FINANCIAL SERVICES REVIEW

Financial Services Review 14 (2005) 269–284

The efficacy of optimization modeling as a retirement strategy in the presence of estimation error

Christine A. McClatchey[a,*], Shawn P. VandenHul[b]

[a]*Department of Finance, University of Northern Colorado, Monfort College of Business, Greeley, CO 80639, USA*
[b]*Equitec, Westminster, CO 80234, USA*

Abstract

We examine the time series performance of mean variance efficient portfolios in the retirement setting. Using a rolling period optimization model we create portfolios with the same ex ante risk as several naïve 1/n strategies to discern whether optimization can improve return performance. Data are simulated from TIAA-CREF retirement accounts during 1994 through 2004. We correct for estimation error using weight constraints and James-Stein adjustments. Overall results indicate optimization does outperform most naïve investment strategies. The investor's terminal wealth improves 2–30%, depending on the underlying asset allocation and assumed time to retirement. Adjustments for estimation error do little to further enhance investment returns. © 2005 Academy of Financial Services. All rights reserved.

JEL classifications: G11; G23; D14

Keywords: Retirement portfolio; Asset allocation; Mean variance optimization; Estimation error

1. Introduction

The recent growth in defined contribution plans that give individuals some responsibility to make asset allocation decisions has brought retirement welfare concerns to the forefront. Six in 10 workers expect employer-sponsored retirement plans to provide one-half or more of their total retirement income in the future, making it the single largest source of retirement funding. Although defined contribution plans provide flexibility to make investment choices

* Corresponding author. Tel.: +1-970-351-1248; fax: +1-970-351-1097.
E-mail address: christine.mcclatchey@unco.edu (C.A. McClatchey).

consistent with each individual's risk preferences and circumstances, the potential for large numbers of retirees with insufficient retirement wealth raises concerns about how well they will perform this task. We examine one aspect of this task, namely, asset allocation.

Regrettably, anecdotal evidence suggests a number of individuals are not prepared to make long-lasting, important financial decisions with respect to retirement investing. Of particular concern is widespread evidence of naïve decision making. Consider the following observations:

- 26% of survey respondents invested their entire IRA holdings in cash; 36% of survey respondents invested their entire IRA holdings in equity; no equity was held in 34% of respondents' accounts (Waggle & Englis, 2000).
- 60% of 401(k) participants never rebalance their accounts after making their initial allocation upon enrollment (DC Plan Investing, 2003).
- Almost one in five participants cannot recall how many investment options they allocated premiums to within their retirement plan (Business Wire, 2003).
- About one-half of participants' investment decisions seem to be advice-based or research-based; of the remainder, 14% picked investments with the best performance at the time, 11% listened to their co-workers or friends' suggestions, 9% simply divided money among a number of options, and 8% took their best guess (Business Wire, 2003).
- Nine of 10 employees think company stock is at least as safe as investment in a diversified portfolio (DC Plan Investing, 2003).
- Roughly one-third of the assets in large retirement savings plans are invested in company stock (Benartzi, 2001).

Collectively these findings suggest many individuals (1) lack diversification across plan assets, (2) fail to monitor or adjust portfolio allocations subsequent to initial decisions, (3) lack reasonable decision-making strategies and personal plan knowledge, and (4) misunderstand the potential consequences of holding company stock in a retirement plan. What is the potential impact for tomorrow's retirees? VanDerhei and Copeland (2003) estimate American retirees will be short $45 billion in retirement income in 2030; the aggregate deficit in retiree income during the decade ending in 2030 will be at least $400 billion. Of course, insufficient or limited savings is a significant contributor to the projected shortfall; however, a general lack of financial sophistication, coupled with the increased responsibility for making retirement decisions, further exacerbates the problem.

Although many individuals employ naïve asset allocation strategies, portfolio theory suggests a more rigorous approach should be taken. This study empirically examines the time series performance of mean variance efficient portfolios relative to several variants of a naïve 1/n investment strategy for participants in the Teachers Insurance and Annuity Association-College Retirement Equities Fund (TIAA-CREF) retirement system. Using simulated data, we employ an active, rolling period optimization strategy whereby individuals rebalance their portfolio monthly to capture changes in recent economic conditions. Ex ante risk is set equal to that of a designated naïve strategy; the optimization model then attempts to enhance portfolio return. Two techniques are used to correct for estimation error in the input parameters: (1) constraints on portfolio weights, and (2) James-Stein estimation. In all cases, optimization is based strictly on historical information that would have been available at the time each portfolio adjustment was made. Although alternative optimization models are available, mean variance methodology is the most frequently used method for selecting investment portfolios (Tew, Reid & Witt, 1991; Thorley, 1995).

