Efficient Retirement Financial Strategies

William F. Sharpe, Jason S. Scott, and John G. Watson

August 2007

PRC WP2007-19
Pension Research Council Working Paper
Pension Research Council
The Wharton School, University of Pennsylvania
3620 Locust Walk, 3000 SH-DH
Philadelphia, PA 19104-6302
Tel: 215.898.7620 Fax: 215.573.3418
Email: prc@wharton.upenn.edu
http://www.pensionresearchcouncil.org

The authors thank John Ameriks, Olivia S. Mitchell, and Stephen Zeldes for many valuable comments on an early draft of this chapter. We also wish to thank our colleagues at Financial Engines Wei Hui and Jim Shearer for their insights and support. All findings, interpretations, and conclusions of this paper represent the views of the authors and not those of the Wharton School or the Pension Research Council.

Efficient Retirement Financial Strategies

Abstract

Today's retirees face the daunting task of determining appropriate investment and spending strategies for their accumulated savings. Financial economists have addressed their problem using an expected utility framework. In contrast, many financial advisors rely instead on rules of thumb. We show that some of the popular rules are inconsistent with expected utility maximization, since they subject retirees to avoidable, non-market risk. We also highlight the importance of *earmarking*—the existence of a one-to-one correspondence between investments and future spending—and show that a natural way to implement earmarking is to create a *lockbox* strategy.

William F. Sharpe
Stanford Graduate School of Business
Stanford University
518 Memorial Way
Stanford, CA 94305-5015
wfsharpe@stanford.edu

John Watson
Financial Engines, Inc.
1804 Embarcadero Rd.
Palo Alto, CA 94303
jwatson@financialengines.com

Jason Scott
Financial Engines, Inc.
1804 Embarcadero Rd.
Palo Alto, CA 94303
jscott@financialengines.com

Efficient Retirement Financial Strategies

William F. Sharpe, Jason S. Scott, and John G. Watson

Today's retirees are making increasingly complex financial decisions. Gone are the days when one could rely solely on government or corporate pensions. The freezing or elimination of pension plans, combined with the rapid introduction of defined contribution plans, has forced retirees to rely more and more on their own investments to fund their retirement spending. Retirees are not only expected to fund a larger portion of their retirement spending, but early retirement and increased longevity imply their assets must support potentially longer retirements as well. To address this responsibility, a retiree has either implicitly or explicitly adopted an *investment strategy* to govern his investment decisions and a *spending strategy* to govern his spending decisions. A pair of investment and spending strategies constitutes a *retirement financial strategy*.

Economists have long explored the issue of optimal spending and investment strategies (Merton 1971). A major theme of their work is that optimal or *efficient* solutions are only achieved when investment and spending decisions are made in tandem as part of a complete retirement financial strategy. Economists use a standard framework in which the retiree's goal is to maximize expected utility in a *complete market*. However, solving such a utility maximization problem requires detailed knowledge about the retiree's preferences. Moreover, one must make assumptions about the trade-offs available in capital markets. Not surprisingly, financial advisors rarely embrace this approach; rather they rely on "rules of thumb". For example, one popular rule suggests annually spending a fixed, real amount equal to four percent of initial wealth and annually rebalancing the remainder to a 40 percent-60 percent mix of bonds and stocks. The goal of this chapter is to consider whether the advice suggested by financial planners is consistent with the approach advocated by financial economists. More specifically, we examine some rules of thumb to see if they are consistent with expected utility maximization, for at least some investor in a standard market setting. If a rule is consistent, we say it is efficient and refer to the underlying utility as the investor's *revealed* utility.

In what follows, we make several key assumptions—the assumptions of our *canonical setting*. Regarding retiree preferences, we assume they are well modeled by additively separable utility functions.[1] Moreover, we assume that spending preferences take into account mortality estimates and the retiree's attitudes concerning his spending relative to that of any beneficiaries, and that the amounts to be spent under the plan will go either to the retiree or to beneficiaries. Further, we assume, as do many rules advocated by financial planners, that no annuities are purchased. Our assumptions about asset prices are consistent with a condition associated with models of asset pricing such as the Capital Asset Pricing Model and a number of Pricing Kernel Models—only risk associated with the overall capital market is compensated. More specifically, we assume that there is no compensation in higher expected return from taking non-market risk (Cochrane 2005; Sharpe 2007). Further, to keep the mathematics as simple as possible, we will develop our results using a simple *complete market* consisting of a riskless asset and a risky asset. Our risky asset tracks the market portfolio, which is assumed to follow a *binomial* process.

In the remainder of this chapter, we first provide the details of our canonical setting and formulate the financial economist's problem—find the investment and spending strategies that maximize a retiree's expected utility. We then develop conditions and tests to determine whether an arbitrary retirement strategy is optimal or efficient, and the equations for its revealed utility, when it exists. We then introduce the simple complete market to be used for illustrative purposes. Next we describe a fundamental spending strategy that employs *lockboxes*. We discuss three efficient lockbox strategies and their revealed utilities. We next look at two popular rules of thumb used by financial planners. We show that the first rule, the investment g*lide path* rule, is efficient provided it is paired with a very specific spending rule. We show that the second rule, the constant four percent spending rule, is only efficient when all investments are in riskless securities. Finally, we conclude with a summary of results and some topics for further investigation.

