

Optimal Asset Allocation for Retirement Saving: Deterministic vs. Adaptive Strategies

Peter A. Forsyth^a

Kenneth R. Vetzal^b

March 5, 2018

Abstract

We consider optimal asset allocation for an investor saving for retirement. The portfolio contains a bond index and a stock index. We use multi-period mean-variance criteria and explore two types of strategies: *deterministic* strategies are based only on the time remaining until the anticipated retirement date, while *adaptive* strategies also consider the investor's accumulated wealth. The vast majority of financial products designed for retirement saving use deterministic strategies (e.g. target date funds). We develop methods which give optimal mean-variance allocations for both deterministic and adaptive strategies. Tests based on both a synthetic market where the stock index is modeled by a jump diffusion process and also on bootstrap resampling of long-term historical data show that the optimal adaptive strategy significantly outperforms the optimal deterministic strategy. This suggests that investors are not being well-served by the strategies currently dominating the marketplace.

Keywords: finance, mean-variance, dynamic asset allocation, jump diffusion, resampled backtests, deterministic strategy, adaptive strategy

1 Introduction

Saving for retirement is one of the most important financial tasks faced by individuals. The total value of retirement assets in the U.S. at the end of 2016 was about \$25 trillion (ICI, 2017), exceeding U.S. GDP for that year by around 35%. More than 60% of these assets were held in individual retirement accounts and defined contribution (DC) pension plans, reflecting the long-term decline in traditional defined benefit (DB) plans. The fundamental reason underlying this trend is that DB plans are seen as a high risk liability for many organizations, and the risk is being transferred to employees through vehicles such as DC plans.

In a DC plan, the employee contributes a fraction of her salary to a tax-advantaged account. This amount is often matched by the employer. The employee is responsible for managing the investments in the account. An accumulation period lasting 30 years would not be unusual, followed by a de-accumulation (retirement) phase of another 20 years, so the employee could end up managing a significant portfolio for 50-60 years. This makes participants in DC plans truly long-term investors.

This study deals with the accumulation phase. Several observers have expressed doubts about the ability of individuals to adequately save for retirement (e.g. Benartzi and Thaler, 2001; 2007;

^aDavid R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1, paforsyt@uwaterloo.ca, +1 519 888 4567 ext. 34415.

^bSchool of Accounting and Finance, University of Waterloo, Waterloo ON, Canada N2L 3G1, kvetzal@uwaterloo.ca, +1 519 888 4567 ext. 36518.

Choi et al., 2004). Three general concerns are (i) whether individuals enrol in savings plans, (ii) if they contribute enough, and (iii) whether they choose appropriate investments. With respect to the first two concerns, significant progress has been achieved through *automatic enrolment* and *automatic escalation*. Automatic enrolment exploits the tendency for individuals to stick with the status quo. Employees are put into a plan by default, while having the choice to easily opt out, instead of having to actively choose to participate. Firms adopting automatic enrolment have seen very strong increases in plan participation rates (Madrian and Shea, 2001; Benartzi and Thaler, 2007). Automatic escalation involves increasing contribution rates over time, as the employee's salary goes up. The Pension Protection Act of 2006 encouraged firms to adopt both automatic enrolment and automatic escalation, and by 2011 over half of firms offering 401(k) plans were doing so (Benartzi and Thaler, 2013).

Offering automatic enrollment entails specifying a default investment option, the third concern noted above. This asset allocation issue is the focus of this study. More than a decade ago, it was common to offer a low-risk default choice such as a money market savings account (Choi et al., 2004). The obvious concern this raised was whether investors could realize high enough returns to accumulate sufficient retirement funds, without taking on more risk. Target date funds (TDFs, also known as lifecycle funds) have become a significant component of the industry's response to doubts that individual investors would be capable of appropriately managing the risk of their retirement savings portfolios. The buyer of a TDF specifies a target date, normally the anticipated retirement date. The most basic TDF consists of a bond index and an equity index. A typical TDF specifies a *glide path*, which determines the fraction of the total portfolio that is invested in the equity index (with the remainder in the bond index) as a function of time. The Pension Protection Act of 2006 permitted TDFs to be used as default investment options in DC plans. Total assets invested in U.S. TDFs have increased dramatically over the past decade, reaching \$887 billion at the end of 2016, up from \$70 billion in 2005 (ICI, 2017, Figure 7.25). The single largest provider of TDFs is Vanguard, with total net assets of about \$280 billion at the end of 2016.¹ Vanguard reports that:

Nine in 10 plan sponsors offered target-date funds at year-end 2016, up over 50% compared with year-end 2007. Nearly all Vanguard participants (97%) are in plans offering target-date funds. 72% of all participants use target-date funds. Two-thirds of participants owning target-date funds have their entire account invested in a single target-date fund. 46% of all Vanguard participants are wholly invested in a single target-date fund, either by voluntary choice or by default (Vanguard, 2017, p. 3).

Moreover, at the end of 2016 83% of Vanguard DC plans specified TDFs as the default investment choice (Vanguard, 2017, Figure 62). Given the propensity of participants to stay with default options, continued strong growth of TDFs appears very likely over the next few years.

The prototypical TDF glide path has a high allocation to stocks during the early years of the accumulation phase. The equity allocation is decreased (and the bond allocation increased) as the time remaining to the target date declines. The underlying rationale is that with many years to retirement, the investor can take on more risk since there is time to recover from adverse market returns. However, as the target date nears, the portfolio is weighted more to bonds as protection against market downturns. This seems to be an intuitively appealing strategy.

The vast majority of TDFs use a *deterministic* glide path. In other words, the bond-stock split is only a function of the time remaining until the target date. This contrasts with an *adaptive* strategy, where the asset allocation can be a function of the time remaining and the accumulated wealth so far. In control terminology, a deterministic strategy can be interpreted as open loop control, while

¹See <http://news.morningstar.com/cover/videocenter.aspx?id=788811>.

an adaptive strategy is a closed loop control. Adaptive strategies have not received much attention to date. One exception is Basu et al. (2011), who consider a type of adaptive strategy using heuristic adjustments based on cumulative investment performance. In particular, they propose strategies that are 100% allocated to equities for a lengthy period, e.g. 20 years. Subsequently, the asset allocation can be switched to 80% equity and 20% in fixed income if overall performance has been satisfactory relative to a specified target; otherwise the portfolio remains completely invested in equities. Portfolio performance is then re-evaluated each year, with similar adjustments based on cumulative performance relative to target. While the adaptive strategies we consider here are similar in spirit, they are based on more robust methods of stochastic optimal control, in contrast to the *ad hoc* adjustments proposed by Basu et al. (2011).

We restrict attention here to an investment portfolio containing a stock and bond index. We model the real (inflation-adjusted) stock index as following a jump diffusion, with the jumps having a double exponential distribution (Kou, 2002; Kou and Wang, 2004). The jump component allows for skewed and leptokurtic returns, and the double exponential distribution fits equity index returns better than a model with lognormally distributed jumps (Ramezani and Zeng, 2007). The diffusion component is simply geometric Brownian motion with constant volatility. An obvious extension would be to allow for random changes in volatility over time, but previous work has shown that mean-reverting stochastic volatility effects are negligible for long-term investors (Ma and Forsyth, 2016), so we use the simpler formulation here. We fit the parameters of the jump diffusion model to 90 years of market data.

We develop strategies based on dynamic (multi-period) mean-variance optimality. In other words, we consider strategies which minimize the variance of real terminal wealth for a given specified expected value of real terminal wealth. This means we are concentrating on the risk of the *outcome*, rather than the risk of the *process* along the way. As an example of process risk, some would argue that we should be concerned with the volatility of the investment portfolio throughout the entire investment period. However, adding constraints on the local volatility will lead to sub-optimal results compared with fixing attention on the terminal wealth distribution. We contend that focusing on the long-term investment goal is appropriate for retirement savings. However, while we focus on outcome risk, we implicitly take process risk into account to some extent through constraints such as not allowing any use of leverage.

We focus on pre-commitment mean-variance strategies. As noted by Menoncin and Vigna (2017), pre-commitment strategies can be alternatively viewed as time consistent target-based policies. This contrasts with time consistent mean-variance strategies, which have the economically unreasonable property that the optimal strategy implies that the amount invested in risky assets does not depend on current wealth (Björk et al., 2014). Björk et al. suggest that this can be addressed by using a wealth dependent risk-aversion parameter. However, this in turn causes undesirable properties when realistic constraints (e.g. leverage and solvency) are imposed (Wang and Forsyth, 2011). Hence we focus on target-based utility functions, which are also pre-commitment mean-variance optimal. Investors saving for retirement are primarily interested in accumulating assets in order to fund a reasonable standard of living post-retirement, at minimal risk. Hence a (real) target-based final wealth strategy seems appropriate in this context. Note that target final wealth based objective functions are entirely different from equilibrium-based, instantaneous return constrained objective functions such as those considered by He and Jiang (2017).

We develop mean-variance optimal deterministic and adaptive strategies, and provide two types of extensive comparisons between them. First, we use a *synthetic market* that relies on Monte Carlo simulations which assume that the stock and bond indexes follow the models with constant parameters fit from the entire historical time series. Second, we compare the strategies using bootstrap resampling of the actual historical data (Politis and Romano, 1994; Cogneau and Zakalmouline,

2013; Dichtl et al., 2016). We emphasize that all strategies enforce realistic constraints, e.g. no short sales or leverage, no trading if insolvent, discrete rebalancing, etc. This is important because unconstrained dynamic mean-variance strategies may involve the use of highly levered portfolios (Lioui and Poncet, 2016).

Our main results are as follows:

- For a lump sum investment in the synthetic market with continuous rebalancing, a constant proportion strategy is superior in the mean-variance sense to any deterministic glide path.
- For a discretely rebalanced long-term portfolio with regular periodic contributions, the optimal deterministic strategy gives only a very slight improvement (under mean-variance criteria) over a constant proportion strategy.
- The risk-reward tradeoff given by the optimal deterministic strategy for a portfolio with regular contributions does not improve much if the portfolio is rebalanced more often than annually. This implies that infrequent rebalancing is not costly in terms of mean-variance criteria, while offering the benefits of lower trading costs.
- The optimal adaptive strategy typically reduces the standard deviation of the terminal wealth by a factor of about two compared to the optimal deterministic strategy having the same expected final wealth. The median terminal wealth for the adaptive strategy is always higher than the mean value. In contrast, in the deterministic strategy case the median terminal wealth is always lower than the mean. The probabilities of shortfall for a wide range of terminal wealth values are also substantially reduced for the adaptive strategy compared to the deterministic strategy. However, this comes at the cost of a lower probability of very high returns and a greater chance of extremely low returns. The low return case occurs when the equity market trends downward throughout a 30 year period, which most would regard as an unlikely scenario.

Our overall conclusion is that the current deterministic strategies used in most TDFs are sub-optimal relative to adaptive strategies. While it is unrealistic to assume that individual investors could determine optimal adaptive strategies themselves, it certainly is possible for sophisticated financial intermediaries to provide them to their clients.