2. Mean variance optimization and estimation error

Markowitz (1952) provides the basic framework for analyzing the risk and return relationships in a portfolio of assets. By selecting assets that are less than perfectly positively correlated, investors can maintain an overall rate of return (or level of risk) while lowering the level of risk (or increasing the level of return). Mean variance optimization assumes the decision-maker maximizes expected utility and either the decision-maker's utility function is quadratic or the probability distribution of returns is normal. Under these assumptions, the only pertinent moments are the mean and variance. Optimization requires knowledge of each security's expected return, standard deviation, and covariance with other security returns, which are used as inputs in a risk minimization construct to determine the optimal portfolio weightings for various risk-return combinations, as illustrated in Eq. (1). The model iteratively solves for fund weights (w_i) that produce the highest return for any specified level of risk or the lowest risk for any specified level of return. The constraints ensure that the portfolio is fully invested among the available funds and prohibit short selling, which is not allowed in the context of portfolio allocation among retirement accounts.

$$\text{Maximize } k_p = \sum_{i=1}^{n} w_i k_i,$$

$$\text{given } \sigma_p = \left[\sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} w_i w_j \sigma_{ij} \right]^{\frac{1}{2}}, \tag{1}$$

$$\text{subject to } 0 \leq w_i \leq 1 \text{ such that } \sum_{i=1}^{n} w_i = 1,$$

where σ_i = standard deviation of asset i,
σ_{ij} = covariance of assets i and j,
k_i = expected return of asset i,
w_i = weight invested in asset i,
n = number of assets in the portfolio.

The problem with mean variance optimization lies with measurement of the model's inputs. In practice, inputs are based on historical and other information and are subject to estimation error. Classic implementation of mean variance analysis replaces expected returns, variances, and covariances with their historical sample values as if they are known with certainty; optimal portfolio weights are then derived without considering uncertainty in these parameters. By construction, optimization produces larger weights for those assets with the most attractive features—high returns, low risk, or low correlation (Scherer, 2002; Michaud, 1989). Although expected variances and covariances are unknown, they are more stable over time (Merton, 1980). Furthermore, Chopra and Ziemba (1991) find that for a typical risk tolerance errors in forecasted means are about 11 times more damaging than errors in forecasted variances, and over 20 times as damaging as errors in covariances. Thus, we focus our analysis on errors in means only.

The impact of estimation error on optimization modeling is threefold. First, optimized portfolios oftentimes exhibit sharp deterioration in performance outside of the estimation period. Assets with high historical returns are more likely to have positive estimation errors, inducing practitioners of mean variance optimization to hold too much of them. In the subsequent period, the mean variance efficient portfolio is likely to perform poorly because the realized returns on these assets are likely to be lower than previously predicted and too much of those assets were held (Chopra, Hensel & Turner, 1992). Second, optimized portfolios may be highly sensitive to slight changes in expected returns. Even small changes in the estimation period can dramatically alter the composition of the optimal portfolio. Finally, optimization oftentimes produces corner solutions whereby the optimized portfolio is heavily weighted in a single security. The end result is a severely under-diversified portfolio. Collectively, optimized portfolios that are constructed purely from historical data (unadjusted for estimation error) can be financially irrelevant and lead to non-intuitive decisions.

Estimation error is inevitable as the historical sample mean invariably differs from the true underlying mean. While using more observations to compute the sample mean would move it closer to the population mean, the distribution of returns may be time varying or non-stationary. Fortunately, several techniques that reduce the input-sensitivity problem of mean variance optimization are available. The first involves the use of reasonable constraints on portfolio weights. Constraints prevent the model from inappropriately magnifying the influence of forecast errors by limiting the optimizer's allocation to any single asset (Chopra et al., 1992). Cohen and Pogue (1967) impose upper bounds of 5% and 2.5% on security populations consisting of 75 and 150 securities. Frost and Savarino (1988) examine constraints of 5%, 2%, 1%, 0.67%, and 0.50% on a security population of 200 securities. Chopra (1991) tests this approach for general asset classes, rather than individual securities, allowing investment in each asset class to vary by 10%, 20%, and 30% from the common 60-40-0 benchmark often used by large pension funds (60% in equities, 40% in bonds, 0% in cash). Although constraints on weights reduce the impact of errors, they may simultaneously reduce the potential gains from information that is not currently reflected in market prices.

A second approach is James-Stein estimation. The basic idea underlying this technique is that similar assets should, on average, behave in a similar manner over the long run. Thus, if assets can be considered to belong to the same basket, their characteristics (returns, variances, and cross-correlations) are likely to be similar. Stein estimation involves shrinking observed sample means for individual assets within the same class toward some global mean for that asset class. The global mean may represent (1) the pooled mean, (2) a Bayesian prior, (3) the mean of the minimum-variance efficient frontier portfolio, or (4) a Sharpe-Litner CAPM-based estimator. Chopra et al. (1992) replace each equity (bond) mean return by the average of individual equity (bond) means, an extreme form of complete shrinkage to the global mean. Overall, shrinkage estimators reduce the tendency of the optimizer to magnify estimation errors and make large allocations to a few assets.