Revealed Utility and Retirement Spending

A retiree, who maximizes his expected utility, is faced with the following problem. For each year in the future, and for all states of the world in each year, our investor must optimally choose how much to consume and an investment policy to support that consumption. If markets are complete, our retiree can purchase contingent claims on the future states, and cash in these securities to pay for consumption. We assume that markets are complete, so that the investment alternatives are known, and only the consumption values are to be determined. Let "t" index future years, "s" index future states, and the pair "t,s" index a state that occurs at time "t". We denote consumption by $C_{t,s}$, the probability that a state occurs by $\pi_{t,s}$, the current price of a contingent claim by $\psi_{t,s}$, utility from consumption at time t by $U_t(C)$, and initial wealth by W_0, Our investor must choose consumption values that maximize the function:

$$(1a) \quad \max \sum \pi_{t,s} \cdot U_t(C_{t,s})$$

and also satisfy the budget constraint:

$$(1b) \quad W_0 = \sum C_{t,s} \cdot \Psi_{t,s}$$

In Eqs.(1), the summations are with respect to all states and times. Note that we have assumed that all states occurring at time t have the same utility function U_t, and that this utility is only a function of consumption at time t. The maximand in Eq.(1a) is the expected utility of the consumption plan, which is assumed to be time separable. We assume that the utility functions $U_t(C)$ are increasing and concave, i.e., $U_t'(C) > 0$ and $U_t''(C) < 0$. In other words, we assume that investors always prefer more to less, and are risk averse. We term this the *canonical retiree problem*: given an initial wealth, to find the set of consumptions at every time and state in the future that will maximize expected utility, where these consumptions are provided by investments in state contingent claims. Throughout, we assume that the retiree has a known separable utility function, knows the probabilities of future states, and knows the prices of contingent claims.

Many economists will find our canonical retiree problem both familiar and sensible, though most practitioners are likely to consider it beyond the pale. How many retirees know their utility functions? Very few, at best. But by choosing a particular retirement financial strategy, a retiree has either made a mistake or revealed something about his preferences. When we examine some popular strategies, in each case we seek to determine (1) whether the strategy is consistent with expected utility maximization, and (2) if so, what are the characteristics of the associated utility function. A strategy that meets condition (1) will be said to *reveal utility* in the sense that we can, if desired, answer question (2) – that is, determine the characteristics of the utility function for which the strategy would be optimal.

The relevant equations for this task are derived from the first order equations for the maximization problem. The full set of such equations includes the budget constraint and the following equations for each time t and state s:

$$(2) \quad U_t'(C_{t,s}) / U_0'(C_{0,0}) = \Psi_{t,s} / \pi_{t,s}$$

In Eq.(2), the right-hand side is the ratio of the state-price to state-probability, sometimes termed the state's price-per-chance (PPC).[2] For any given time in the future, a retirement strategy prescribes a set of consumption values—one for each state. Using these choices and a model that specifies the price-per-chance for each state, we can infer the marginal utility function $U_t'(C)$ for that time period, if it exists. Such a function exists if two conditions are met. First, the strategy must provide a single consumption value for each time and state. Secondly, in order to recover a concave utility function, consumption must be higher in states with lower price-per-chance and the same in all states with the same price-per-chance. More succinctly, if for a given time we rank the states in order of increasing consumption, this must be equivalent to ranking the states in order of decreasing price-per-chance. If such an ordering is possible, we can then integrate the marginal utility function to get a revealed utility function. We note that the revealed utility functions $U_t(C)$ are not completely unique; for each time we can add an arbitrary integration constant, and all times can have a common positive multiple, namely $U_0'(C_{0,0})$. Fortunately this non-uniqueness is economically immaterial.

A key ingredient in our analyses is a model of the characteristics of asset prices. Since contingent claims prices are not observable, we need to make an assumption about the nature of equilibrium in capital markets. We adopt a multi-period generalization of the results obtained with several standard models of asset pricing, such as the Capital Asset Pricing Model, some Pricing Kernel models, and the binomial model employed below. In particular, we assume that the explicit or implicit contingent claim prices at any given time t are a decreasing function of the cumulative return on the overall market portfolio from the present time to that time period. Equivalently, if for a given time we rank the states in order of increasing market return, this must be equivalent to ranking the states in order of decreasing price-per-chance. This is the market setting for our canonical retiree problem.

For the remainder of the chapter, we will consider a state “s” at time “t” to be synonymous with cumulative market return at time “t”. So a retirement strategy must predict a single consumption value for any particular market return, and be independent of prior market returns, i.e., the particular *paths* that led to the final cumulative return. Further, since we assume that the price-per-chance is a decreasing function of market returns, the existence of a revealed utility requires that consumption be an increasing function of market return. Although our illustrations utilize a binomial process, the results apply in other settings as well.

Not all retirement strategies have revealed utilities. Three straightforward tests can be used to identify obvious violators. First, the retirement strategy cannot lead to multiple values of consumption for the same cumulative market return. Consumption must be path independent. Second, the present level of wealth for every state (before or after consumption) must also be path independent. If not, either consumption will ultimately be path dependent (a violation of the first test), or it must be the case that for one or more paths not all wealth will have been spent (a violation of optimality).

The third test used to identify violators is non-obvious, subtle, and exceedingly powerful. We term it the principle of *earmarking*. If knowledge of a state provides knowledge of consumption, and state prices exist, then at any point in time, the retiree’s portfolio can be subdivided into assets that are earmarked for consumption in that state and time. It follows that assets allocated to all states at a given time can be aggregated so that our retiree also can identify the assets earmarked to support spending in any given year. Maximizing expected utility implies that our retiree knows at any point in time how much wealth is currently earmarked for consumption at each future date. If the wealth allocated to consumption at a specific time is uncertain, this uncertainty must translate to uncertainty regarding consumption in at least one state at that time, which necessarily violates maximizing expected utility.