2 Formulation

For simplicity we assume that there are only two assets available in the financial market, namely a risky asset and a risk-free asset. In practice, the risky asset would be a broad market index fund. The investment horizon is T . S_t and B_t respectively denote the *amounts* invested in the risky and risk-free assets at time t , $t \in [0, T]$. In general, these amounts will depend on the investor's strategy over time, including contributions, withdrawals, and portfolio rebalances, as well as changes in the unit prices of the assets. The investor can control all of these factors except for the unit prices. To clarify our assumptions regarding asset price dynamics, suppose for the moment that the investor does not take any action with respect to the controllable factors. We refer to this as the absence of control. It implies that all changes in S_t and B_t result from changes in asset prices. In this case, we assume that S_t follows a jump diffusion process. Let $t^- = t - \epsilon$, $\epsilon \rightarrow 0^+$, i.e. t^- is the instant of time before t , and let ξ be a random number representing a jump multiplier. When a jump occurs, $S_t = \xi S_{t^-}$. Allowing discontinuous jumps lets us explore the effects of severe market crashes on the risky asset holding. We assume that ξ follows a double exponential distribution (Kou, 2002; Kou

and Wang, 2004). If a jump occurs, p_{up} is the probability of an upward jump, while $1 - p_{up}$ is the chance of a downward jump. The density function for $y = \log(\xi)$ is

$$f(y) = p_{up}\eta_1 e^{-\eta_1 y} \mathbf{1}_{y \geq 0} + (1 - p_{up})\eta_2 e^{\eta_2 y} \mathbf{1}_{y < 0}. \quad (2.1)$$

For future reference, note that

$$\begin{aligned} E[y = \log \xi] &= \frac{p_{up}}{\eta_1} - \frac{(1 - p_{up})}{\eta_2}, \\ E[\xi] &= \frac{p_{up}\eta_1}{\eta_1 - 1} + \frac{(1 - p_{up})\eta_2}{\eta_2 + 1}, \\ E[(\xi - 1)^2] &= \frac{p_{up}\eta_1}{\eta_1 - 2} + \frac{(1 - p_{up})\eta_2}{\eta_2 + 2} - 2 \left(\frac{p_{up}\eta_1}{\eta_1 - 1} + \frac{(1 - p_{up})\eta_2}{\eta_2 + 1} \right) + 1. \end{aligned} \quad (2.2)$$

In the absence of control, S_t evolves according to

$$\frac{dS_t}{S_{t-}} = (\mu - \lambda E[\xi - 1]) dt + \sigma dZ + d \left(\sum_{i=1}^{\pi_t} (\xi_i - 1) \right), \quad (2.3)$$

where μ is the (uncompensated) drift rate, σ is the volatility, dZ is the increment of a Wiener process, π_t is a Poisson process with positive intensity parameter λ , and ξ_i are i.i.d. positive random variables having distribution (2.1). Moreover, ξ_i , π_t , and Z are assumed to all be mutually independent.

As an aid to carrying out algebraic manipulations, we can write (2.3) more informally as

$$\frac{dS_t}{S_{t-}} = (\mu - \lambda E[\xi - 1]) dt + \sigma dZ + (\xi - 1) dQ, \quad (2.4)$$

where $dQ = 1$ with probability λdt and $dQ = 0$ with probability $1 - \lambda dt$.

In the absence of control, we assume that the dynamics of the amount B_t invested in the risk-free asset are

$$dB_t = rB_t dt, \quad (2.5)$$

where r is the (constant) risk-free rate.

We define the investor's total wealth at time t as

$$\text{Total wealth} \equiv W_t = S_t + B_t. \quad (2.6)$$

Given a specified expected value of terminal wealth $E[W_T]$, the investor wants to minimize the risk of achieving this expected terminal wealth. We impose the constraints that shorting stock and using leverage (i.e. borrowing) are not permitted, which would be typical of a retirement savings account.

3 Deterministic Glide Paths

Let p denote the fraction of total wealth that is invested in the risky asset, i.e.

$$p = \frac{S_t}{S_t + B_t}. \quad (3.1)$$

A *deterministic* glide path restricts the admissible strategies to those with $p = p(t)$, i.e. the optimal strategy cannot take into account the actual value of W_t at any time. Clearly this is a very restrictive assumption, but it is commonly used in TDFs. Although a constant proportion strategy can be seen as a special case of a deterministic glide path where $p(t) = \text{const.}$, it is simpler here for expository reasons to reserve the label “deterministic glide path” for cases where $p(t)$ is time-varying.

3.1 Lump Sum Investment with Continuous Rebalancing

To gain some intuition about deterministic strategies, we consider first a simple case with a lump sum initial investment and no further cash injections or withdrawals. We also assume here that the portfolio is continuously rebalanced. Under these conditions, we can derive:

Proposition 3.1 (Inefficiency of glide path strategies for lump sum investments). *Consider a market with two assets following the processes (2.4) and (2.5). Suppose we invest a lump sum W_0 at $t = 0$ in a continuously rebalanced portfolio using a deterministic glide path strategy $p = p(t)$, where p is the fraction of total wealth invested in the risky asset. Also consider a strategy with a constant proportion p^* invested in the risky asset, where*

$$p^* = \frac{1}{T} \int_0^T p(s) ds. \quad (3.2)$$

Then:

- (i) the expected value of the terminal wealth is the same for both strategies; and
- (ii) the standard deviation of terminal wealth for the glide path strategy cannot be less than that of the constant proportion strategy.

Proof. Equations (2.4) and (2.5) imply

$$\frac{dW_t}{W_{t-}} = [p(t)(\mu - r) + r] dt - \lambda p(t)E[\xi - 1] dt + p(t)\sigma dZ + p(t)(\xi - 1) dQ. \quad (3.3)$$

Letting $\bar{W}_t = E[W_t]$ and noting that $p(t)$ is deterministic, we have

$$d\bar{W}_t = [p(t)(\mu - r) + r] \bar{W}_t dt \quad (3.4)$$

and

$$\bar{W}_T = E[W_T] = W_0 e^{[p^*(\mu - r) + r]T}, \quad (3.5)$$

where p^* is defined in equation (3.2). Write equation (3.3) as

$$\frac{dW_t}{W_{t-}} = \hat{\mu} dt + p(t)\sigma dZ + p(t)(\xi - 1) dQ, \quad (3.6)$$

where $\hat{\mu} = [p(t)(\mu - r) + r] - \lambda p(t)E[\xi - 1]$. Let $G_t = W_t^2$. From equation (3.6) and Itô's Lemma for jump processes,

$$\frac{dG_t}{G_{t-}} = \left[2\hat{\mu} + (p(t)\sigma)^2 \right] dt + 2p(t)\sigma dZ + [p(t)^2(\xi - 1)^2 + 2p(t)(\xi - 1)] dQ. \quad (3.7)$$

Let $\bar{G}_t = E[G_t] = E[W_t^2]$. Equation (3.7) and the fact that $p(t)$ is deterministic imply

$$\begin{aligned} \frac{d\bar{G}_t}{\bar{G}_t} &= \left[2\hat{\mu} + (p(t)\sigma)^2 \right] dt + \left(\lambda p(t)^2 E[(\xi - 1)^2] + 2\lambda p(t)E[(\xi - 1)] \right) dt \\ &= (2[p(t)(\mu - r) + r] + p(t)^2\sigma_e^2) dt \end{aligned} \quad (3.8)$$

where $\sigma_e^2 = \sigma^2 + \lambda E[(\xi - 1)^2]$. This in turn gives

$$\bar{G}_T = G_0 \exp\left(2[p^*(\mu - r) + r]T + \sigma_e^2 \int_0^T p(s)^2 ds\right), \quad (3.9)$$

or

$$E[W_T^2] = (E[W_T])^2 \exp\left[\sigma_e^2 \int_0^T p(s)^2 ds\right]. \quad (3.10)$$

From $Var[W_T] = E[W_T^2] - (E[W_T])^2$, we obtain

$$std[W_T] = E[W_T] \left(\exp\left[\sigma_e^2 \int_0^T p(s)^2 ds\right] - 1 \right)^{1/2} \quad (3.11)$$

where $std[\cdot]$ denotes standard deviation. By the Cauchy-Schwartz inequality

$$(p^*)^2 T \leq \int_0^T p(s)^2 ds, \quad (3.12)$$

and Proposition 3.1 follows immediately. \square

This proposition suggests that deterministic glide path strategies may have been oversold. A similar result for the geometric Brownian motion case (i.e. no jumps) was noted by Graf (2017). Furthermore, several authors have suggested that deterministic glide path strategies do not appear to offer many advantages based on Monte Carlo and historical simulations. For example, Poterba et al. (2009) simulate scenarios involving periodic contributions based on a sample of household earnings trajectories and investment returns based on resampled annual returns. They find that allocating wealth to assets based on age does not outperform a simple constant proportion strategy, noting that

The similarity of the retirement wealth distributions from the life-cycle portfolios, and from strategies that allocate a constant portfolio share to equities, is one of the central findings of our analysis. This result calls for further work to evaluate the extent to which life-cycle strategies offer unique opportunities for risk reduction relative to simpler strategies that allocate a constant fraction of portfolio assets to equities at all ages. (Poterba et al., 2009, p. 38)

Basu et al. (2011) and Esch and Michaud (2014) also find that glide paths do not seem to provide significant benefits in comparison to simpler fixed proportion strategies. Under some simplified assumptions, Proposition 3.1 shows that this result must hold: for any glide path, there is an equivalent constant weight strategy that offers the same expected final wealth at equal or lower risk. It is not surprising, then, to find that this is approximately correct in more complex and realistic simulations.

Along somewhat different lines, Arnott et al. (2013) simulate an inverse glide path which starts out with a low equity allocation that is increased over time. Their simulations show that this results in, if anything, better performance than the standard glide path which reduces equity exposure over time. Arnott et al. attribute this counterintuitive result to the effect of contributions on portfolio size over time. The standard glide path is most heavily invested in equities early on when the portfolio is fairly small. It does not benefit as much in monetary terms from high equity returns as

the inverse glide path strategy, which has higher wealth when it is most exposed to equities. Basu et al. (2011) make a similar point, noting that the standard glide path approach can perform poorly because switching out of equities into bonds at a time when accumulated wealth (and possibly also contributions, if these are a fixed percentage of salary which has increased over time) is relatively large, “the investor may be foregoing the opportunity to earn higher returns on a larger sum of money invested” (Basu et al., 2011, p. 84). However, we point out that even in the case of a single lump sum contribution, the standard glide path intuition fails. Note that $\int_0^T p(s) ds = \int_0^T p(T-s) ds$ and $\int_0^T [p(s)]^2 ds = \int_0^T [p(T-s)]^2 ds$, so by equations (3.5) and (3.11) the glide path results are the same in this case if we reverse the strategy. In other words, if our glide path starts with a high allocation to stocks and finishes with a low allocation to stocks, we can achieve exactly the same mean-variance result in terms of final wealth by beginning with a low equity allocation and ending with a high equity allocation.

3.2 Discrete Rebalancing and Periodic Contributions

The results in Section 3.1 are useful for gaining some intuition about the performance of glide path strategies, but the assumptions of no cash injections and continuous rebalancing are unrealistic. We now consider the implications of periodic cash injections and discrete portfolio rebalancing.

Let the inception time of the investment be $t_0 = 0$. We consider a set \mathcal{T} of pre-determined *rebalancing times*,

$$\mathcal{T} \equiv \{t_0 = 0 < t_1 < \dots < t_M = T\}. \quad (3.13)$$

For simplicity, we specify \mathcal{T} to be equidistant with $t_i - t_{i-1} = \Delta t = T/M$, $i = 1, \dots, M$. At each rebalancing time t_i , $i = 0, 1, \dots, M-1$, the investor injects an amount of cash q_i into the portfolio and then rebalances the portfolio. At $t_M = T$, the portfolio is liquidated. Let $t_i^- = t_i - \epsilon$ ($\epsilon \rightarrow 0^+$) be the instant before rebalancing time t_i , and $t_i^+ = t_i + \epsilon$ be the instant after t_i . Let $p(t_i^+) = p_i$ be the fraction in the risky asset at t_i^+ . This fraction is deterministic, so we can find some simple recursive expressions for the mean and variance of terminal wealth at $t = t_M$.