Although previous studies have offered alternative methods for minimizing estimation risk, they have done so in contexts quite different from the present study. Larsen and Resnick (2001) optimize over portfolios stratified by some measure providing return persistence (in their case, size-based portfolios). Jorion (1985, 1986), Solnik and Noetzliin (1982), and Fletcher and Hillier (2001) examine international asset allocation. Altay-Salih, Muradoglu and Mercan (2002) examine emerging markets. The dynamics underlying the retirement setting have practical implications for mean variance optimization that differ from other investment settings. First, the available investment opportunity set is not comprised of individual securities; rather it is comprised of a set of mutual funds. Diversification across mutual funds is different than diversification across individual securities, in that each fund itself is well diversified. Furthermore, participants are limited to the set of investment choices offered by their specific plan that may be quite varied or skewed toward a single asset class. It may be that estimation error is significantly lower for investors allocating premiums among mutual funds, rather than among individual securities, and out-of-sample performance gains cannot be significantly increased. Second, periodic portfolio rebalancing is generally allowed free of transaction costs, although the frequency may be constrained. Rebalancing outside of a retirement setting may entail significant transaction costs that could exceed any gains from an active strategy that requires rebalancing. Finally, we have no problems with indivisibilities and limited portfolio sizes. It is possible for investors to hold a portfolio containing all of the mutual funds used in our calculations. The same is not true for studies that employ market index returns as proxies for benchmark investment opportunities.

3. Data and summary statistics

Daily net asset value data were obtained from www.tiaa-cref.org from May 31, 1994 through June 30, 2004 for the following CREF retirement accounts: Bond Market, Stock, Equity Index, Growth, Global Equities, Social Choice, and Money Market.[1] The year 1994 was chosen as a starting point because it was the most distant, yet common, inception date for all seven accounts. Appendix A details each account's objective and inception date. Because Social Choice is a balanced account that invests in both fixed income and equity securities, we denote the subset Stock, Equity Index, Growth, and Global as "all-equity" accounts. Monthly returns were obtained by computing the monthly change in net asset value for each of the seven accounts. Summary statistics are provided in Table 1.

Monthly returns ranged from 0.346% (Money Market) to 0.972% (Equity Index). As expected, the Money Market account exhibited the lowest volatility, followed by Bond (monthly $\sigma_{\text{Bond}} = 1.14\%$). As a balanced account, the volatility of Social Choice was less than the four all-equity accounts but greater than Bond and Money Market. The correlation coefficients between the all-equity accounts ranged from +0.91 to +0.99. The strong relationship between Stock and Equity Index (+0.99) is not surprising since a significant component of the Stock account is indexed. The Bond account was negatively correlated

Table 1
Summary statistics for TIAA-CREF retirement accounts 5/31/94–6/30/04

	Stock	Global	Growth	Equity	Social	Bond	Money
Monthly return, standard deviation, and coefficient of variation							
Return (%)	0.866	0.672	0.830	0.972	0.840	0.571	0.346
SD (%)	4.31	4.44	5.77	4.48	2.71	1.14	0.163
CV	4.98	6.61	6.95	4.61	3.23	2.00	0.471
Correlation coefficients							
Stock	1.000						
Global	0.947	1.000					
Growth	0.940	0.911	1.000				
Equity	0.994	0.920	0.946	1.000			
Social	0.958	0.853	0.893	0.967	1.000		
Bond	−0.029	−0.086	−0.053	−0.018	0.178	1.00	
Money	0.051	0.018	0.074	0.073	0.105	0.186	1.00

with all four all-equity accounts and exhibited a positive but weak correlation with Social Choice and Money Market.

4. Methodology

Historical data are used to simulate 180 monthly returns for each CREF account.[2] We employ a 60-month rolling estimation period; for each month, the prior 60 months of return data are used to (1) construct the variance-covariance matrix, (2) compute historical mean returns, and (3) compute the variance of a designated naïve portfolio. The optimization process employs this data to construct the efficient frontier and to identify account weights that maximize return for a level of risk equal to the naïve portfolio's risk. The active strategy is to invest according to these weights for the next month, ignoring any transaction costs. Using overlapping periods, the process is repeated for each consecutive month, using optimized investment weights from the most recent 60-month period, leaving 10 years (120 months) to measure out-of-sample return performance. New data are then simulated, and the entire process is repeated 1,000 times.

We use the sample covariance matrix for the prior 60 months as the estimate for the expected covariance matrix. Three models of expected returns are examined.

- Certainty Equivalence Method (HIST): HIST assumes there is no estimation risk in classical sample estimates. Historical mean returns over the most recent 60-month period are used.
- Constrained Certainty Equivalence Method (CONST): CONST also uses historical means; however, feasible investment weights are constrained to 50% in any single account. As a result, the optimizer is forced to allocate across a *minimum* of two accounts.
- James-Stein Method (JSM): Define the n-vector of historic means at the beginning of period t as:

$$\mu_{Ht} = (r_{it}, \ldots, r_{nt})', \tag{2}$$

where: $\bar{r}_{it} = \frac{1}{k} \sum_{r=\tau-k}^{t-1} r_{i\tau},$

k = the number of periods in the estimation period,

n = the number of funds.

The James-Stein estimator shrinks historic mean returns toward their arithmetic average, taking the form:

$$\mu_{JSt} = (1 - w_t)\mu_{Ht} + w_t\bar{r}_{Gt}e, \tag{3}$$

where:

$$w_t = \min\left[1, \frac{(n-2)/k}{(\mu_{Ht} - \bar{r}_{Gt}e)'S_t^{-1}(\mu_{Ht} - \bar{r}_{Gt}e)}\right], \tag{4}$$

S is the sample covariance matrix calculated from the k-period estimation period,

$\bar{r}_{Gt} = \frac{1}{n} \sum_i \bar{r}_{it}$ is the grand mean,

e is a n-vector of 1's.