A Simple Complete Market

In this section, we describe the simple *complete market* we use in the remainder of this chapter. Generally, a market is complete if the set of all contingent claims can be constructed using its assets. Our simple market has just two assets, a deterministic risk-free asset and a stochastic risky asset. The yearly returns on the risk-free asset are assumed constant, while the returns on the risky asset will track the returns of the total market portfolio. We assume that in any year the market is equally likely to move up or down, that the characteristics of the movements in each year are the same, and that the movement in any year is independent of the actual movements in prior years. More succinctly, the market moves are independent and identically distributed coin-flips, and thus the total number of up-moves (or down-moves) over a span of years has a *binomial distribution*.[3]

Given the above assumptions, our complete market is specified by three parameters: (1) the total annual market return R_u for an up-move, (2) the total annual market return R_d for a down-move, and (3) the total annual return R_f on a risk-free asset. All three of these annual returns are assumed to be real. For example, the values $R_u = 1.18$, $R_d = 0.94$, and $R_f = 1.02$ give a market portfolio with an annual expected real-rate of return of six percent, a volatility of 12 percent, and a Sharpe Ratio of 1/3. These values roughly correspond to an aggregate market portfolio made up of 40 percent bonds and 60 percent equities.

We take the initial value of the market portfolio to be one, which is the value at the root of the binomial tree. After one year, the market value is equal to $\mathbf{V}_{m,1} = \mathbf{R}_{m,1}$. This value, the random total market return for the first year, can equal either R_u or R_d. In Figure 1, we draw two paths emanating from the initial market value, one up and one down, that connect the initial value to the two possible values at t = 1. After two years, the market value is equal to the random product $\mathbf{V}_{m,2} = \mathbf{R}_{m,1} \cdot \mathbf{R}_{m,2}$. It can have one of three possible values $\{ R_u^2, R_u \cdot R_d, R_d^2 \}$, but there are now four different, equally likely paths, { up-up, up-down, down-up, down-down } connecting the initial value with the three final values. The up-down path and the down-up path lead to the same market return, namely $R_u \cdot R_d$, and this value is twice as likely as either of the two possible paths that lead to it. After t years, the market value $\mathbf{V}_{m,t} = \mathbf{R}_{m,1} \cdot \mathbf{R}_{m,2} \cdots \mathbf{R}_{m,t}$ can have one of (t+1) possible values, $\{ V_{t,s} = R_u^{s} \cdot R_d^{t-s} \mid 0 \leq s \leq t \}$, where the parameter "s" is the total number of "up-moves"—a useful parameter for indexing the market values. On the other hand, there are 2^t paths between the initial state and the final states. The number of paths that have the market value indexed by s at time t will equal the binomial coefficient for "t-choose-s". Hence, the state probability $\pi_{t,s}$, the probability that the market's value is equal to $V_{t,s}$, is given by the expression:

$$(3) \quad \pi_{t,s} = \frac{t!}{s! \cdot (t-s)!} \cdot 2^{-t}$$

In Eq.(3), an exclamation point is used to denote a factorial.

Figure 1 here

Associated with every path of the binomial tree is the price today of a security that pays \$1 if and only if that path is realized. We term these securities path-contingent claims. We can use standard arbitrage pricing techniques to compute any such price. The price of a claim to receive \$1 at a given time and state is the cost of a dynamic strategy using the market and the risk-free asset that will provide this amount and nothing at any other time and state. For example, the current option price f_u for the first period up-move path and the current option price f_d for the first period down-move path are the following functions of R_f, R_u, and R_d:

$$\text{(4a)} \quad f_u = \frac{R_f - R_d}{R_f \cdot (R_u - R_d)}$$

$$\text{(4b)} \quad f_d = \frac{R_u - R_f}{R_f \cdot (R_u - R_d)}$$

The inequality $R_d < R_f < R_u$ is a necessary condition for positive prices. The prices for the two period paths can be written in terms of the one-period prices: $f_{uu} = f_u^2$, $f_{ud} = f_{du} = f_u \cdot f_d$, and $f_{dd} = f_d^2$. More generally, for a t-year path, the price is equal to $(f_u^s \cdot f_d^{t-s})$, where s is the number of up-moves. Hence, the option price for all paths that end at the market state "s" are the same, and depend only on the total number of up-moves and down-moves, not on the particular sequence of up-moves and down-moves. Thus today's price $\psi_{t,s}$ of a state claim that pays \$1 if and only if state s occurs is equal to the number of paths to the state times the price of each path.

$$\text{(5)} \quad \psi_{t,s} = \frac{t!}{s! \cdot (t-s)!} \cdot f_u^s \cdot f_d^{t-s}$$

We assume that markets are complete, or at least sufficiently complete, so dynamic strategies involving the market and the risk-free asset can replicate any state-claim. Finally for a fixed value of t, the sum of all the state prices is equal to $1/R_f^t$, the price of a risk-free dollar, t-years from now. This must be the case, since purchasing all the state claims available at year t guarantees the investor a dollar in year t, no matter which state is realized.

We are now in a position to show that for our binomial model, the price-per-chance is a decreasing function of cumulative market return. First, the price-per-chance and the cumulative market value are given by the formulas:

$$\text{(6a)} \quad \psi_{t,s}/\pi_{t,s} = 2^t \cdot f_u^s \cdot f_d^{t-s}$$

$$\text{(6b)} \quad V_{t,s} = R_u^s \cdot R_d^{t\text{-}s}$$

If we take the logarithm of each equation, we obtain two equations that are linear in s. After we eliminate the parameter s from this pair, we get the following simple relation between price-per-chance and cumulative market return:

$$\text{(7a)} \quad \psi_{t,s}/\pi_{t,s} = a^t / V_{t,s}^p$$

In Eq.(7a), the power p and time-factor a are constants defined by:

$$\text{(7b)} \quad p = \ln\left(\frac{R_u - R_f}{R_f - R_d}\right) \div \ln(R_u/R_d)$$

$$\text{(7c)} \quad a = 2 \cdot f_d \cdot R_d^p$$

Generally, p and a are positive, and so price-per-chance is a decreasing function of total cumulative market return. In our numeric example, p = 3.05 and a = 1.08.