Similarly, let $S_i^+ = S_{t_i^+}$, $S_i^- = S_{t_i^-}$, $B_i^+ = B_{t_i^+}$, and $B_i^- = B_{t_i^-}$. From equations (2.4) and (2.5) we obtain

$$\begin{aligned} E[S_{i+1}^-] &= E[S_i^+] \exp[\mu \Delta t] \\ E[B_{i+1}^-] &= E[B_i^+] \exp[r \Delta t]. \end{aligned} \quad (3.14)$$

Since $W_i^- = S_i^- + B_i^-$,

$$\begin{aligned} W_i^+ &= W_i^- + q_i = S_i^- + B_i^- + q_i \\ E[W_i^+] &= E[S_i^-] + E[B_i^-] + q_i. \end{aligned} \quad (3.15)$$

Then

$$\begin{aligned} S_i^+ &= p_i W_i^+ \\ B_i^+ &= (1 - p_i) W_i^+ \\ E[S_i^+] &= p_i E[W_i^+] \\ E[B_i^+] &= (1 - p_i) E[W_i^+], \end{aligned} \quad (3.16)$$

since p_i is deterministic. Define

$$\begin{aligned} \mathcal{G}_t &= S_t^2 \\ \mathcal{F}_t &= B_t^2 \\ \mathcal{H}_t &= S_t \cdot B_t. \end{aligned} \quad (3.17)$$

Following similar steps as used to obtain equation (3.9), we can see that

$$\begin{aligned} E [\mathcal{G}_{i+1}^-] &= E [\mathcal{G}_i^+] \exp[(2\mu + \sigma_e^2)\Delta t] \\ E [\mathcal{F}_{i+1}^-] &= E [\mathcal{F}_i^+] \exp[2r\Delta t] \\ E [\mathcal{H}_{i+1}^-] &= E [\mathcal{H}_i^+] \exp[(r + \mu)\Delta t]. \end{aligned} \quad (3.18)$$

Noting that

$$\begin{aligned} (W_i^+)^2 &= (S_i^- + B_i^- + q_i)^2 \\ (W_i^-)^2 &= (S_i^- + B_i^-)^2, \end{aligned} \quad (3.19)$$

we obtain

$$\begin{aligned} E [(W_i^+)^2] &= E [(W_i^-)^2] + q_i^2 + 2E [S_i^-] q_i + 2E [B_i^-] q_i \\ E [(W_i^-)^2] &= E [\mathcal{G}_i^-] + E [\mathcal{F}_i^-] + 2E [\mathcal{H}_i^-]. \end{aligned} \quad (3.20)$$

From equations (3.16), (3.17), and (3.19), we obtain (again noting that p_i is deterministic)

$$\begin{aligned} E [\mathcal{G}_i^+] &= p_i^2 E [(W_i^+)^2] \\ E [\mathcal{F}_i^+] &= (1 - p_i)^2 E [(W_i^+)^2] \\ E [\mathcal{H}_i^+] &= (1 - p_i)p_i E [(W_i^+)^2]. \end{aligned} \quad (3.21)$$

Given a deterministic glide path $\{p_0, \dots, p_{M-1}\}$, the mean and variance of terminal wealth can be easily computed using Algorithm 3.1.

For a given specified expected terminal wealth $E [W_M^-] = d$, the mean-variance optimization problem to determine the optimal glide path can be stated as

$$\begin{aligned} \min_{\{p_0, p_1, \dots, p_{M-1}\}} \quad & \text{Var} (W_M^-) = E [(W_M^-)^2] - d^2 \\ \text{subject to} \quad & \begin{cases} E [W_M^-] = d \\ E [W_M^-], E [(W_M^-)^2] \text{ given by Algorithm 3.1} \\ p_i = p_i(t_i^+); 0 \leq p_i \leq 1 \end{cases} \end{aligned} \quad (3.22)$$

Note that we impose no-shorting and no-borrowing constraints $0 \leq p_i \leq 1$, which would be typical in the context of an investor saving for retirement.

3.3 Numerical Solution for the Deterministic Strategy

The objective function for Problem (3.22) can be evaluated very rapidly using Algorithm 3.1, so we can solve for the optimal controls $\{p_0, p_1, \dots, p_{M-1}\}$ using a numerical optimization technique. We use a Sequential Quadratic Programming (SQP) algorithm (Nocedal and Wright, 2006). Problem (3.22) is not in standard convex optimization form, since the expected value equality constraint is a nonlinear function of the controls p_i . If an SQP algorithm converges, it will converge to a local minimum, and there is no guarantee of convergence to the global minimum. In our numerical tests, we check for possible convergence to local minima by carrying out 10,000 tests, each starting with a different random initial starting guess for the optimal controls $\{p_0, p_1, \dots, p_{M-1}\}$. In all cases reported here, the SQP algorithm converged to the same solution vector, to within the specified convergence tolerance. This obviously is not a guarantee of convergence to the global minimum, but it is strongly suggestive.

```

input:  $\{p_0, p_1, \dots, p_{M-1}\}$  {glide path};
        $\{q_0, q_1, \dots, q_{M-1}\}$  {contributions};
        $\{\mu, r, \sigma_e^2, \Delta t\}$  {parameters};
initialize:  $E[S_0^-] = E[B_0^-] = E[\mathcal{G}_0^-] = E[\mathcal{F}_0^-] = E[\mathcal{H}_0^-] = 0$ ;

for  $i = 0, 1, \dots, M - 1$  do {Timestep loop}
   $E[W_i^+] = E[S_i^-] + E[B_i^-] + q_i$ ;
   $E[(W_i^+)^2] = E[\mathcal{G}_i^-] + E[\mathcal{F}_i^-] + q_i^2 + 2E[\mathcal{H}_i^-] + 2E[S_i^-]q_i + 2E[B_i^-]q_i$ ;
   $E[S_i^+] = p_i E[W_i^+]$ ;  $E[B_i^+] = (1 - p_i)E[W_i^+]$ ;
   $E[\mathcal{G}_i^+] = p_i^2 E[(W_i^+)^2]$ ;  $E[\mathcal{F}_i^+] = (1 - p_i)^2 E[(W_i^+)^2]$ ;  $E[\mathcal{H}_i^+] = (1 - p_i)p_i E[(W_i^+)^2]$ ;
   $E[S_{i+1}^-] = E[S_i^+] e^{\mu\Delta t}$ ;  $E[B_{i+1}^-] = E[B_i^+] e^{r\Delta t}$ ;
   $E[\mathcal{G}_{i+1}^-] = E[\mathcal{G}_i^+] e^{[(2\mu + \sigma_e^2)\Delta t]}$ ;  $E[\mathcal{F}_{i+1}^-] = E[\mathcal{F}_i^+] e^{[2r\Delta t]}$ ;  $E[\mathcal{H}_{i+1}^-] = E[\mathcal{H}_i^+] e^{[(r + \mu)\Delta t]}$ 
end for {End Timestep loop}

{Determine mean and variance at  $t_M$ }
 $E[W_M^-] = E[S_M^-] + E[B_M^-]$ ;
 $E[(W_M^-)^2] = E[\mathcal{G}_M^-] + E[\mathcal{F}_M^-] + 2E[\mathcal{H}_M^-]$ ;
return mean =  $E[W_M^-]$ ; variance =  $E[(W_M^-)^2] - (E[W_M^-])^2$ ;

```

ALGORITHM 3.1: *An algorithm for determining the mean and variance of terminal wealth for a given deterministic discrete rebalancing strategy $\{p_0, p_1, \dots, p_{M-1}\}$ and a schedule of contributions $\{q_0, q_1, \dots, q_{M-1}\}$, assuming the stochastic processes (2.4) and (2.5).*

4 Adaptive Strategies

We now allow the admissible set of controls to depend on the state of the investment portfolio, i.e. $p_i = p_i(S_i^+, B_i^+, t_i^+)$. Since we find the optimal strategy amongst all strategies with constant wealth, this is equivalent to $p_i = p_i(W_i^+, t_i^+)$. We consider the realistic case with discrete rebalancing and periodic contributions.

In the case of adaptive strategies, in some circumstances it can be optimal to withdraw cash from the portfolio (Cui et al., 2014; Dang and Forsyth, 2016). We denote this optimal cash withdrawal as $c_i \equiv c(W_i^- + q_i, t_i)$. Since we only allow cash withdrawals, $c_i \geq 0$. The control at rebalancing time t_i now consists of the pair (p_i, c_i) , i.e. after withdrawing c_i from the portfolio, rebalance to fraction p_i .

The optimization problem can now be written as

$$\begin{aligned}
& \min_{\{(p_0, c_0), \dots, (p_{M-1}, c_{M-1})\}} \text{Var}(W_T) = E[W_T^2] - d^2 \\
& \text{subject to } \begin{cases} E[W_T = S_T + B_T] = d \\ (S_t, B_t) \text{ follow processes (2.4)-(2.5); } t \notin \mathcal{T} \\ W_i^+ = W_i^- + q_i - c_i; S_i^+ = p_i W_i^+; B_i^+ = W_i^+ - S_i^+; t \in \mathcal{T} \\ p_i = p_i(W_i^+, t_i); 0 \leq p_i \leq 1 \\ c_i = c_i(W_i^- + q_i, t_i); c_i \geq 0 \end{cases} \quad (4.1)
\end{aligned}$$

4.1 Embedding Approach

To solve Problem (4.1), we use the embedding result of Li and Ng (2000) and Zhou and Li (2000). Consider a control set $P = \{(p_0(W_0^+, t_0^+), c_0(W_0^- + q_0, t_0)), \dots, \}$. Informally, if P^* is an optimal control for Problem (4.1), then there exists a W^* such that P^* is also the optimal control for the problem

$$\begin{aligned} & \min_{\{(p_0, c_0), \dots, (p_{M-1}, c_{M-1})\}} E \left[(W^* - W_T)^2 \right] \\ & \text{subject to } \begin{cases} (S_t, B_t) \text{ follow processes (2.4)-(2.5); } t \notin \mathcal{T} \\ W_i^+ = W_i^- + q_i - c_i; S_i^+ = p_i W_i^+; B_i^+ = W_i^+ - S_i^+; t \in \mathcal{T} \\ p_i = p_i(W_i^+, t_i); 0 \leq p_i \leq 1 \\ c_i = c_i(W_i^- + q_i, t_i); c_i \geq 0 \end{cases} \end{aligned} \quad (4.2)$$

Problem (4.2) can be solved by dynamic programming. If Problem (4.1) is not convex, there may be solutions to problem (4.2) which are not solutions to Problem (4.1). However, these spurious solutions can easily be eliminated (Tse et al., 2014; Dang et al., 2016).