As a practical matter, Stein estimators should be used for assets that belong to the same asset class—it would be inappropriate to apply the adjustment to financial instruments with very different characteristics, e.g., stocks and fixed income securities. To this end, we employ historical mean return data for the Money Market, Bond, and Social Choice accounts and Stein estimators for the four all-equity accounts (Stock, Equity Index, Global, and Growth).

5. Benchmark portfolios

To make meaningful comparisons, the mean variance efficient portfolio that has the same variance as each benchmark portfolio in the most recent 60-month estimation period is chosen. We examine several naïve portfolio allocations as our benchmarks, as described in Table 2. Each of the naïve strategies applies some variant of the commonly employed 1/n heuristic, as documented by Benartzi and Thaler (2001).

The first portfolio, denoted Strategy 1, invests equally across all of the accounts except for Money Market. (This account is generally considered a tool appropriate for short-term money management rather than long-term retirement investing; hence naïve investors are likely to exclude it from their allocation decision.) Bernartzi and Thaler (2001) find some investors divide their contributions evenly across all funds offered in the plan, irrespective of the mix of funds offered by the plan; as the number of available stock (fixed income) funds increases, so does the allocation to equities (fixed income).

Although robust across several model specifications, the 1/n heuristic documented by

Table 2
Asset allocation of 1/n benchmark portfolios

1/n Strategy	Stock	Equity Index	Growth	Global	Social	Bond
S1	1/6	1/6	1/6	1/6	1/6	1/6
S2	1/2					1/2
S3		1/2				1/2
S4	1/8	1/8	1/8	1/8		1/2
S5	1/4		1/4			1/2
S6		1/4	1/4			1/2
S7	1/4	1/4	1/4	1/4		

Bernartzi and Thaler (2001) is not supported by other researchers who find that individuals may use the 1/n approach across broad asset categories, rather than individual funds. That is, many individuals hold portfolios allocated 50% in equities and 50% in fixed income, regardless of the types and mix of funds offered. The simplest strategy would be to invest the entire equity position in a single fund. In fact, TIAA-CREF suggested precisely this strategy in a recent article distributed to all of its participants (TIAA-CREF, 2004). Of the available equity accounts, Stock and Equity Index are managed as broad diversified equity portfolios and represent the most likely candidates. To this end, we examine Strategy 2 (invested one-half in Stock and one-half in Bond) and Strategy 3 (invested one-half in Equity Index and one-half in Bond).

Research by Goodfellow and Schieber (1997) and Yakoboski and VanDerhei (1996) suggests the allocation within the equity position may be more complex; that is, investors may spread their equity position equally across several accounts. To replicate this position, we examine Strategy 4, which invests one-eighth in each of the four all-equity accounts and the remaining one-half in Bond. The choice of investing equally across all of the equity accounts contradicts some research that finds allocations to international funds are zero for a substantial number of investors (Goodfellow & Schieber, 1997). Furthermore, Stock and Equity Index are both well-diversified equity accounts; although Equity Index is managed to track the Russell 3000, the largest segment of the Stock account is promoted with the same directive. To the extent investors perceive these accounts as substitutes for one another, they may choose to include only one in their portfolio. To this end, we examine Strategy 5 (one-half in Bond, one-quarter in Stock, one-quarter in Growth), and Strategy 6 (one-half in Bond, one-quarter in Equity Index, one-quarter in Growth).

The final naïve strategy applies the 1/n heuristic to the four all-equity accounts. Agnew, Balduzzi and Sunden (2000) and VanDerhei, Galer, Quick and Rea (1999), among others, find a fair proportion of investors holding a 100% equity position, albeit this tendency seems to decline with age (Goodfellow & Schieber, 1997). Given the available CREF investments, we examine Strategy 7, which invests one-quarter in Equity Index, one-quarter in Growth, one-quarter in Stock, and one-quarter in Global.

For each strategy we allow the optimized portfolio to invest in any combination of the seven CREF accounts (Money Market is allowed in the optimized portfolio). We compare the out-of-sample realized return, standard deviation, and terminal wealth of the optimized portfolios with their seven equal-risk naïve strategies over a 10-year investment horizon.[3]

6. Results

We first examine the out-of-sample performance of mean variance optimization before any corrections for estimation error (HIST vs. 1/N). Results are reported in Table 3. The first four columns provide monthly percentages (average return, standard deviation, minimum return, and maximum return), averaged across the 1,000 simulations. Column 5 annualizes monthly return data. The ex-post standard deviation of the optimized and passive strategies are very similar within each of the seven comparisons; recall the optimization process sets the ex ante, or estimation period, variances equal. As expected, the portfolios obtained via Strategy 7 bear the highest risk and highest returns, followed by Strategy 1. Strategy 7 invests one-quarter in each of the all-equity accounts; thus, it is 100% invested in equities. Strategy 1 invests equally across six of the CREF accounts; four are all-equity accounts, while Social Choice is a balanced account, creating a heavily laden equity allocation.