Lockbox Spending Strategies

Next we introduce and illustrate *lockbox* strategies, an approach to spending which divides a retiree's initial wealth among separate accounts, one account for each future year of spending. The assets in each account are dynamically managed according to the account's exogenous investment rule. When an account reaches its target year, our retiree cashes out its investments, closes the account, and spends all of its proceeds. The accounts can be real or virtual, and we collectively call them lockboxes—a term that emphasizes the retiree's implicit obligation to yearly spend all the assets from the target account and to never co-mingle or spend the assets of any of the remaining accounts.

All efficient strategies adhere to the earmarking principle and have a lockbox formulation, however, there are inefficient lockbox strategies. The test for efficiency is simple; each lockbox's value must be a path-independent, increasing function of the cumulative market. For example, lockboxes that alternate investments in the risk-free and market assets are obviously path dependent and inefficient. In the remainder of this section, we pair three different investment strategies with lockbox spending. For each pair we show that the resulting retirement strategy is efficient and derive its revealed utility.

Consider the lockbox strategy where each lockbox is invested in the market portfolio. Suppose our retiree has a planning horizon of T years, and has assigned F_0 dollars for today's consumption, placed F_1 dollars in the first lockbox, F_2 in the second and, more generally, F_t in the t-th year lockbox. The total assigned dollars will sum to the initial wealth, i.e., $W_0 = F_0 + F_1 + \ldots + F_T$. At the end of each year, the consumption from the t-th lockbox will equal F_t times the cumulative return of the market:

(8a) $$C_{t,s} = F_t \cdot V_{t,s}$$

Note that for each state at each time there will be a unique amount of consumption, and this will be an increasing function of the market value. We see immediately that this investor has a revealed utility. The revealed marginal utility follows from Eqs.(2) and (7):

(8b) $$U'_t(C) = a^t \cdot \left(F_t/C\right)^p$$

To obtain Eq.(8b), we set $U_0'(C_{0,0})$ equal to one since the entire set of an investor's utility functions can be multiplied by a constant without changing the implied optimal strategy. This strategy is thus optimal for an investor with a utility function that exhibits constant relative risk-aversion, generally abbreviated as CRRA, with risk-aversion parameter p. Further, our investor's attitudes towards consumption in future states relative to the present are revealed by the dollars assigned to the lockboxes.

Now, suppose that instead of investing solely in the market, our retiree invests $F_{m,t}$ dollars of lockbox t in the market and $F_{f,t}$ dollars in the risk-free asset, with the sum of the dollars invested equal to W_0. Once the initial allocation is made, our investor adopts a buy and hold investment strategy. In practice, this investment strategy could be implemented by purchasing a zero-coupon bond and a market exchange-traded fund for each lockbox. When the t-th lockbox is opened and cashed out, the consumption will be:

$$\text{(9a)} \quad C_{t,s} = F_{f,t} \cdot R_f^t + F_{m,t} \cdot V_{t,s}$$

As long as at least some dollars are allocated to the market, consumption will be an increasing function of market returns. Solving for the revealed marginal utility we obtain:

$$\text{(9b)} \quad U_t'(C) = a^t \cdot \left(\frac{F_{m,t}}{C - F_{f,t} \cdot R_f^t} \right)^p$$

In this case, the strategy is optimal for an investor with a HARA utility function—one that exhibits hyperbolic absolute risk aversion. In effect, the investor requires a minimum subsistence level equal to the amount provided by the allocation to the risk-free asset and has constant relative risk aversion with respect to the amount provided by the allocation to the market.

Our third example has lockboxes invested in constant-mix, constant-risk portfolios. Specifically, we annually rebalance a lockbox's assets so that a fraction β is invested in the market portfolio and the remaining fraction (1-β) is invested in the risk-free asset. We impose a no-bankruptcy condition; hence the total return must be positive in either an up or down state, and so β is limited to the range:

$$\text{(10a)} \quad -R_f / (R_u - R_f) = \beta_{min} < \beta < \beta_{max} = +R_f / (R_f - R_d)$$

For our parameter choices, the bounds are β_{min} = -6.38 and β_{max} = 12.75. The total annual returns of the mix follow a binomial model, and the cumulative return $M_{t,s}(\beta)$ of the mix at time t and in state s is given by:

$$(10b) \quad M_{t,s}(\beta) = [(1-\beta)\cdot R_f + \beta \cdot R_u]^s \cdot [(1-\beta)\cdot R_f + \beta \cdot R_d]^{t-s}$$

Again, s denotes the number of up-moves in the path to time t.

As before, we let F_t's be the amounts of initial wealth allocated to the lockboxes, and we introduce β_t's as the constant mixes for the lockboxes. Though the risk in any given lockbox is constant, the risks among all the lockboxes are allowed to vary. It follows that the spending from a constant-mix lockbox is given by:

$$(11a) \quad C_{t,s} = F_t \cdot M_{t,s}(\beta_t)$$

We see from the previous equations that consumption will be an increasing function of s, provided the market exposures are non-negative. But Eq.(6a) showed that price-per-chance is a decreasing function of s. Thus for a constant mix strategy, consumption at time t will be a decreasing function of price-per-chance. Therefore the investor's utility function for that period will be revealed. Moreover, Eq.(6b) showed that market return is an increasing function of s. Thus a constant-mix strategy will have no non-market risk and will be efficient.