As noted above, it is optimal to withdraw cash from the portfolio under some conditions. This is easily seen in the context of problem (4.2). Let

$$Q_\ell = \sum_{j=\ell+1}^{j=M-1} e^{-r(t_j - t_\ell)} q_j \quad (4.3)$$

be the discounted future contributions as of time t_ℓ . If

$$(W_i^- + q_i) > W^* e^{-r(T-t_i)} - Q_i, \quad (4.4)$$

then the optimal strategy is to (i) withdraw cash $c_i = W_i^- + q_i - (W^* e^{-r(T-t_i)} - Q_i)$ from the portfolio; and (ii) invest the remainder $(W^* e^{-r(T-t_i)} - Q_i)$ in the risk-free asset. This is optimal in this case since $E \left[(W^* - W_T)^2 \right] = 0$, which is the minimum of Problem (4.2).

In the following, we will refer to any cash withdrawn from the portfolio as a *surplus* cash flow. For the sake of discussion, we assume that any surplus cash flow is invested in the risk-free asset, but does not contribute to the computation of the terminal mean and variance. Other possibilities are discussed in Dang and Forsyth (2016).

Since we do not impose any further constraints on the control set P , the solution of Problem (4.2) is the so-called *pre-commitment* solution, which is not classically time-consistent (Basak and Chabakauri, 2010). However, the time-consistent solution can be obtained from the pre-commitment solution by imposing a time constraint (Wang and Forsyth, 2011), so it is obvious that a time-consistent solution will generally be sub-optimal (in terms of the variance of final wealth) compared to the pre-commitment solution.

Remark 4.1 (Pre-committing to a target). *It is easy to see from Problem (4.2) that the pre-commitment aspect comes about from using a constant (real) value of W^* set at $t = 0$. However, if we imagine that this target is based on a desired salary replacement level in retirement, then this pre-commitment to a savings target would appear to be a sensible goal. In addition, the strategy that solves Problem (4.2) de-risks as the target wealth is approached (Wang and Forsyth, 2010), thereby following the investment maxim “if you have won the game, stop playing”.*

Remark 4.2 (Time-consistency). *Some authors object to the pre-commitment policy on the grounds that it is not time-consistent (e.g. Basak and Chabakauri, 2010). However, the pre-commitment policy can be found by solving Problem (4.2), which can be determined by dynamic programming. It is then immediately obvious that the pre-commitment policy is the optimal time-consistent strategy for the objective function in Problem (4.2). Under this policy, the investor’s current wealth is monitored and the strategy will depend on the distance of the accumulated wealth from the target.*

Björk et al. (2014) note that the time-consistent strategy for multi-period mean-variance optimality has the property that the amount invested in the risky asset is deterministic (i.e. not a function of W_t). Björk et al. argue that this is economically unreasonable, and suggest using a wealth-dependent risk-aversion parameter to ameliorate this difficulty. However, Wang and Forsyth (2011) show that using a wealth-dependent risk-aversion has strange effects. In particular, adding constraints to the strategy results in an efficient frontier which plots higher than the unconstrained efficient frontier. This seemingly contradictory result can be explained by noting that the time-consistent mean-variance policy, as defined in Björk et al. (2014), no longer produces the optimal efficient frontier (i.e. it does not maximize expected value for fixed variance). This fact has also been observed by Vigna (2017), who notes that: “in order to be time consistent in the consistent planning sense, the investor has to choose a different functional, in other words different preferences.” (Vigna, 2017, p. 21) Vigna also points out that the time-consistent mean-variance policy is equivalent to specifying a CARA utility function, at least in some special cases.

One way to think of the solution of Problem (4.2) is that we determine the optimal pre-commitment policy at $t = 0$ (which maximizes expected wealth for fixed variance), which is Problem (4.1). This sets the value of W^* for the equivalent Problem (4.2). For times $t > 0$, the optimal strategy is the time-consistent solution of Problem (4.2). Menoncin and Vigna (2017) suggest that reasoning about pre-commitment mean-variance in terms of the time-consistent target-based Problem (4.2) is useful for DC pension plan members.

In other words, both the pre-commitment and time-consistent mean-variance solutions are time-consistent policies for alternative utility functions. But the pre-commitment policy has the additional feature that it is the optimal policy for the original problem at $t = 0$. Our main focus in this work concerns the comparison of adaptive with deterministic strategies. In view of the above remarks, we concentrate on target-based adaptive strategies. However, we give a brief numerical comparison of pre-commitment and time-consistent strategies (with leverage constraints) in Section 8.

Other definitions of time-consistent mean-variance are also possible. For example, He and Jiang (2017) suggest using an expected value constraint at each instant in time based on current wealth and a desired growth rate. However, this objective does not attempt to hit a fixed target, in the sense that the target simply adjusts to current wealth. This type of objective function does not seem appropriate in the DC plan context.

4.2 Numerical Solution for the Adaptive Strategy

We formulate Problem (4.2) as the solution of a nonlinear Hamilton-Jacobi-Bellman (HJB) partial-integro differential equation (PIDE). We refer the reader to Dang and Forsyth (2014); Forsyth and Labahn (2017) for details concerning the numerical solution. Given an arbitrary value of W^* , we can solve Problem (4.2) for the optimal control, which we denote by $P^*(W^*)$.

However, we want to find the solution to Problem (4.1), which is expressed in terms of a specified expected value $E[W_T] = d$. To determine the value of W^* for Problem (4.2) which satisfies the constraint $E[W_T] = d$, we solve for the value of W^* such that

$$f(W^*) = E_{P=P(W^*)}[W_T] - d = 0. \tag{4.5}$$

We solve equation (4.5) by Newton iteration. Each evaluation $f(W^*)$ requires a PIDE solve. This can be done efficiently by determining an approximate value for W^* on a coarse grid, and then using this estimate as the initial guess for the Newton iteration on a sequence of finer grids. Typically, only one Newton iteration is required on the finest grid. Since we use dynamic programming to solve Problem (4.2), we are guaranteed to obtain the globally optimal solution.

5 Data and Parameters

The parameters of equations (2.4) and (2.5) are estimated using data from the Center for Research in Security Prices (CRSP) on a monthly basis over the 1926-2015 period.² Our base case tests use the CRSP 3-month Treasury bill (T-bill) index for the risk-free asset and the CRSP value-weighted total return index for the risky asset. This latter index includes all distributions for all domestic stocks trading on major U.S. exchanges. As an alternative case for additional illustrations, we replace the above two indexes by a 10-year Treasury index and the CRSP equal-weighted total return index.³ All of these various indexes are in nominal terms, so we adjust them for inflation by using the U.S. CPI index, also supplied by CRSP. We use real indexes since investors saving for retirement should be focused on real (not nominal) wealth goals.

Appendix A discusses the methods used to calibrate the model parameters to the historical data. We use both a threshold technique (Cont and Mancini, 2011) and maximum likelihood (ML) estimation. The threshold estimator requires a parameter α , described in Appendix A. Briefly, we identify a jump if the magnitude of the observed return in a month is greater than α standard deviations from the mean expected return assuming geometric Brownian motion. Given our data frequency, setting $\alpha = 3$ is a sensible choice (Forsyth and Vetzal, 2017). Annualized estimated parameters using both the threshold method with $\alpha = 3$ and ML for both the value-weighted and equal-weighted indexes are provided in Table 5.1. As might be expected due to the small firm effect, the equal-weighted index has slightly higher estimated diffusion parameters (μ and σ). It also has a higher estimated probability of an upward jump, and jumps that tend to be a little larger in magnitude. More importantly for our purposes, the ML parameter estimates imply much more frequent and smaller jumps on average for both indexes. From the perspective of a long-term investor, it is probably more appropriate to model infrequent larger jumps. Hence we have a preference for the threshold estimates, so we use them in the numerical examples below. We also note that Dang et al. (2017) and Forsyth and Vetzal (2017) conduct some tests using both ML and threshold techniques. A range of values for α are used to estimate the jump diffusion parameters. As one example, Forsyth and Vetzal (2017) compute the optimal adaptive strategy using ML estimates, and then apply this control in a synthetic market where the stochastic process follows parameters which are estimated by thresholding. The investment results are robust to this form of parameter mis-specification.

Table 5.2 shows the average real interest rates for the 3-month T-bill and 10-year U.S. Treasury indexes over the entire sample period from 1926 to 2015. The 10-year index earned an average return of about 130 basis points per year over the 3-month index during this time.

²More specifically, results presented here were calculated based on data from Historical Indexes, ©2015 Center for Research in Security Prices (CRSP), The University of Chicago Booth School of Business. Wharton Research Data Services was used in preparing this article. This service and the data available thereon constitute valuable intellectual property and trade secrets of WRDS and/or its third-party suppliers.

³The 10-year Treasury index was constructed from monthly returns from CRSP back to 1941. The data for 1926-1941 were interpolated from annual returns in Homer and Sylla (2005).

| Method | μ | σ | λ | p_{up} | η_1 | η_2 |
|--------------------------------|--------|----------|-----------|----------|----------|----------|
| Real CRSP Value-Weighted Index | | | | | | |
| ML | .08326 | .12611 | 3.0881 | 0.09963 | 10.837 | 18.913 |
| threshold ($\alpha = 3$) | .08889 | .14771 | .32222 | 0.27586 | 4.4273 | 5.2613 |
| Real CRSP Equal-Weighted Index | | | | | | |
| ML | .10735 | .14256 | 2.8166 | .14407 | 8.3486 | 14.963 |
| threshold ($\alpha = 3$) | .11833 | .16633 | .40000 | .33334 | 3.6912 | 4.5409 |

TABLE 5.1: *Estimated annualized parameters for double exponential jump diffusion model. Value-weighted and equal-weighted CRSP indexes, deflated by the CPI. Sample period 1926:1 to 2015:12.*

| | Real 3-month T-bill Index | Real 10-year Treasury Index |
|-------------|---------------------------|-----------------------------|
| Mean return | .00827 | .02160 |

TABLE 5.2: *Mean annualized real rates of return for bond indexes ($\log[B(T)/B(0)]/T$). Sample period 1926:1 to 2015:12.*

6 Numerical Examples

6.1 Base Case: CRSP Value-Weighted Index and 3-month T-bill Index

As a first example, we consider the base case input data summarized in Table 6.1. An investor with a horizon of 30 years makes real contributions each year of \$10, allocated between the CRSP value-weighted and 3-month T-bill indexes and rebalanced annually.

6.1.1 Synthetic Market - Base Case

We refer to a market where the underlying stock and bond indexes follow processes (2.4) and (2.5), with fixed parameters given in Tables 5.1 and 5.2, as a *synthetic market*. In other words, this is a market based on the historical (constant) estimated parameters. We are careful to distinguish tests in a synthetic market with tests that use actual historical returns (*bootstrap resampling*), as discussed below in Section 6.1.2.

We first use a constant proportion strategy ($p = 0.5$) and determine the expected value of the terminal real wealth for this strategy. We then use this expected value as a constraint and determine the optimal deterministic strategy, which is the solution of problem (3.22). Finally, we use the same expected value as a constraint and solve for the optimal adaptive strategy (4.1), by using the embedded formulation (4.2).

We evaluate the performance of the various strategies using Monte Carlo simulation in the synthetic market. This case constitutes the best possible context for both the optimal deterministic and the optimal adaptive strategies since the associated control parameters are based on perfect knowledge of the stochastic properties of the market.