Overall performance comparisons indicate mean variance optimization does generally enhance return performance relative to a naïve strategy; the optimized portfolio earned a higher average monthly return in six of the seven strategies examined. For ease of comparison, we subtract each naïve portfolio's annual return from its corresponding optimized annual return. The differences range from −0.199% (Strategy 3) to +0.835% (Strategy 5). Optimization improves return by 75 basis points or more in over one-half of the comparisons, consistent with the results of Canner, Mankiw and Weil (1997).

Similar to Larsen and Resnick (2001), we perform a dominance analysis on the ex-post average returns to measure each strategy's consistency. Superior performance due to a few high values may not be preferred to a strategy that provides enhanced performance at a lower level a majority of the time. To this end, we track the number of times (out of 1,000 simulations) the average return of the optimized strategy exceeds the average return of the passive strategy. For example, the results from Strategy 1 suggest mean variance optimization (on average) improves return by 79.6 basis points; furthermore, optimization outperformed the naïve strategy (earned a higher average monthly return) in 83% of the simulations. Other results are similar, with the exception of Strategy 3; that is, optimization generally outperformed the naïve strategy the majority of the time.

We do not conduct statistical significance tests, because of their poor performance. As noted by Chopra et al. (1992), statistical tests for return differences are weak for monthly data because the magnitude of the variance of returns is much greater than the mean. The out-of-sample time horizon would have to be quite long (e.g., beyond human life expectancy) to detect statistical significance. As an alternative to *statistical* significance, we examine the *economic* significance of our results. The final column in Table 3 reports the terminal wealth of each strategy, assuming a $1 investment is made each month over a 10-year investment horizon. All computations assume the portfolio is rebalanced monthly (existing dollars in the optimized and the naïve portfolios are reset each month). Consistent with return results, we find terminal wealth is enhanced by optimization in all cases except Strategy 3. The largest

Table 3
Average monthly return, standard deviation, and minimum/maximum returns of mean variance optimized and 1/n investment strategies during 10-year out-of-sample period

Passive 1/n benchmark strategy	Allocation strategy	Monthly return %	Monthly standard deviation %	Minimum monthly return %	Maximum monthly return %	Average annual return %	Annualized difference from 1/n %	Percent of simulations optimized return exceeded 1/n return	Terminal wealth
S1: 1/6 Stock, 1/6 Equity, 1/6 Growth, 1/6 Global, 1/6 Social, 1/6 Bond	1/N	0.788	3.50	−15.22	15.54	9.455			$200.40
	HIST	0.854	3.50	−14.78	16.18	10.251	0.796	83%	$210.22
S2: 1/2 Stock, 1/2 Bond	1/N	0.716	2.20	−8.60	10.45	8.595			$190.66
	HIST	0.749	2.21	−8.66	10.46	8.984	0.389	77%	$195.01
S3: 1/2 Equity, 1/2 Bond	1/N	0.776	2.28	−8.59	11.10	9.309			$198.73
	HIST	0.759	2.29	−9.02	11.00	9.110	−0.199	33%	$196.51
S4: 1/8 Stock, 1/8 Equity, 1/8 Growth, 1/8 Global, 1/2 Bond	1/N	0.701	2.35	−9.97	11.10	8.413			$188.53
	HIST	0.767	2.36	−9.23	11.05	9.202	0.788	85%	$197.37
S5: 1/4 Stock, 1/4 Growth, 1/2 Bond	1/N	0.712	2.50	−10.71	11.85	8.548			$190.07
	HIST	0.782	2.51	−9.95	11.73	9.383	0.835	85%	$199.72
S6: 1/4 Equity, 1/4 Growth, 1/2 Bond	1/N	0.740	2.55	−10.69	12.17	8.886			$193.67
	HIST	0.782	2.56	−10.22	11.90	9.384	0.498	73%	$199.62
S7: 1/4 Stock, 1/4 Equity, 1/4 Growth, 1/4 Global	1/N	0.829	4.61	−20.44	20.72	9.951			$206.68
	HIST	0.897	4.60	−19.79	21.29	10.770	0.819	85%	$217.06

The optimized portfolio has the same ex-ante variance as the 1/n strategy. The annualized difference is equal to the average annual return of each optimized strategy minus the average annual return of its corresponding 1/n strategy. Terminal wealth is the accumulated value of the portfolio assuming a $1 investment each month over a 10-year out-of-sample period. Return, risk, and terminal wealth measures are averages across 1,000 simulations.

improvement is obtained for Strategy 7—the optimized portfolio's terminal wealth is $10.38 higher than its naïve counterpart, a 5% improvement with no additional risk. Expanding the contribution period—a reasonable assumption for the retirement setting—produces even more dramatic results. If we assume monthly contributions of $1,000 for 20 years, we find a mean variance investor matching Strategy 7 would have about $91,000 more at retirement (an 11.8% improvement). After 30 years, the mean variance investor would be $470,000 wealthier (a 20% improvement) and, after 40 years, $1.9 million wealthier (a 30% improvement). Results for other strategies are similar; optimization improves terminal wealth by 2–30%, depending on the underlying asset allocation and assumed time to retirement.