Using Eqs.(6b) and (10b), we can eliminate the parameter s from Eq.(11a) and write $C_{t,s}$ as an increasing function of $V_{t,s}$. The revealed marginal utility then follows:

$$(11b) \quad U'_t(C) = a_t^t \cdot (F_t/C)^{\gamma_t}$$

$$(11c) \quad \gamma_t = \ln\left[\frac{R_u - R_f}{R_f - R_d}\right] \div \ln\left[\frac{(1-\beta_t)\cdot R_f + \beta_t \cdot R_u}{(1-\beta_t)\cdot R_f + \beta_t \cdot R_d}\right]$$

$$(11d) \quad a_t = 2 \cdot f_d \cdot [(1-\beta_t)\cdot R_f + \beta_t \cdot R_d]^{\gamma_t}$$

Again, our retiree has a CRRA utility, but in this case, the retiree's choice for the exposures β_t determines the risk-aversions γ_t. Both the exposures β_t and initial allocations F_t determine the retiree's relative preference for consumption today, versus the future. We note if all the exposures are equal to one, the market exposure, then Eq.(11b) reduces to Eq.(8b), the result for the market only strategy.

The above three examples illustrate efficient financial retirement strategies and their revealed utility functions. Since financial economists often use CRRA or HARA models for utility, they may likely suggest one of our example strategies to a retiree. On the other hand, financial planners, who tend to rely on rules of thumb for investing and spending, would rarely advise one of the above combinations of investing and spending strategies. In the next two sections, we evaluate the efficiency of two of the most common rules.

Glide Path Investment Strategies

Many advisors recommend that retirees annually adjust their portfolios by decreasing their exposure to equities, and thus reducing their overall risk. This rule of thumb, often called a *glide path* strategy, is an age-based investment strategy. A classic example is the oft-quoted 100-minus-age rule for the percentage of assets allocated to equities, e.g., 60 year-olds should hold 60 percent of their assets in bonds and 40 percent of their assets in equities. Many retirees follow a glide path strategy by investing in *lifecycle funds*—age-targeted, managed funds intended to serve as the sole investment vehicle for all of a retiree's assets. In recent years, interest in lifecycle funds has exploded. Jennings and Reichenstein (2007) analyzed the policies of some leading lifecycle funds and found that a 120–minus-age equity allocation describes the typical management rule. An earlier paper by Bengen (1996) suggested that the target equity allocation should equal 128-minus-age for clients up to age 80 and 115-minus-age afterwards. The advocates of glide path strategies often pair this investment rule with one or more options for a spending rule. However, there is only one spending rule that makes the complete retirement strategy efficient, and that rule is the focus of this section.

The investing rules described above specify equity percentages, but our market model deals more conveniently with market fractions. However, there is a simple linear relationship between the two descriptions. For example, our sample parameters roughly correspond to a market of 60 percent equities. In this case, the equity mix is 0 percent when $\beta = 0$, is 60 percent when $\beta = 1$, and is 100 percent when equals $\beta = {}^{5}/_{3}$. Now, consider a 65 year-old retiree following the 120-minus-age rule. This retiree has the annual equity percentage targets of 55 percent, 54 percent, 53 percent, etc. and the annual market fraction targets of ${}^{55}/_{60}$, ${}^{54}/_{60}$, ${}^{53}/_{60}$, etc. Because age-based rules are easily translated into a market-fraction time series, we generally use the latter to describe a glide path.

Consider the generic glide path investment and spending strategy. At the beginning of each year, some portion of the portfolio is spent; the fraction β_t of the remainder is invested in the market, and the rest is invested in the risk-free asset. When this total portfolio strategy is efficient, it has a lockbox equivalent. We use this equivalence principle to derive the optimal spending rule. We start by choosing any one of the lockboxes, say the j-th box, and virtually combine its contents and the contents of all succeeding lockboxes. Initially, the j-th combined portfolio will have value $F_j + \ldots + F_T$, where the F_t's are again the initial lockbox allocations. The future values of a combined portfolio must satisfy two requirements. First, they must evolve in a path independent manner, just like the values of the constituent lockboxes. Secondly, since the j-th combined portfolio *is* the total portfolio for the j-th year, this combined portfolio must have the glide path's market fraction β_j in the j-th year, independent of the market state. Now, as we saw in the previous section, a constant-mix portfolio with market exposure β_j satisfies both of these requirements; in fact, it can be shown that every combined portfolio is a constant-mix portfolio. If we let the random variable $\mathbf{\Sigma}_{j,t}$ be the value of the j-th combined portfolio at year $0 \leq t \leq j$, then we have:

$$\text{(12a)} \quad \mathbf{\Sigma}_{j,t} = \left(F_j + \ldots + F_T\right) \cdot \mathbf{M}_t\left(\beta_j\right)$$

where $\mathbf{M}_t(\beta)$ is the random cumulative return at year t for the constant-mix portfolio with market weight β; its value in state s at time t is given by Eq.(10a).