Table 6.2 compares the results for the constant proportion, optimal deterministic, and optimal adaptive strategies. By design, all three strategies have the same expected real terminal wealth. The optimal deterministic standard deviation is about 0.98 times that of the constant proportion strategy, so the optimal deterministic strategy offers little improvement over a simpler constant

| | Base Case | Alternative Case |
|--------------------------------|----------------|------------------|
| Investment horizon (years) | 30 | 30 |
| Equity market index | Value-weighted | Equal-weighted |
| Risk-free asset index | 3-month T-bill | 10-year Treasury |
| Initial investment W_0 (\$) | 0.0 | 0.0 |
| Real investment each year (\$) | 10.0 | 10.0 |
| Rebalancing interval (years) | 1 | 1 |

TABLE 6.1: *Input data for examples. Cash is invested at $t = 0, 1, \dots, 29$ years. Market parameters are provided in Tables 5.1 and 5.2.*

| Strategy | $E[W_T]$ | $Median[W_T]$ | $std[W_T]$ | Probability of Shortfall | |
|-----------------------------------|----------|---------------|------------|--------------------------|-------------|
| | | | | $W_T < 500$ | $W_T < 600$ |
| Constant proportion ($p = 0.5$) | 705.6 | 628 | 349 | .28 | .45 |
| Optimal deterministic | 705.6 | 630 | 341 | .27 | .45 |
| Optimal adaptive | 705.6 | 776 | 154 | .12 | .17 |

TABLE 6.2: *Synthetic market results from 160,000 Monte Carlo simulation runs for base case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. The expected surplus cash flow for the optimal adaptive strategy is 16.7, assumed to be invested in the risk-free asset.*

proportion strategy with the same expected terminal wealth. In stark contrast, the optimal adaptive policy standard deviation is about 0.44 times that of the constant proportion strategy. The probabilities of shortfall shown in the table are reduced by a factor of more than two for the optimal adaptive strategy compared to the constant weight strategy. In addition, the $Median[W_T]$ for the adaptive strategy is more than 20% higher than the $Median[W_T]$ for the deterministic strategies.

Recall that Proposition 3.1 shows that a constant proportion strategy dominates the optimal deterministic glide path by mean-variance criteria, assuming that the portfolio is continuously rebalanced and that there is a lump sum initial investment. That result clearly does not hold in current context with annual rebalancing and contributions. However, the results from Table 6.2 are not very encouraging for the optimal deterministic strategy as it gives just very slight improvement over the simpler constant proportion alternative. Moreover, this is in a context that is tailor made for the deterministic strategy because the market simulations here use parameters and stochastic processes that exactly match those assumed when determining the optimal controls.

The intuition underlying the marginal improvement of the optimal deterministic strategy compared to the constant proportion strategy is as follows. As the time in the strategy becomes large, the marginal amount contributed is small compared to the accumulated wealth (on average), hence the optimal strategy tends to a constant proportion (i.e. this begins to resemble the lump sum case, and we know from Proposition 3.1 that a constant proportion strategy will be superior to any glide path in this case).

Figure 6.1 shows the optimal controls for both the deterministic and adaptive strategies. As a comparison, we show the deterministic control for $T = 15, 30, 50$ years in Figure 6.1(a). In each case, $E[W_T]$ is set to the expected final wealth for the constant proportion $p = 0.5$ case. Note that $p(t) \rightarrow 0.5$ as (T, t) increase, consistent with the intuition given above.. In the adaptive case, the

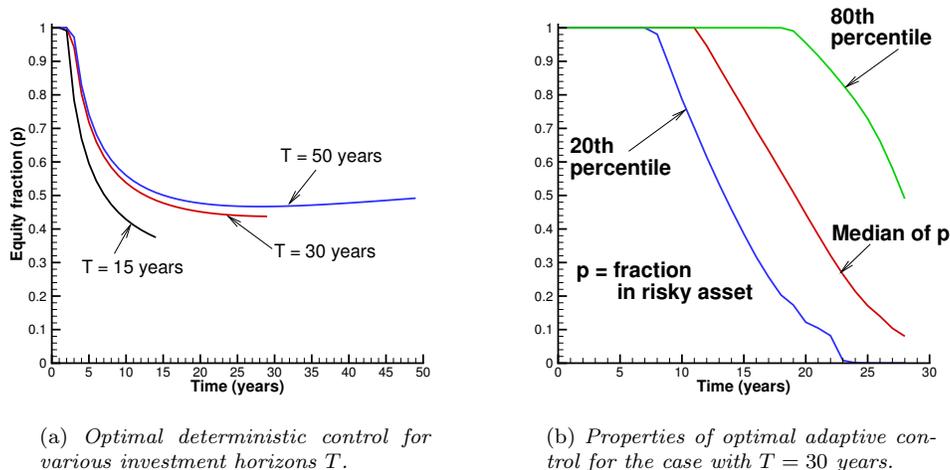


FIGURE 6.1: *Properties of optimal strategies using base case input data from Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. In each case, $E[W_T]$ is constrained to match that of a constant proportion strategy with $p = 0.5$. Figure 6.1(b) is based on 160,000 Monte Carlo simulation runs.*

control is a function of the current wealth. For ease of illustration, we show the median and the 20th and 80th percentiles of $p(W_t, t)$ for the case with $T = 30$ years in Figure 6.1(b), which we compute by Monte Carlo simulation. Although the median value of p corresponds in a general way to the standard glide path (starting with a high equity allocation and declining as the investment horizon is approached), the wide range of values between the two percentiles shown for values of $t > 10$ years shows that the optimal adaptive strategy depends significantly on accumulated wealth.

Figure 6.2 plots the cumulative distribution functions for the three strategies. The constant proportion and optimal deterministic strategies are virtually indistinguishable, reinforcing the conclusion that deterministic strategies offer at best slight benefits over simpler constant weight alternatives. The optimal adaptive strategy sacrifices large possible gains ($W_T > 800$) in order to reduce probability of shortfall over a wide range of terminal wealth values $360 < W_T < 800$. However, for low values of W_T , the deterministic strategy has smaller shortfall probability. A standard metric for measuring tail risk is the 95% conditional tail expectation (CTE), which is the mean of the worst 5% of the outcomes. The 95% CTE is 306 for the deterministic strategy, compared with 240 for the optimal adaptive strategy.

6.1.2 Resampled Historical Data - Base Case

Although it is useful to examine strategies for synthetic markets with parameters obtained from historical data, it is perhaps more convincing to see how the various strategies would have performed on actual historical data. We use bootstrap resampling to study this.

A single bootstrap resampled path is constructed as follows. Suppose the investment horizon is T years. We divide this total time into k blocks of size b years, so that $T = kb$. We then select k blocks at random (with replacement) from the historical data (from both the deflated stock and bond indexes). Each block starts at a random month. We then form a single path by concatenating these blocks. Since we sample with replacement, the blocks can overlap. To avoid end effects, the historical data is wrapped around, as in the circular block bootstrap (Politis and White, 2004; Patton et al., 2009). We repeat this procedure for many paths. The sampling is done in blocks in

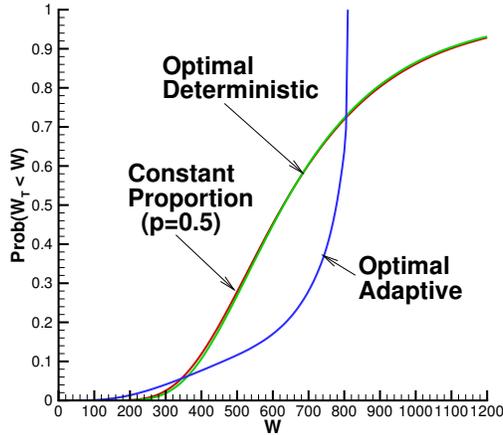


FIGURE 6.2: Cumulative distribution functions using base case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Distributions are computed using 160,000 Monte Carlo simulation runs in a synthetic market. Surplus cash flow is excluded from the distribution functions. $E[W_T] = 705.6$ for all strategies.

| Data series | Optimal expected block size \hat{b} (months) |
|--------------------------------|--|
| Real 3-month T-bill index | 50.1 |
| Real 10-year Treasury index | 4.7 |
| Real CRSP value-weighted index | 1.8 |
| Real CRSP equal-weighted index | 10.4 |

TABLE 6.3: Optimal expected blocksize $\hat{b} = 1/v$ when the blocksize follows a geometric distribution $Pr(b = k) = (1 - v)^{k-1}v$. The algorithm in Patton et al. (2009) is used to determine \hat{b} .

order to account for possible serial dependence effects in the historical time series. The choice of blocksize is crucial and can have a large impact on the results (Cogneau and Zakalmouline, 2013). We simultaneously sample the real stock and bond returns from the historical data. This introduces random real interest rates in our samples, in contrast to the constant interest rates assumed in the synthetic market tests and in the determination of the optimal controls.

To reduce the impact of a fixed blocksize and to mitigate the edge effects at each block end, we use the stationary block bootstrap (Politis and White, 2004; Patton et al., 2009). The blocksize is randomly sampled from a geometric distribution with an expected blocksize \hat{b} . The optimal choice for \hat{b} is determined using the algorithm described in Patton et al. (2009). This approach has also been used in other tests of portfolio allocation problems recently (e.g. Dichtl et al., 2016). Calculated optimal values for \hat{b} for the various indexes are given in Table 6.3.

We compute and store the optimal strategies (deterministic and adaptive) for the base case input data from Table 6.1 and the corresponding market parameters from Tables 5.1 (threshold) and 5.2. All strategies are constrained to have $E[W_T] = 705.6$ (in the synthetic market). We then apply these strategies using bootstrap resampling, based on the historical monthly data from January 1926 to December 2015. Of course, the resampled means will not be precisely the same and equal to 705.6

| Strategy | $E[W_T]$ | Median $[W_T]$ | std $[W_T]$ | Probability of Shortfall | |
|---|----------|----------------|-------------|--------------------------|-------------|
| | | | | $W_T < 500$ | $W_T < 600$ |
| Expected Blocksize $\hat{b} = 0.25$ years | | | | | |
| Constant proportion ($p = .5$) | 677 | 621 | 276 | .27 | .46 |
| Optimal deterministic | 676 | 623 | 268 | .27 | .46 |
| Optimal adaptive | 698 | 761 | 146 | .11 | .17 |
| Expected Blocksize $\hat{b} = 0.5$ years | | | | | |
| Constant proportion ($p = .5$) | 680 | 627 | 278 | .28 | .46 |
| Optimal deterministic | 679 | 624 | 272 | .28 | .46 |
| Optimal adaptive | 695 | 758 | 147 | .12 | .18 |
| Expected Blocksize $\hat{b} = 1.0$ years | | | | | |
| Constant proportion ($p = .5$) | 680 | 626 | 278 | .28 | .45 |
| Optimal deterministic | 679 | 625 | 270 | .27 | .45 |
| Optimal adaptive | 695 | 757 | 146 | .12 | .18 |
| Expected Blocksize $\hat{b} = 2.0$ years | | | | | |
| Constant proportion ($p = .5$) | 677 | 628 | 264 | .27 | .46 |
| Optimal deterministic | 676 | 625 | 257 | .26 | .45 |
| Optimal adaptive | 700 | 757 | 137 | .10 | .17 |
| Expected Blocksize $\hat{b} = 5.0$ years | | | | | |
| Constant proportion ($p = .5$) | 675 | 636 | 250 | .27 | .44 |
| Optimal deterministic | 674 | 635 | 246 | .26 | .44 |
| Optimal adaptive | 708 | 766 | 130 | .09 | .16 |

TABLE 6.4: *Stationary moving block bootstrap resampling results for base case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Strategies are rebalanced annually and are based on the synthetic market with $E[W_T] = 705.6$ in all cases. Calculations based on 10,000 bootstrap resamples of historical data for the period 1926:1 to 2015:12.*

for this test. The results for various block sizes are shown in Table 6.4. Choosing a block size that is too large will result in artificially low standard deviations. Table 6.4 indicates that the results are not too sensitive to expected block sizes in the range of 0.5 to 2 years. Generally, the results in Table 6.4 are quite comparable to those from the synthetic market reported in Table 6.2.