We next examine the sensitivity of our results to estimation error using the two methods described previously: (1) constraining portfolio weights (CONST), and (2) incorporating statistical corrections in the mean inputs via James-Stein estimation (JSM). Table 4 reports annualized out-of-sample performance results similar to the structure in Table 3. The James-Stein estimator (JSM) produced a higher average return than the HIST estimator (unadjusted historical inputs) and weight constraint (CONST) methods in five of the seven strategies, suggesting such adjustments may enhance the performance of mean variance optimization. However, the magnitude of enhanced return performance, whether measured by the increase in monthly return or increase in terminal wealth, is quite small, suggesting the improvement is insignificant in economic terms. Furthermore, the dominance analysis finds such corrections do not enhance consistency.

7. Summary and conclusions

We employ a rolling period mean variance optimization model to data simulated from historical TIAA-CREF retirement accounts with the goal of evaluating its effectiveness as a return-enhancing strategy for retirement investors. We examine the retirement setting for two reasons. First, it is unique from other investment situations, which may limit the generalization of previous research findings. Specifically, retirement investors generally choose from a pre-specified set of investment funds, rather than individual securities, and are allowed periodic transfers among funds without incurring transaction costs. Second, the importance of understanding the behavior of retirement investors has escalated with an increase in the number of defined contribution plans and the current state of our social security system. Defined contribution plans that assign critical decisions (such as participation, contributions, and asset allocation) to the individual will likely fund a significant portion of future retirement income. Evidence to date suggests many individuals are not making sensible, value-maximizing decisions (Brennan & Torous, 1999). From a public policy perspective, it is important to investigate all facets of the decision-making process to ensure tomorrow's retirees have sufficient wealth.

Overall, we find ex ante optimization is generally able to outperform an equal-risk naïve strategy. Depending on the investor's allocation strategy and time to retirement (contribution horizon), we find mean variance optimization can improve terminal wealth by 2–30% with no additional risk, despite the strong historical correlation among several of the CREF accounts. Our results are consistent with other studies that have successfully employed mean

Table 4
Summary out-of-sample performance comparisons

		Average annual return %	Annualized difference % from 1/n	Percent of simulations optimized return exceeded 1/n return	Terminal wealth
S1: 1/6 Stock, 1/6 Equity, 1/6 Growth, 1/6 Global, 1/6 Social, 1/6 Bond	HIST	10.251	0.796	83%	$210.22
	CONST	10.244	0.789	85%	$209.79
	JSM	10.185	0.730	82%	$209.01
S2: 1/2 Stock, 1/2 Bond	HIST	8.984	0.389	77%	$195.01
	CONST	8.991	0.396	77%	$195.04
	JSM	8.996	0.401	78%	$195.05
S3: 1/2 Equity, 1/2 Bond	HIST	9.110	−0.199	33%	$196.51
	CONST	9.102	−0.207	33%	$196.21
	JSM	9.116	−0.193	34%	$196.48
S4: 1/8 Stock, 1/8 Equity, 1/8 Growth, 1/8 Global, 1/2 Bond	HIST	9.202	0.788	85%	$197.37
	CONST	9.163	0.750	84%	$196.94
	JSM	9.223	0.810	85%	$197.71
S5: 1/4 Stock, 1/4 Growth, 1/2 Bond	HIST	9.383	0.835	85%	$199.72
	CONST	9.276	0.728	84%	$198.49
	JSM	9.416	0.868	84%	$200.03
S6: 1/4 Equity, 1/4 Growth, 1/2 Bond	HIST	9.384	0.498	73%	$199.62
	CONST	9.364	0.478	73%	$199.33
	JSM	9.392	0.506	74%	$199.72
S7: 1/4 Stock, 1/4 Equity, 1/4 Growth, 1/4 Global	HIST	10.770	0.819	85%	$217.06
	CONST	10.674	0.723	87%	$215.66
	JSM	10.604	0.653	54%	$214.57

Returns and terminal wealth are averages across 1,000 simulations. The optimized portfolio has the same ex-ante variance as the 1/n strategy. The annualized difference is equal to the average annual return of each optimized strategy minus the average annual return of its corresponding 1/n strategy. Terminal wealth is the accumulated value of the portfolio assuming a $1 investment each month over a 10-year out-of-sample period.

variance optimization in other investment settings (Brennan & Torous, 1999; Larsen & Resnick, 2001; Jorion, 1985; Levy & Lerman, 1988); however, they do contradict some empirical evidence that finds correcting for estimation error, particularly in the means, can improve investment performance substantially (Jobson, Korkie & Ratti, 1979; Jobson & Korkie, 1980, 1981; Jorion, 1985, 1991). That is, we find little or no improvement in terminal wealth using weight constraints or James-Stein estimators when compared to using historical sample estimates during optimization. This disparity may suggest that estimation error is already tempered when investors allocate funds across diversified portfolios, rather than individual securities, making further adjustments unnecessary.

We make our overall recommendations with three important caveats in mind. First, our results should not be generalized to all retirement investors or to all retirement settings. The return enhancing potential of mean variance optimization depends on the specific mix of available funds and their cross-correlation with one another. Second, our results do not suggest equal-risk optimized strategies are *appropriate* strategies. There are two potential allocation errors that can result in welfare losses: (1) choosing an inefficient portfolio that plots below the frontier and (2) choosing an inappropriate point along the frontier. We examine only the first—an optimized strategy that is *efficient* is not necessarily *appropriate* for a given individual's risk-return preferences. Finally, our analysis did not examine the effect of company stock as an available investment option. Many individuals make a variety of problematic decisions with respect to company stock—oftentimes investing a significant portion of their wealth in a single security.[4] Portfolio theory suggests employer stock would consistently plot inside of the efficient frontier, as its returns contain significant diversifiable risk. Most likely a naïve allocation strategy that includes employer stock as a significant component would be consistently (and economically) dominated by mean variance optimization.