Given the combined portfolios for an efficient glide path, the individual lockbox holdings follow immediately. First, the lockbox for T is just the combined portfolio for T; a constant-mix portfolio with exposure β_T and initial allotment F_T. The remaining lockbox portfolios are obtained by differencing successive combined portfolios. Let the random variable $\mathbf{\Lambda}_{j,t}$ be the value of j-th lockbox at time t:

$$\begin{aligned} \text{(12b)} \quad \mathbf{\Lambda}_{j,t} &= \mathbf{\Sigma}_{j,t} - \mathbf{\Sigma}_{j+1,t} \\ &= \left(F_j + \ldots + F_T\right) \cdot \mathbf{M}_t\left(\beta_j\right) - \left(F_{j+1} + \ldots + F_T\right) \cdot \mathbf{M}_t\left(\beta_{t+1}\right) \\ &= F_j \cdot \mathbf{M}_t\left(\beta_j\right) + \left(F_{j+1} + \ldots + F_T\right) \cdot \left[\mathbf{M}_t\left(\beta_j\right) - \mathbf{M}_t\left(\beta_{t+1}\right)\right] \end{aligned}$$

The initial lockbox holds cash, the last lockbox holds a constant-mix portfolio, and the middle lockboxes hold a combination of assets; the first is a constant-mix asset, and the second is a "swap" between two constant-mix assets. Finally the efficient spending is given by $\mathbf{C}_t = \mathbf{\Lambda}_{t,t}$, or in terms of states:

$$\text{(12c)}\quad C_{t,s} = \begin{cases} F_0, & t = 0 \\ F_t \cdot M_{t,s}(\beta_t) + (F_{t+1} + \ldots + F_T) \cdot [M_{t,s}(\beta_t) - M_{t,s}(\beta_{t+1})], & 0 < t < T \\ F_T \cdot M_{T,s}(\beta_T), & t = T \end{cases}$$

It is tedious, but straightforward, to directly verify that the above spending rule, coupled with its glide-path investment rule, is efficient. Further, though there is no simple function to describe the revealed utility, its values can be easily computed numerically.

Glide paths may well reflect the desires of many retirees to take less risk concerning their investments as they age, but these retirees' retirement strategies will be inefficient unless spending follows Eq.(12c). Glide path rules are ubiquitous, but their complementary spending rules are rare. In fact we are unaware of any retiree that computes his annual spending according to the above rule. As an alternative to the glide path strategy, we recommend the constant-mix lockbox strategy discussed in the previous section. If a retiree decreases the market fractions for successive lockboxes, then his total portfolio risk will tend to decrease over time. Thus, a retiree can retain the desired feature of the glide path, but have a much simpler spending rule.

The Four Percent Rule

Many recent articles in the financial planning literature have attempted to answer the question: "How much can a retiree safely spend from his portfolio without risking running out of money?" Bengen (1994) examined historical asset returns to determine a constant spending level that would have had a low probability of failure. He concluded that a real value equal to approximately four percent of initial wealth could be spent every year, assuming that funds were invested with a constant percentage in equities within a range of 50 percent to 75 percent. Cooley, Hubbard, and Walz (1998) used a similar approach and found that a four percent spending rule with inflation increases had a high degree of success assuming historical returns and at least a 50 percent equity allocation. Later, Pye (2000) concluded that with a 100 percent allocation to equities, the four percent rule would be safe enough if equity returns were log-normally distributed with a mean return of eight percent and a standard deviation of return equal to 18 percent.[4] Based on this research, there is a growing consensus that newly retired individuals with funding horizons of thirty to forty years can safely set their withdrawal amount to four percent of initial assets and increase spending annually to keep pace with inflation. This is the foundation for the now common *four percent rule* of thumb for retirement spending.

An efficient retirement strategy must be totally invested in the risk-free asset to provide constant spending in every future state.[5] However, the generic four percent rule couples a risky, constant-mix investment strategy with a riskless, constant spending rule. There is a fundamental mismatch between its strategies, and as a result it is inefficient. The following simple example illustrates these points. Consider a retiree who, whether the market goes up or down, wants to spend only \$1 next year. He can achieve this goal by investing $1/R_f$ dollars in the risk-free asset. On the other hand, if he uses the market asset, he must increase his investment to $1/R_d$ dollars, so that if the market goes down, the investment pays the required \$1. However, if the market goes up, the investment pays (R_u/R_d) dollars, and there is an unspent surplus. So, if our retiree truly requires just \$1, then investing in the market is less efficient than investing in the risk-free asset because of the greater cost and the potential unspent surplus.

We can use the above argument to investigate a more general case. Suppose a retiree wants to support a constant spending level $C_{t,s} = f \cdot W_0$ for T years from a portfolio with initial wealth W_0 that is invested in a possibly, time-dependent strategy, e.g., a glide path. Further, let D_t equal the minimum total return of the portfolio in year t. These minimums will correspond to down (up) moves for portfolios with positive (negative) market fractions β_t. Then to insure against the worst-case scenario, a safe spending fraction f must satisfy the equation:

$$(13) \quad \frac{1}{f} = 1 + \frac{1}{D_1} + \frac{1}{D_1 \cdot D_2} + \ldots + \frac{1}{D_1 \cdot D_2 \cdot \ldots \cdot D_T}$$

The most efficient investment will yield the largest spending fraction f, which corresponds to maximizing the minimum returns D_t. However, in any period, the best of the worst is achieved by investing in the risk-free asset, and thus $\beta_t = 0$. For example, suppose a retiree has a planning horizon of 35 years and invests in the risk-free asset, then $D_t = R_f = 1.02$ and f = 3.85 percent. On the other hand, if the retiree insists on investing in the market portfolio, then $D_t = R_d = 0.94$ and f = 0.77 percent, a five-fold decrease in spending. For the risk-free asset ($D_t > 1$), each successive year is cheaper to fund, but for the market portfolio ($D_t < 1$), each successive year is more costly.