Figure 6.3 shows the cumulative distribution functions for the various strategies computed using bootstrap resampling of the actual historical data. Again, the cumulative distribution function for the optimal deterministic strategy is very close to that for the constant proportion strategy. If we include the surplus cash flow which is available for the adaptive strategy (assumed here to be invested in the risk-free asset), then there is some chance of obtaining $W_T > 800$. The surplus cash flow is a potential benefit for an investor following the adaptive strategy, so it makes sense to include it in the distribution function. Looking at Figures 6.2 and 6.3, we can see that the left tail risk of the adaptive strategy (relative to the optimal deterministic strategy) is somewhat reduced in the bootstrap simulations compared to the synthetic market tests. In this case, the 95% CTE for the optimal adaptive strategy is 279 compared to 316 for the optimal deterministic strategy (expected block size $\hat{b} = 2$ years). The median terminal wealth is more than 20% higher for the optimal adaptive strategy compared to the deterministic strategies for all block sizes.

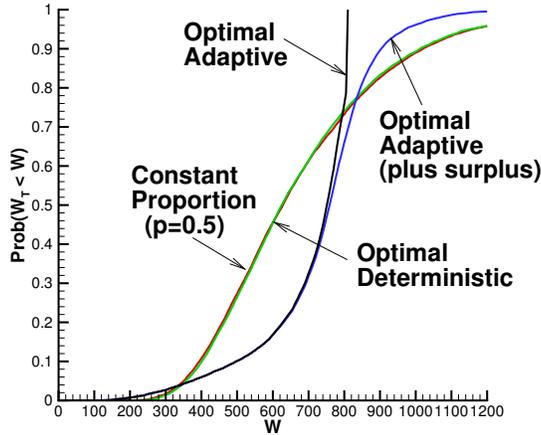


FIGURE 6.3: *Cumulative distribution functions using base case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Distributions are computed using 10,000 bootstrap resamples historical data from 1926:1 to 2015:12. Expected blocksize $\hat{b} = 2$ years.*

6.2 Alternative Case: CRSP Equal-Weighted Index and 10-year Treasury Index

To provide a second set of examples, we use alternative assets. In particular, as indicated in Table 6.1, we replace the CRSP value-weighted index with its equal-weighted counterpart, and we substitute the 10-year Treasury bond index for the 3-month Treasury bill index. See Tables 5.1 and 5.2 for relevant corresponding parameter estimates. We retain the same assumptions regarding investment horizon, rebalancing frequency, and real cash contributions as for the base case.

6.2.1 Synthetic Market - Alternative Case

Table 6.5 presents the results for the constant proportion, optimal deterministic, and optimal adaptive strategies. The results are very similar in qualitative terms to those seen earlier for the base case in Table 6.2, though investing in these two assets leads to a terminal wealth distribution with a higher mean and standard deviation relative to using the value-weighted index and 3-month T-bills. We continue to observe that the optimal deterministic strategy barely outperforms a simpler constant weight alternative, while the optimal adaptive strategy offers dramatically lower standard deviation and shortfall probabilities (except for the extreme left tail, as discussed shortly below). In this case, the median terminal wealth for optimal adaptive strategy exceeds the median terminal wealth for the deterministic strategy by more than 40%.

Figure 6.4(a) shows the optimal controls for the deterministic strategy. This is similar to the plot shown earlier in Figure 6.1(a) for the value-weighted index, but here we focus only on the case with $T = 30$ years. Again, over time the additional contributions tend to get small relative to the accumulated wealth, so the fraction invested in the equity index tends to a constant proportion. Figure 6.4(b) shows the median as well as the 20th and 80th percentiles of the optimal adaptive control $p(W_t, t)$. As with the value-weighted case shown above in Figure 6.1(b), there is a wide range between the 20th and 80th percentiles. Reflecting accumulated wealth from realized returns, the optimal adaptive strategy often departs from the median allocation after about the first decade.

Figure 6.5 compares the cumulative distribution functions for the three strategies. Once again,

| Strategy | $E[W_T]$ | Median $[W_T]$ | std $[W_T]$ | Probability of Shortfall | |
|-----------------------------------|----------|----------------|-------------|--------------------------|-------------|
| | | | | $W_T < 700$ | $W_T < 900$ |
| Constant proportion ($p = 0.5$) | 1085.2 | 874 | 860 | .33 | .52 |
| Optimal deterministic | 1085.2 | 878 | 846 | .32 | .52 |
| Optimal adaptive | 1085.2 | 1243 | 342 | .17 | .23 |

TABLE 6.5: Synthetic market results from 160,000 Monte Carlo simulation runs for alternative case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. The expected surplus cash flow for the optimal adaptive strategy is 51, assumed to be invested in the risk-free asset.

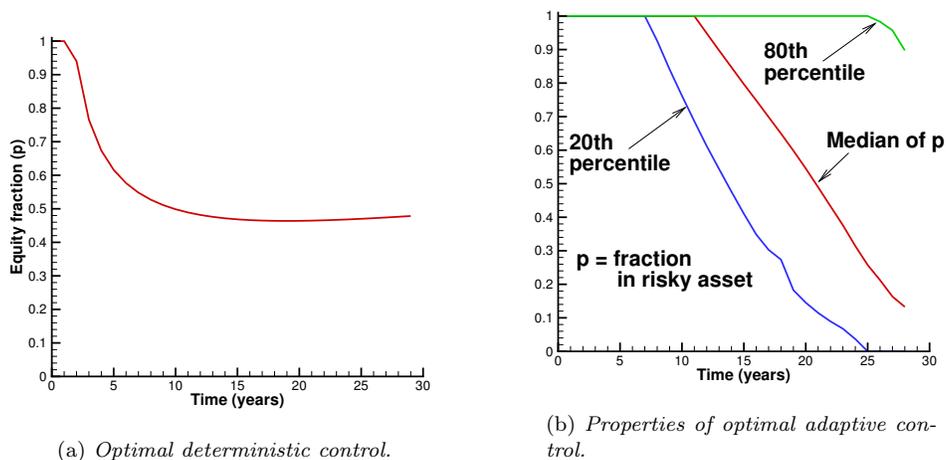


FIGURE 6.4: Properties of optimal strategies using alternative case input data from Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. In each case, $E[W_T]$ is constrained to match that of a constant proportion strategy with $p = 0.5$. Figure 6.4(b) is based on 160,000 Monte Carlo simulation runs.

the cumulative distribution for the optimal deterministic strategy is almost identical to that for the constant proportion strategy. The optimal adaptive strategy again sacrifices the extreme upside for protection against a wide downside range, but remains exposed to more left tail risk. In this case, the 95% CTE is 226 for the optimal adaptive strategy and 345 for the optimal deterministic strategy.

6.2.2 Resampled Historical Data - Alternative Case

We use similar bootstrap resampling procedures as described above in Section 6.1.2, but this time for the alternative case with the equal-weight equity and 10-year Treasury indexes. Table 6.6 shows the results for expected blocksizes ranging from 0.25 to 5.0 years. In all cases, the optimal adaptive strategy has higher average real terminal wealth with significantly lower standard deviation and shortfall probabilities for $W_T = 700$ and $W_T = 900$. In addition, the median final wealth is about 30% higher for the optimal adaptive strategy compared to the deterministic strategies for all blocksizes.

Figure 6.6 shows the cumulative distribution functions for the various strategies computed using

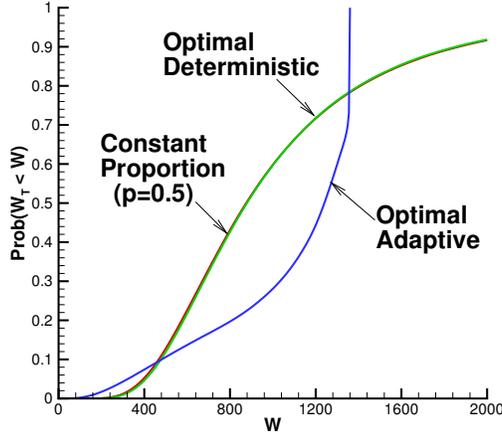


FIGURE 6.5: Cumulative distribution functions using alternative case input data from Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Distributions computed using 160,000 Monte Carlo simulation runs in a synthetic market. Surplus cash flow is excluded from the distribution functions. $E[W_T] = 1085.2$ for all strategies.

bootstrap resampling of the historical data. If we include the surplus cash flow, it appears that the adaptive strategy is almost first order stochastically dominant compared to the optimal deterministic strategy. The left tail risk measure (95% CTE) is 336 for the optimal adaptive strategy and 382 for the optimal deterministic case. The optimal deterministic strategy again offers at most a marginal advantage over the simpler constant proportion alternative.

7 Deterministic Strategy with Periodic Contributions and Continuous Rebalancing

It is interesting to determine the loss of efficiency in the deterministic case due to discrete rebalancing compared to continuous rebalancing. Of course, in practice trading costs can make high frequency rebalancing very expensive. We consider continuously rebalanced strategies, but with periodic contributions. As in Section 3.2, we specify contributions q_i at times $t_i, i = 0, \dots, M - 1$. There is no contribution at the terminal time $t_M = T$. We assume that the contributions are evenly spaced, so that $t_i - t_{i-1} = \Delta t$. Let $t_i^- = t_i - \epsilon, \epsilon \rightarrow 0^+$, and $t_i^+ = t_i + \epsilon$. Define the total wealth $W_t = S_t + B_t$, and let $\mathcal{G}_t = W_t^2$. Let

$$W_i^+ = W_{t_i^+}; W_i^- = W_{t_i^-}; \mathcal{G}_i^+ = \mathcal{G}_{t_i^+}; \mathcal{G}_i^- = \mathcal{G}_{t_i^-}. \quad (7.1)$$

At each contribution date t_i we have

$$W_i^+ = W_i^- + q_i; \mathcal{G}_i^+ = \mathcal{G}_i^- + 2q_i W_i^- + q_i^2, \quad (7.2)$$

so that

$$E[W_i^+] = E[W_i^-] + q_i; E[\mathcal{G}_i^+] = E[\mathcal{G}_i^-] + 2q_i E[W_i^-] + q_i^2. \quad (7.3)$$

| Strategy | $E[W_T]$ | $Median[W_T]$ | $std[W_T]$ | Probability of Shortfall | |
|---|----------|---------------|------------|--------------------------|-------------|
| | | | | $W_T < 700$ | $W_T < 900$ |
| Expected Blocksize $\hat{b} = 0.25$ years | | | | | |
| Constant proportion ($p = .5$) | 1015.4 | 863 | 615 | .34 | .53 |
| Optimal deterministic | 1013.6 | 865 | 602 | .33 | .53 |
| Optimal adaptive | 1043.9 | 1171 | 316 | .16 | .25 |
| Expected Blocksize $\hat{b} = 0.5$ years | | | | | |
| Constant proportion ($p = .5$) | 1005.4 | 868 | 585 | .33 | .53 |
| Optimal deterministic | 1004.0 | 869 | 582 | .33 | .53 |
| Optimal adaptive | 1040.8 | 1163 | 314 | .16 | .25 |
| Expected Blocksize $\hat{b} = 1.0$ years | | | | | |
| Constant proportion ($p = .5$) | 983.5 | 864 | 526 | .33 | .54 |
| Optimal deterministic | 981.8 | 863 | 516 | .32 | .54 |
| Optimal adaptive | 1046.3 | 1155 | 305 | .16 | .25 |
| Expected Blocksize $\hat{b} = 2.0$ years | | | | | |
| Constant proportion ($p = .5$) | 961.4 | 865 | 465 | .31 | .54 |
| Optimal deterministic | 958.7 | 861 | 457 | .31 | .54 |
| Optimal adaptive | 1064.0 | 1148 | 277 | .13 | .22 |
| Expected Blocksize $\hat{b} = 5.0$ years | | | | | |
| Constant proportion ($p = .5$) | 936 | 869 | 382 | .29 | .54 |
| Optimal deterministic | 936 | 866 | 380 | .29 | .54 |
| Optimal adaptive | 1089.5 | 1155 | 241 | .09 | .19 |

TABLE 6.6: Stationary moving block bootstrap resampling results for alternative case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Strategies are rebalanced annually and are based on the synthetic market with $E[W_T] = 1085.2$ in all cases. Calculations based on 10,000 bootstrap resamples of historical data for the period 1926:1 to 2015:12.