Notes

1. We exclude Real Estate, Inflation-Linked Bond, and TIAA Traditional Annuity. Real Estate and Inflation-Linked Bond were not introduced until October 2, 1995 and May 1, 1997, respectively. Including them necessitates the use of a shorter time period for all funds (1997–2004); because this period includes an extended bear market, the model produces non-intuitive results (e.g., fixed income returns that are greater than equity returns). TIAA Traditional Annuity guarantees a minimum 3% return, with the opportunity to earn higher rates if the underlying portfolio performs well. The separation theorem underlying the Markowitz framework and the resulting Capital Asset Pricing Model finds the decision to select the risky portfolio of stocks is distinct and separate from the borrowing or lending decision. That is, mean variance optimization should be applied only to the portfolio of risky assets. Once it is selected, investors can borrow or lend (in this case investors would be limited to lending via the Traditional Annuity account) to achieve their desired risk-return position.
2. Data were simulated to follow a multivariate normal distribution.
3. Weekly portfolio revisions rule out traditional measures of portfolio performance such as Sharpe, Jensen, or Treynor measures (Klemkosky & Bharati, 1995). Terminal wealth computations assume portfolios are re-balanced monthly; for the naïve investor, this implies that both new and existing dollars are invested according to that strategy's 1/n allocation.
4. Brennan and Torous (1999) and Benartzi (2001) find that roughly a third of the assets in large retirement savings plans are invested in company stock. In extreme cases, such as Coca-Cola, the allocation reaches 90%. When the employer's contributions are automatically directed to company stock, employees invest more of their own contributions in company stock (Benartzi, 2001). Nearly one-half of participants with a choice of company stock have more than one-fifth of their plan assets allocated to

company stock; 16% have at least 80% of their retirement assets invested in company stock (DC Plan Investing, 2003).

Acknowledgments

We would like to thank James Larsen and other session participants at the 2004 Academy of Financial Services Annual Meeting, the Editor, and an anonymous reviewer for their helpful comments in subsequent versions of the manuscript. Also, we would like to thank Michelle Ermentraut, TIAA-CREF Individual Consultant, for her diligence in gathering the data necessary for our empirical analyses.

Appendix: CREF variable annuity account descriptions and inception dates*

CREF Stock

This widely diversified account seeks long-term returns through capital appreciation and current income while avoiding extremes of conservatism and high risk. It invests primarily in U.S. stocks with approximately 13% in foreign securities. Inception date: July 1, 1952.

CREF Global Equities

This actively managed fund offers participation in stock markets around the world, including the United States, for diversification and growth potential. The account's returns should reflect returns across many foreign securities markets and may be significantly higher or lower than those of U.S. markets at any time. Inception date: May 1, 1992.

CREF Growth

This actively managed account looks for favorable long-term returns, mainly through capital appreciation, from a portfolio of stocks believed to be poised for superior growth in light of economic and market conditions. Inception date: April 29, 1994.

CREF Equity Index

This highly diversified variable annuity account is designed to track the overall U.S. stock market as represented by the Russell 3000. Because of the account's comprehensive holdings, with individual issues purchased in proportion to their overall market value, it should generally parallel overall market performance. Inception date: April 29, 1994.

CREF Bond Market

This account seeks favorable long-term returns, mainly through high current income consistent with capital appreciation. It holds bonds of many different companies and government agencies, all with varying maturities; its returns depend on interest income and price changes in the bond market, themselves ordinarily dependent on interest rate changes. Inception date: March 1, 1990.

CREF Social Choice

This account is a balanced fund, holding stocks, bonds, and money market instruments issued by companies included in the Russell 3000 Stock Index that pass two sets of social screens. The fixed-income portion, which can also invest in foreign and domestic government securities, attempts to track the Lehman Brothers Aggregate Bond Index. Inception date: March 1, 1990.

CREF Money Market

This account seeks high current income consistent with liquidity and capital preservation. Yields reflect current short-term interest rates; they should keep pace with the cost of living and may provide some real rate of return over longer periods. The account is neither insured nor guaranteed by the government. Inception date: April 1, 1988.

* Note: Account descriptions taken from www.tiaa-cref.org (9/30/03).

References

Agnew, J., Balduzzi, P., & Sunden, A. (2000). *Portfolio choice, trading, and returns in a large 401(k) plan.* Working Paper CRR WP 2000–06. Boston, MA: Boston College, Center for Retirement Research.

Altay-Salih, A., Muradoglu, G., & Mercan, M. (2002). Performance of the efficient frontier in an emerging market setting. *Applied Economic Letters, 9,* 177–183.

Benartzi, S. (2001). Excessive extrapolation and the allocation of 401(k) accounts to company stock. *Journal of Finance, 56,* 1747–1764.