The safe spending fraction satisfies Eq.(13). With this spending level, all scenarios, other than the worst-case scenario, will have an unspent surplus. If we raise the spending fraction just a bit, then the worst-case scenario will be under funded and the spending plan will collapse if this path is realized. As we continue to raise the level, more and more scenarios will be under funded, a few may be spot on, and the remaining will have a surplus. If our example retiree insists both on investing in the market and increasing his spending fraction to four percent, then approximately 10 percent of scenarios will be under-funded and the remaining 90 percent of scenarios result in an unspent surplus. Further, more than 50 percent of the scenarios will have a surplus more than twice initial wealth! It is very unlikely that this retiree, who desired a riskless spending plan, would find such an eschewed-feast or famine plan acceptable.

This type of analysis generalizes to any given desired spending plan. With complete markets, any given spending plan has a unique companion investment plan that delivers the spending at minimum cost. With state-contingent securities, the minimum cost investment plan involves simply purchasing the contingent claims that deliver the desired spending. Given our simple complete market, the contingent claims must be translated into dynamic strategies utilizing the market and riskless assets. Deploying this minimum required wealth using any other investment strategy necessarily results in surpluses and deficits relative to the desired spending plan. Extra wealth must then be introduced to eliminate all deficits.

The preceding assumed individual preferences were consistent with a fixed spending plan and demonstrated the inefficiency of a market investment plan. If we instead assume the investment plan is indicative of preferences, then we need to find a spending plan consistent with a market portfolio investment plan. This problem was previously analyzed, and the spending solution is reported in Eq.(8a). If a market investment plan is indicative of preferences, then all efficient spending plans require spending that is proportional to cumulative market returns.

The four percent rule does not generate a revealed utility because the investment and spending rules do not correspond to an efficient retirement strategy. Retirees interested in fixed retirement spending should invest in the risk-free asset. Anyone who chooses to invest in the market should be prepared for more volatile spending. Either can adopt an efficient strategy. However, a retiree who plans to spend a fixed amount each period, while investing some or all funds in the market, faces a very uncertain future. Markets could perform well, and his wealth would far exceed the amount needed to fund his desired spending, or they could perform poorly, and his entire spending plan would collapse.

Conclusion and Discussion

Virtually all retirees have an explicit or implicit retirement spending and investment strategy. What is striking is the gulf that exists between how financial economists approach the problem of finding optimal retirement strategies and the rules of thumb typically utilized by financial advisors. Aside from identifying this gap, our objective with this chapter has been to evaluate the extent to which several popular retirement spending and investment strategies are consistent with expected utility maximization. This evaluation has two stages. First, is the given rule of thumb consistent with expected utility maximization for *any* investor? Second, if it is, how must the rule's investment and spending strategies be integrated to achieve and maintain efficiency?

By and large, we find that the strategies analyzed fail one or more of our tests. Investment rules suggesting risk glide paths pass the first assessment in that they are not *per se* inconsistent with expected utility maximization. However, the conditions on the implied spending rule required by efficiency seem onerous and unlikely to be followed by virtually any retirees. While risk glide paths only specify suggested investments, the four percent rule is fairly explicit about both the recommended spending and investment strategy. Unfortunately, the four percent rule represents a fundamental mismatch between a riskless spending rule and a risky investment rule. This mismatch renders the four percent rule inconsistent with expected utility maximization. Either the spending or the investment rule can be a part of an efficient strategy, but together they create either large surpluses or result in a failed spending plan.

While most of our results are obtained using a simple binomial model of the evolution of asset returns, many hold in more general settings, as we intend to show in subsequent research. Our results suggest a reliance on lockbox spending strategies, a very different type of retirement financial strategy than those currently advocated by practitioners. To an extent, this may be attributable to the assumptions we have made concerning both the nature of the capital markets and the objectives of the retiree. It is at least possible that one or more of the standard rules may be appropriate if prices are set differently in the capital markets and/or the investor has a different type of utility function. For example, one might posit that returns are not independent, but negatively serially correlated. Or one might focus on the efficiency of a strategy for an investor whose utility for consumption at a given time depends on both the consumption at that time and consumption in prior periods. However, we suspect that it may be difficult to prove that the practitioner rules we have analyzed are efficient even in such settings.

Much of the analysis in this chapter relates to identifying problems with existing rules of thumb, but thus far, we have only hinted at ways to remedy the situation. An interesting line of inquiry would address this gap by finding an efficient strategy that strictly dominates an inefficient strategy such as one of those advocated by practitioners. There are two types of inefficiencies that could be introduced. First, a given retirement strategy could inefficiently allocate resources. That is, the same set of outcomes could be purchased with fewer dollars.[6] Given this inefficiency, a revised strategy could be constructed that strictly dominates the original strategy in that the revised strategy would increase spending in at least one state without decreasing spending in any state. A second type of inefficiency occurs when a strategy entails multiple spending levels for a given market return. If the total present value allocated to purchase the multiple spending levels were instead used to purchase a single spending amount, then as long as the expected returns in all such states are the same, any such replacement would be preferred by any risk-averse investor (formally, the revised set of spending amounts would exhibit second-degree stochastic dominance over the initial set). By making all such possible replacements, an inefficient strategy could be converted to a dominating efficient strategy. Another line of inquiry involves the examination of the properties of the revealed utility function associated with any efficient strategy, whether advocated initially or derived by conversion of an inefficient strategy. Such examination might reveal preferences that are inconsistent with those of a particular retiree and hence the strategy, while efficient, would not be appropriate in the case at hand.