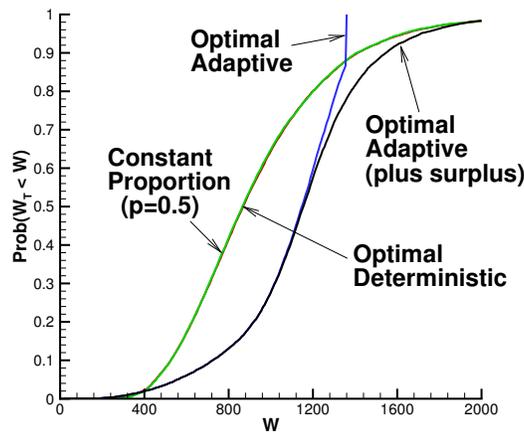


FIGURE 6.6: Cumulative distribution functions using alternative case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Distributions are computed using 10,000 bootstrap resamples historical data from 1926:1 to 2015:12. Expected blocksize $\hat{b} = 2$ years.

From the results in Section 3.1, it is easy to see that for a continuously rebalanced deterministic strategy with equity fraction $p(t)$

$$\begin{aligned} E [W_{i+1}^-] &= E [W_i^+] e^{(p_i^*(\mu-r)+r)\Delta t} \\ E [\mathcal{G}_{i+1}^-] &= E [\mathcal{G}_i^+] \exp \left[2(p_i^*(\mu-r) + r)\Delta t + \sigma_e^2 \int_{t_i}^{t_{i+1}} p(s)^2 ds \right], \end{aligned} \quad (7.4)$$

where

$$p_i^* = (1/\Delta t) \int_{t_i}^{t_{i+1}} p(s) ds. \quad (7.5)$$

Note that we can consider the continuously rebalanced strategy as the limit of a discretely rebalanced strategy, where we divide the contribution interval $(t_i, t_{i+1}]$ into sub-timesteps, and let the size of the sub-timesteps tend to zero. We allow different controls during each sub-timestep. Since the set of admissible controls for the limiting continuously rebalanced strategy is clearly larger than for the discretely rebalanced strategy, the variance of the continuously rebalanced strategy (for a fixed expected value) cannot exceed the variance of the discretely rebalanced strategy.

Before proceeding with our computations, the following result will be useful:

Proposition 7.1 (Optimal strategy: continuously rebalanced, deterministic case). *Consider a market with two assets following the processes (2.4) and (2.5), with periodic contributions at discrete times t_i . The mean-variance optimal continuously rebalanced deterministic strategy is to rebalance to a constant equity fraction between contribution times.*

Proof. Consider any strategy $p(t)$. Replace this strategy by the piecewise constant strategy

$$\hat{p}(t) = p_i^*; t \in (t_i, t_{i+1}] , \quad (7.6)$$

with p_i^* given in equation (7.5). Equations (7.4) now become

$$\begin{aligned} E [W_{i+1}^-]^* &= E [W_i^+]^* e^{(p_i^*(\mu-r)+r)\Delta t} \\ E [\mathcal{G}_{i+1}^-]^* &= E [\mathcal{G}_i^+]^* \exp \left[2(p_i^*(\mu-r) + r)\Delta t + \sigma_e^2 (p_i^*)^2 \Delta t \right], \end{aligned} \quad (7.7)$$

where $E[\cdot]^*$ indicates that the strategy (7.6) is used. This new strategy has the same expected value as the original strategy, so that $E [W_i^\pm]^* = E [W_i^\pm], \forall i$. From $\text{Var}[W_T] = E[W_T^2] - (E[W_T])^2$, we need only to show that $E [\mathcal{G}_M^-]^* \leq E [\mathcal{G}_M^-]$. Assume that $E [\mathcal{G}_i^+]^* \leq E [\mathcal{G}_i^+]$. From equations (3.12), (7.4), (7.5), and (7.7), we have $E [\mathcal{G}_{i+1}^-]^* \leq E [\mathcal{G}_{i+1}^-]$. From equation (7.3) and the fact that $E [W_i^-]^* = E [W_i^-]$, we have $E [\mathcal{G}_{i+1}^+]^* \leq E [\mathcal{G}_{i+1}^+]$. Finally, noting that $E [\mathcal{G}_0^+]^* = E [\mathcal{G}_0^+]$, the result follows. \square

From Proposition 7.1, we can use Algorithm 7.1 to calculate the mean and variance of terminal wealth for a given strategy $\{p_0^*, p_1^*, \dots, p_{M-1}^*\}$. The optimal continuously rebalanced strategy can be found by using Algorithm 7.1 and solving the optimization problem (3.22), using the methods described in Section 3.3. Table 7.1 compares the optimal mean-variance results for the deterministic strategies for both discretely and continuously rebalanced cases. As expected, the continuously rebalanced strategy is superior to the discretely rebalanced policy, but not by much. This has the practical implication that infrequent rebalancing does not reduce efficiency to a large degree, while reducing transaction costs.

```

input:  $\{p_0^*, p_1^*, \dots, p_{M-1}^*\}$  {glide path};
        $\{q_0, q_1, \dots, q_{M-1}\}$  {contributions};
        $\{\mu, r, \sigma_e^2, \Delta t\}$  {parameters};
initialize:  $E[W_0^-] = E[\mathcal{G}_0^-] = 0$ ;

for  $i = 0, 1, \dots, M - 1$  do {Timestep loop}
   $E[W_i^+] = E[W_i^-] + q_i$ ;
   $E[\mathcal{G}_i^+] = E[\mathcal{G}_i^-] + 2q_i E[W_i^-] + q_i^2$ ;
   $E[W_{i+1}^-] = E[W_i^+] e^{(p_i^*(\mu-r)+r)\Delta t}$ ;
   $E[\mathcal{G}_{i+1}^-] = E[\mathcal{G}_i^+] e^{(2\Delta t(p_i^*(\mu-r)+r)+(p_i^*)^2\sigma_e^2\Delta t)}$ ;
end for {End Timestep loop}

{Determine mean and variance at  $t_M$ }
return mean =  $E[W_M^-]$ ; variance =  $E[\mathcal{G}_M^-] - (E[W_M^-])^2$ ;

```

ALGORITHM 7.1: An algorithm for determining the mean and variance of a given deterministic continuously rebalanced strategy $\{p_0^*, p_1^*, \dots, p_{M-1}^*\}$ and a schedule of contributions $\{q_0, q_1, \dots, q_{M-1}\}$, assuming the stochastic processes (2.4) and (2.5).

| Strategy | Base Case | | Alternative Case | |
|------------------------------------|-----------|------------|------------------|------------|
| | $E[W_T]$ | $std[W_T]$ | $E[W_T]$ | $std[W_T]$ |
| Optimal deterministic (discrete) | 705.6 | 340.6 | 1085.2 | 846 |
| Constant proportion (continuous) | 705.6 | 337.6 | 1085.2 | 814 |
| Optimal deterministic (continuous) | 705.6 | 329.5 | 1085.2 | 802 |

TABLE 7.1: Comparison of discretely and continuously rebalanced strategies for input data given in Table 6.1 and corresponding parameters from Table 5.1 (threshold) and 5.2. In each case, $E[W_T]$ is set equal to that for a discretely rebalanced constant proportion strategy, as in Tables 6.2 and 6.5. The constant proportion (continuously rebalanced) weights which generate these expected values of terminal wealth are $p = .510$ (base case) and $p = .512$ (alternative case).

8 Comparison With Other Time-Consistent Strategies

Recall from Remark 4.2 that our optimal adaptive strategy is a pre-commitment policy which is also time-consistent for a quadratic utility function, i.e. it is the time-consistent solution to Problem (4.2). However, as noted earlier, other time-consistent strategies have also been suggested in the literature. In this section, we compare our optimal adaptive strategy with the time-consistent alternatives defined in Basak and Chabakauri (2010) and Björk et al. (2014).

The optimization problem for these alternatives can be written as

$$\begin{aligned}
& \max_{\{(p_0, c_0), \dots, (p_{M-1}, c_{M-1})\}} E[W_T] - \lambda^*(W_t, t) \text{Var}[W_T] \\
& \text{subject to } \begin{cases} (S_t, B_t) \text{ follow processes (2.4)-(2.5); } t \notin \mathcal{T} \\ W_i^+ = W_i^- + q_i; S_i^+ = p_i W_i^+; B_i^+ = W_i^+ - S_i^+; t \in \mathcal{T} \\ p_i = p_i(W_i^+, t_i); 0 \leq p_i \leq 1 \\ \text{time-consistent constraint, see Wang and Forsyth (2011)} \end{cases} \quad (8.1)
\end{aligned}$$

| Strategy | $E[W_T]$ | Median $[W_T]$ | std $[W_T]$ | Probability of Shortfall | |
|----------|----------|----------------|-------------|--------------------------|-------------|
| | | | | $W_T < 500$ | $W_T < 600$ |
| Case 1 | 705.6 | 688 | 234 | .17 | .33 |
| Case 2 | 705.6 | 592 | 477 | .38 | .51 |

TABLE 8.1: *Synthetic market results from 160,000 Monte Carlo simulation runs for base case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Compare with Table 6.2.*

No cash withdrawal is permitted for Problem (8.1). The time-consistent constraint is imposed using the method in (Wang and Forsyth, 2011), which is easily generalized for the time discrete control case. Due to the leverage constraints, closed form solutions of Problem (8.1) are unavailable. We use the algorithms in Wang and Forsyth (2011) and Forsyth and Labahn (2017) to solve Problem (8.1) numerically.

