in Sharpe ratios was significantly different.²⁰

Intersection Tests

When we use intersection tests, we find that when short sales are not allowed, adding ICDI indexes to the RB indexes does not shift the efficient frontier, while adding RB indexes to the ICDI indexes shifts the efficient frontier, and the shift is statistically significant.

²⁰ We thank Andrew Lo for developing, at our request, tests for the difference in Sharpe ratios.

Based on both the Sharpe tests and the intersection tests, the eight RB indexes provide a better set of alternatives than the 14 ICDI indexes. Hence we use them in the body of the paper.²¹

²¹ We also examined whether disaggregation of the RB indexes into a general bond index and a separate mortgage and separate mid-cap and small stock index shifted the efficient frontier. It didn't. We also employed factor analysis and found the number of factors matched the smaller division and factors could be rotated so that they were highly correlated with the smaller set. Finally, we performed cluster analysis and found the larger set quickly combined to our eight indexes.

References

Ameriks, John and Stephen P. Zeldes, 2001, How do household portfolio shares vary with age? Unpublished manuscript, Columbia University.

Agnew, Julie and Pierluigi Balduzzi, 2003, What do we do with our pension money? Recent evidence from 401(k) plans. Unpublished manuscript, Boston College.

Bekaert, Geert and Michael S. Urias, 1996, Diversification, integration, and emerging market closed-end funds. *Journal of Finance* 51, 835-870.

Benartzi, Shlomo and Richard Thaler, 2001, Naïve diversification strategies in retirement saving plans. *American Economic Review* 91(1), 78-98.

Blake, Christopher R., Edwin J. Elton and Martin J. Gruber, 1993, The performance of bond mutual funds. *Journal of Business*.

Choi, James David Laibson, Brigitte Madrian and Andrew Metrick, 2004, Active decisions: a natural experiment in savings. Harvard University mimeo.

De Roon, Frans, Theo E. Nijman and Bas J. M. Werker, 2001, Testing for mean variance spanning with short sales constraints and transaction costs: the case of emerging markets. *Journal of Finance* 56(2), 721-742.

De Santis, Giorgio, 1994, Asset pricing and portfolio diversification: Evidence from emerging financial markets. Working paper, University of Southern California.

Elton, Edwin J. and Martin J. Gruber, 1970, Homogeneous groups and the testing of economic hypothesis. *Journal of Financial and Quantitative Analysis*.

Elton, Edwin J. and Martin J. Gruber, 1977, Risk reduction and portfolio size: an analytical solution. *Journal of Business* 50, 415-434.

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake, 1996, The Persistence of Risk-Adjusted Mutual Fund performance. *Journal of Business* 69(2), 133-157.

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake, 1999, Common factors in fund returns. *European Financial Review* 3(1), 1-23.

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake, 2001, A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases. *Journal of Finance* 56(6).

Elton, Edwin J., Martin J. Gruber and Jeffrey A. Busse, 2003, Are investors rational? Choices among index funds. *Journal of Finance* 59, 261-288.

Elton, Edwin J., Martin J. Gruber and T. Clifton Green, 2004, The impact of mutual fund family membership on investor risk. Unpublished manuscript, New York University.

Fama, Eugene and Ken French, 1994, Book-to-market in earnings and returns, *Journal of Finance*.

Gallant, A. Ronald, 1987, *Nonlinear Statistical Models*, John Wiley and Sons, New York.

Harman, Harry, 1976, Modern factor analysis. University of Chicago Press, Chicago, IL.

Huberman, Gur and Shmuel Kandel, 1987, Mean-variance spanning, *Journal of Finance* 42, 873-888.

Huberman, Gur and Sengmuller, 2003, Company stock in 401(k) plans. Unpublished manuscript, Columbia University.

Investment Company Institute, 2000, 401(k) plan participants, characteristics contributions, and accounts activity (Spring 2000).

Liang, Nellie and Scott Weisbenner, 2002, Investor behavior and the purchase of company stock in 401(k) plan design. Unpublished manuscript, University of Illinois.

Lo, Andrew, 2002, The statistics of Sharpe ratios, *Financial Analysts Journal* 58, 36-50.

Madrian, Brigitte and Dennis F. Shea, 2001, The power of suggestion: inertia in 401(k) participation and savings behavior, *Quarterly Journal of Economics* 116, 1149-1187.

Newey, W. and K. West, 1987, A simple positive definite heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-705.

Roll, Richard, 1977. A critique of the asset pricing theory's tests, *Journal of Financial Economics* 4, 129-176.

Table 1

Percentages of 680 401K Plans Offering Different Numbers of Investment Choices
(Number of choices and percentages include mutual funds, stable value funds, GICs and
company stock.)

Number of Investment Choices	Percentage of Plans
1	2.21%
2	2.35%
3	3.09%
4	4.85%
5	8.97%
6	12.06%
7	12.06%
8	13.82%
9	11.76%
10	9.85%
11	5.59%
12	2.21%
13	2.50%
14	1.91%
15	1.18%
16	1.03%
17 or more	4.56%