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

Brennan, M. J., & Torous, W. N. (1999). Individual decision making and investor welfare. *Economic Notes by Banca Monte dei Paschi di Siena SpA, 28,* 119–143.

Business Wire. (2003). Employer-sponsored retirement plans have become the most important source of future retirement income. December 18, 2003. Lexis-Nexis Academic Universe. Available at: http://web.lexis-nexis.com/universe. Accessed January 20, 2004.

Canner, N. N., Mankiw, G., & Weil, D. N. (1997). An asset allocation puzzle. *The American Economic Review, 87,* 181–191.

Chopra, V. K. (1991). Mean-variance revisited: near-optimal portfolios and sensitivity to input variations. *Russell Research Commentaries, December,* 1–17.

Chopra, V. K., Hensel, C. R., & Turner, A. L. (1992). Massaging mean-variance inputs: returns from alternative global investment strategies in the 1980s. *Russell Research Commentaries, May,* 1–19.

Chopra, V. K., & Ziemba, W. T. (1991). The effect of errors in means, variances, and covariances on optimal portfolio choice. *Russell Research Commentaries, November,* 1–11.

Cohen, K. J., & Pogue, J. A. (1967). An empirical evaluation of alternative portfolio selection models. *Journal of Business, 40,* 166–189.

DC Plan Investing. (2003). New asset allocation strategy targets overexposure to company stock in DC plans, May 13, 2003. Lexis-Nexis Academic Universe. Available at: http://web.lexis-nexis.com/universe. Accessed January 20, 2004.

Fletcher, J., & Hillier, J. (2001). An examination of resampled portfolio efficiency. *Financial Analysts Journal, 57*, 66–74.

Frost, P. A., & Savarino, J. E. (1988). For better performance: constrain portfolio weights. *Journal of Portfolio Management, 15*, 29–34.

Goodfellow, G. P., & Schieber, S. J. (1997). Investment of assets in self-directed retirement plans. In M. S. Gordon, O. S. Mitchell, & M. M. Twinney (Eds.), *Positioning pensions for the twenty-first century. Pension Research Council* (pp. 67–90). Philadelphia, PA: University of Pennsylvania Press.

Jobson, J. D., & Korkie, B. (1980). Estimation for Markowitz efficient portfolios. *Journal of American Statistical Association, 75*, 544–554.

Jobson, J. D., & Korkie, B. (1981). Putting Markowitz theory to work. *Journal of Portfolio Management, 7*, 70–74.

Jobson, J. D., Korkie, B., & Ratti, V. (1979). Improved estimation for Markowitz portfolios using James-Stein type estimators. *Proceedings of the American Statistical Association. Business and Economics Statistics Section, 41*, 279–284.

Jorion, P. (1985). International portfolio diversification with estimation risk. *Journal of Business, 58*, 259–278.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis, 21*, 279–292.

Jorion, P. (1991). Bayesian and CAPM estimators of the means: implications for portfolio selection. *Journal of Banking and Finance, 15*, 717–727.

Klemkosky, R. C., & Bharati, R. (1995). Time varying expected returns and asset allocation. *Journal of Portfolio Management, 21*, 80–87.

Larsen, G. A., & Resnick, B. G. (2001). Parameter estimation techniques, optimization frequency, and portfolio return enhancement. *Journal of Portfolio Management, 27*, 27–34.

Levy, H., & Lerman, Z. (1988). Testing the predictive power of ex-post efficient portfolios. *Journal of Financial Research, 11*, 241–254.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Merton, R. C. (1980). On estimating the expected return on the market. *Journal of Financial Economics, 8*, 323–361.

Michaud, R. (1989). The Markowitz optimization enigma: is 'optimized' optimal? *Financial Analysts Journal, 45*, 31–42.

Scherer, B. (2002). Portfolio resampling: review and critique. *Financial Analysts Journal, 58*, 98–109.

Solnik, B., & Noetzliin, B. (1982). Optimal international asset allocation. *Journal of Portfolio Management, 9*, 11–21.

Tew, B. V., Reid, D. W., & Witt, C. A. (1991). The opportunity cost of a mean-variance efficient choice. *The Financial Review, 26*, 31–43.

Thorley, S. R. (1995). The time-diversification controversy. *Financial Analysts Journal, 51*, 68–76.

TIAA-CREF Advance. (2004). The virtue of simplicity. *Quarterly News and Tools from TIAA-CREF, Summer*, 12–15.

VanDerhei, J., Galer, R., Quick, C., & Rea, J. (1999). *401(k) plan asset allocation, account balances, and loan activity.* EBRI Issues Brief No. 205, January (Employee Benefits Research Institute).

VanDerhei, J., & Copeland, C. (2003). *Can America afford tomorrow's retirees: results from the EBRI-ERF Retirement Security Projection Model.* EBRI Issues Brief No. 263, November (Employee Benefits Research Institute).

Waggle, D., & Englis, B. (2000). Asset allocation decisions in retirement accounts: an all-or-nothing proposition? *Financial Services Review, 9*, 79–92.

Yakoboski, P., & VanDerhei, J. (1996). *Worker investment decisions: an analysis of large 401(k) plan data.* EBRI Issue Brief No. 176, August (Employee Benefits Research Institute). AQ1: Please provide a page range for this reference.