Overall, our findings suggest that it is likely to be more fruitful to clearly specify one's assumptions about a retiree's utility function, then to establish the optimal spending and investment strategy directly. Of course, one should take into account more aspects of the problem than we have addressed in this chapter. Annuities should be considered explicitly, rather than ruled out *ex cathedra*. Separate utility functions for different personal states (such as "alive" and "dead") could be specified rather than using a weighted average using mortality probabilities, as we have assumed here. Yet our analysis suggests that rules of thumb are likely to be inferior to approaches derived from the first principles of financial economics.

The authors thank John Ameriks, Olivia S. Mitchell, and Stephen Zeldes for many valuable comments on an early draft of this chapter. We also wish to thank our colleagues at Financial Engines Wei Hui and Jim Shearer for their insights and support.

References

Bengen, William P. (1994). 'Determining Withdrawal Rates Using Historical Data.' *Journal of Financial Planning*, 7(4): 171-180.

Bengen, William P. (1996). 'Asset Allocation for a Lifetime.' *Journal of Financial Planning,* 9(4): 58-67.

Cochrane, John H. (2005). *Asset Pricing: Revised Edition*. Princeton, NJ: Princeton University Press: 3-147.

Cooley, Phillip L., Carl M. Hubbard and Daniel T. Walz. (1998). 'Retirement Savings: Choosing a Withdrawal Rate That Is Sustainable.' *The American Association of Individual Investors Journal*, February: 16-21.

Dybvig, Phillip H. (1988a). 'Distributional Analysis of Portfolio Choice.' *Journal of Business*, 61: 369-93.

Dybvig, Phillip H. (1988b). 'Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market.' *Review of Financial Studies*, 1:67-88.

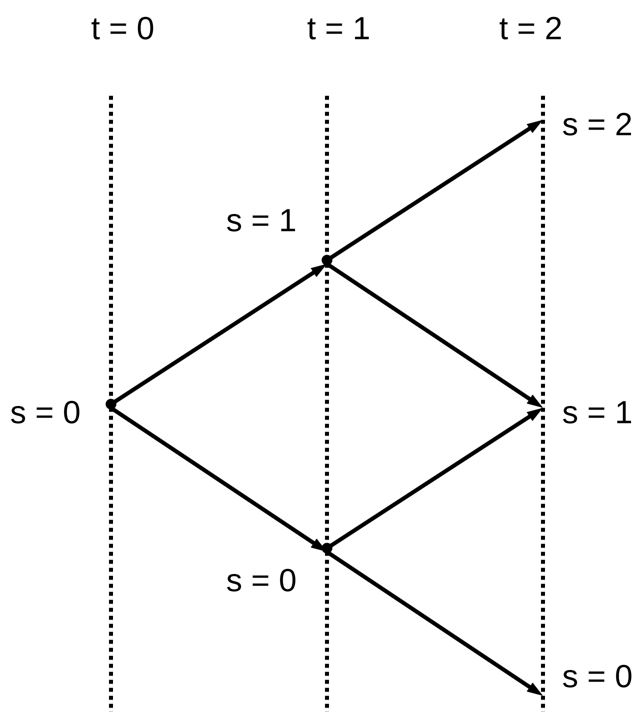
Jennings, William W. and William Reichenstein. (2007). 'Choosing the Right Mix: Lessons From Life Cycle Funds.' *The American Association of Individual Investors Journal,* January: 5-12.

Merton, Robert C. (1971). 'Optimum Consumption and Portfolio Rules in a Continuous-Time Model.' *Journal of Economic Theory*, 3: 373-413.

Pye, Gordon B. (2000). 'Sustainable Investment Withdrawals.' *The Journal of Portfolio Management*, 26(4): 73-83.

Sharpe, William F. (2007). *Investors and Markets: Portfolio Choices, Asset Prices, and Investment Advice*. Princeton, NJ: Princeton University Press: 74-100.

Figure 1. An illustration of a two-year binomial tree.
Source: Authors' conceptions.
Note: Each successive year (t) has one more possible market state (s). In the second year (t = 2) the middle state (s = 1) can be reached by either following the up-down path or down-up path.



Endnotes

[1] We are assuming that for each time period, there is a utility function that gives the utility *measured today* as a function of the amount consumed in that period. Moreover, we assume that the investor prefers more to less and is risk-averse, so the utility function for a time period increases with consumption at a decreasing rate. The expected utility of consumption in a time period is simply the probability-weighted average of the utilities of the amounts consumed in different scenarios at the time. Finally, the expected utility of the retirement plan is the sum of the expected utilities for each of the time periods.

[2] Sharpe uses the term price-per-chance or PPC for the ratio of a state's price to its probability (Sharpe 2007). As discussed by Cochrane (2005), this quantity is also called the marginal rate of substitution, the pricing kernel, a change of measure, and the state-price density.

[3] Though our binomial model for annual market returns may appear highly restrictive, similar models using shorter time periods are often used in both the academy and financial sector for pricing options and predicting the results of investment strategies.

[4] Pye (2000) also shows that a 60 percent initial allocation to TIPS improves the allowable withdrawal to 4.5 percent, while simultaneously lowering the measured downside-risk.

[5] As the market fraction β approaches zero, constant-mix lockboxes are invested in just the risk-free asset and provide constant spending. Further, the risk aversion parameter of the underlying CRRA utility approaches infinity in this limit. Alternatively, state-independent spending can be viewed as the limit of the buy and hold lockbox for which all wealth is allocated to the risk-free asset and none in the market asset. Here, the subsistence levels of the underlying HARA utility exhaust the budget.

[6] Dybvig (1988a, b) explored inefficient portfolio strategies in a pair of papers. His approach is very useful for analyzing retirement strategies such as the four percent rule.