The approaches of Basak and Chabakauri (2010) and Björk et al. (2014) differ in terms of specifying $\lambda^*(W_t, t)$. Letting $\lambda_v > 0$ denote a constant, these two cases are given by

$$\lambda^*(W_t, t) = \begin{cases} \lambda_v & \text{Case 1} \\ \lambda_v/W_t & \text{Case 2} \end{cases}$$

Basak and Chabakauri (2010) suggest Case 1, while Björk et al. (2014) advocate Case 2.⁴

To compare with our previous findings, we determine λ_v to match the expected value of final wealth. The results are presented in Tables 8.1 and 8.2 for the base case and alternative case input data respectively. Consider first the base case results in Table 8.1, which should be compared with those reported earlier in Table 6.2. By the measures reported in these two tables, the results for the time-consistent Case 2 strategy are relatively poor. Although this strategy produces the same expected value of final wealth as for all of the other strategies in Tables 6.2 and 8.1, it offers the lowest median final wealth, the highest standard deviation of final wealth, and the highest shortfall probabilities for final wealth of 500 or 600. Moreover, although not reported in the tables, the 95% CTE for the time-consistent Case 2 strategy is just 177 (recall that the corresponding values reported above for the 240 for the optimal adaptive strategy and 306 for the optimal deterministic strategy). The time-consistent Case 1 strategy performs somewhat better, outperforming all of the other strategies except for the optimal adaptive strategy in terms of median final wealth, standard deviation of final wealth, and shortfall probabilities for $W_T = 500$ and 600. The 95% CTE for the Case 1 strategy is 305, slightly below that for the optimal deterministic strategy and somewhat higher than that for the optimal adaptive strategy.

Turning to the alternative case input data, we compare the results from Table 8.2 with those given above in Table 6.5. To be brief, the same general conclusions emerge here. The time-consistent Case 2 strategy gives the worst performance in terms of median, standard deviation, and shortfall probabilities. In addition, its 95% CTE of 171 is also the lowest of any strategy considered. The time-consistent Case 1 strategy gives somewhat better results, though it is still worse than the optimal adaptive strategy by any measure other than the 95% CTE.⁵

⁴Case 2 was investigated in Wang and Forsyth (2011), which resulted in the rather strange property that the efficient frontier with leverage constraints plotted higher than the efficient frontier without constraints.

⁵The 95% CTE for the time-consistent Case 1 strategy was 310, compared to 226 for the optimal adaptive strategy and 345 for the optimal deterministic strategy.

| Strategy | $E[W_T]$ | Median[W_T] | std[W_T] | Probability of Shortfall | |
|----------|----------|-----------------|--------------|--------------------------|-------------|
| | | | | $W_T < 700$ | $W_T < 900$ |
| Case 1 | 1085.2 | 1029 | 483 | .20 | .38 |
| Case 2 | 1085.2 | 781 | 1261 | .44 | .58 |

TABLE 8.2: *Synthetic market results from 160,000 Monte Carlo simulation runs for alternative case input data given in Table 6.1 and corresponding parameters from Tables 5.1 (threshold) and 5.2. Compare with Table 6.5.*

Overall, when all strategies are constrained to produce the same level of expected final wealth and have the same applicable leverage restrictions, the optimal adaptive strategy consistently offers the best performance by most measures (median final wealth, standard deviation of final wealth, and shortfall probability for moderate levels of final wealth). It does, however, underperform in the extreme left tail. The time-consistent Case 2 strategy gives the worst performance across the board. The time-consistent Case 1 strategy gives better left tail performance than the optimal adaptive strategy, but its performance is not as good as the optimal deterministic strategy by this criterion.⁶ Qualitatively similar results are obtained for the resampled historical data tests, which we omit for brevity.

9 Conclusion

We compare optimal deterministic strategies to simpler constant proportion alternatives, based on minimizing the variance of terminal wealth for fixed expected terminal wealth. We find that the best possible deterministic strategy gives at most very slight improvement over the simpler constant proportion strategy. Moreover, the efficiency of these strategies is not compromised in any significant way by relatively infrequent (i.e. annual) rebalancing, as opposed to being continuously rebalanced.

We also compare optimal deterministic strategies to optimal adaptive strategies, based on the same type of mean-variance criteria. Under both synthetic markets and bootstrap resampling of historical data, we observe that:

- The standard deviation of terminal wealth (for fixed mean wealth) is reduced by a factor $\simeq 2$ for the adaptive strategy compared to the optimal deterministic strategy.
- Over a wide range of terminal wealth values, the probability of shortfall for the adaptive strategy is much reduced compared to the deterministic strategy.
- The median value of the final real wealth for the adaptive strategy is 20 – 40% higher than the median values for the deterministic strategies.

In addition, we show that enforcing time-consistency leads to significantly reduced performance in terms of final wealth, except for the extreme left tail of the distribution. This is especially true for the case considered with non-constant risk-aversion.

However, there are some disadvantages for the adaptive strategies:

- There is a smaller probability of very large gains. This is to be expected from the form of the embedded mean-variance problem: we try to minimize the quadratic shortfall with respect to

⁶Further comparisons of pre-commitment and time-consistent mean variance solutions for a variety of problems with realistic constraints can be found in Wang and Forsyth (2012).

W^* , i.e. we sacrifice large gains in exchange for downside protection. We believe that this is a reasonable compromise for retirement saving.

- The 95% CTE level is smaller for the optimal adaptive control compared to the optimal deterministic strategy (i.e. there is larger left tail risk). An analysis of the cases which generate these poor results shows that this occurs for 30 year paths where the total return on equities is zero or negative. In this case, of course, there is some protection with the deterministic glide path, which moves into bonds as time goes on. In contrast, the adaptive strategy is fully invested in equities, since the accumulated wealth is always well below the target. This has historically been a good bet, but in the case of a 30 year stagnation in equities, it will certainly underperform.

Note that the 95% CTE for the adaptive strategy is higher for the bootstrap resampled simulations compared to the synthetic market (i.e. there is less left tail risk in the resamples). In the resampled case, long periods of low returns can occur if, for example, we repeatedly sample from the 1930s to form a 30 year chain of very poor returns for equities. In the synthetic market, with i.i.d. returns, such a sequence of low returns occurs with higher probability than in the historical data set. If we believe that long periods of very low returns for equities are unlikely, then adaptive strategies are well worth considering as an alternative to the ubiquitous deterministic strategies used in TDFs.

In short, over the past decade U.S. individuals have invested heavily in TDFs, which are now commonly offered as a default choice. This is a clear improvement over the situation around the turn of the century, where the default allocation was to a money market account. However, our results strongly suggest that TDFs themselves may be far from an optimal solution for investors saving for retirement.

Appendix

A Calibration of Model Parameters

In this Appendix, we discuss the estimation of the parameters of the jump diffusion process given by equations (2.1) and (2.3). Consider a discrete series of index prices $S(t_i) = S_i, i = 1, \dots, N + 1$ that are observed at equally spaced time intervals $\Delta t = t_{i+1} - t_i, \forall i$, with $T = N\Delta t$. We assume equal spacing for simplicity. Given log returns $\Delta X_i = \log(S_{i+1}/S_i)$, define detrended log returns as $\Delta \hat{X}_i = \Delta X_i - \hat{m} \Delta t$, where $\hat{m} = [\log(S_{N+1}) - \log(S_1)]/T$.

Figure A.1(a) shows a histogram of the monthly log returns from the real value-weighted CRSP total return index, scaled to zero mean and unit standard deviation. We superimpose a standard normal density onto this histogram. We also superimpose the fitted density for the double exponential jump diffusion model. The plot shows that the empirical data is leptokurtic, having a higher peak and fatter tails than a normal distribution, consistent with previous empirical findings for virtually all financial time series. Figure A.1(b) zooms in on these two densities, to better reveal the fat tails of the jump diffusion model.

A standard technique for parameter estimation is maximum likelihood (ML). However, it is well-known that the use of ML estimation for a jump diffusion model is problematic, due to multiple local maxima and the ill-posedness of trying to distinguish high frequency small jumps from diffusion (Honore, 1998). Alternative econometric techniques have been developed for detecting the presence of jumps in high frequency data, i.e. on a time scale of seconds (Aït-Sahalia and Jacod, 2012). However, from the perspective of a long-term investor, the most important feature of a jump diffusion

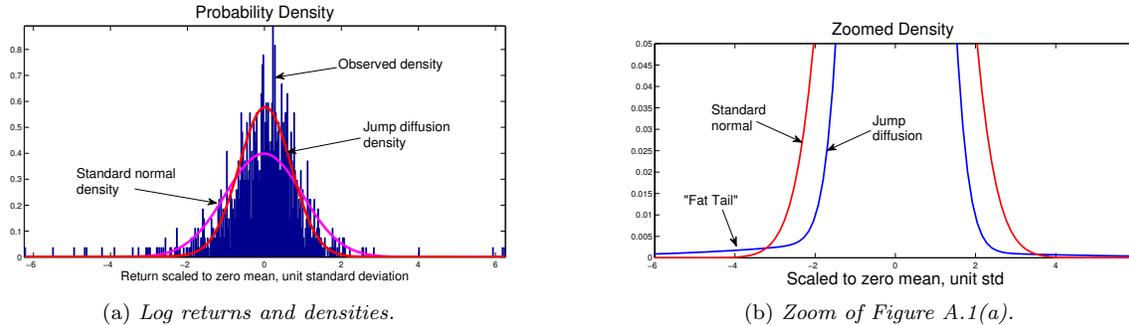


FIGURE A.1: Actual and fitted log returns for real CRSP value-weighted index. Monthly data, 1926:1-2015:12, scaled to unit standard deviation and zero mean. Standard normal density and fitted double exponential density (threshold, $\alpha = 3$) also shown.

model is that it allows modelling of infrequent large jumps in asset prices. Small and frequent jumps look like enhanced volatility when examined on a large scale, hence these effects are probably insignificant when constructing a long-term investment strategy. Consequently, as an alternative to ML estimation, we use the thresholding technique described in Mancini (2009) and Cont and Mancini (2011). This procedure is considered to be more efficient for low frequency data.

Suppose we have an estimate for the diffusive volatility component $\hat{\sigma}$. Then we detect a jump in period i if

$$\left| \Delta \hat{X}_i \right| > \mathcal{A} \hat{\sigma} \frac{\sqrt{\Delta t}}{(\Delta t)^\beta} \quad (\text{A.1})$$

where $\beta, \mathcal{A} > 0$ are tuning parameters (Shimizu, 2013), and $\hat{\sigma}$ is our most recent estimate of volatility. An iterative method is used to determine the parameters (Clewlow and Strickland, 2000). The intuition behind equation (A.1) is simple. If we choose $\mathcal{A} = 3$, say, and $\beta \ll 1$, then equation (A.1) identifies an observation as a jump if the observed log return exceeds a 3 standard deviation geometric Brownian motion change. Typically, β in equation (A.1) is quite small, $\beta \simeq .01 - .02$. For details, we refer the reader to Dang and Forsyth (2016). As described in Dang and Forsyth (2016), we replace $\mathcal{A}/(\Delta t)^\beta$ by the parameter α . Use of $\alpha = 3$ for monthly data results in fairly infrequent, large jumps. Additional details concerning the ML and threshold estimators can be found in Dang and Forsyth (2016) and Forsyth and Vetzal (2017).

Acknowledgements

Peter Forsyth acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC). Both authors acknowledge support from the University of Waterloo.

Declarations of Interest

The authors have no conflicts of interest to declare.

References

Aït-Sahalia, Y. and J. Jacod (2012). Analysing the spectrum of asset returns: jump and volatility of components of high frequency data. *Journal of Economic Literature* 50, 1007–1050.

- Arnott, R. D., K. F. Sherrerd, and L. Wu (2013). The glidepath illusion and potential solutions. *The Journal of Retirement* 1(2), 13–28.
- Basak, S. and G. Chabakauri (2010). Dynamic mean-variance asset allocation. *Review of Financial Studies* 23, 2970–3016.
- Basu, A. K., A. Byrne, and M. E. Drew (2011). Dynamic lifecycle strategies for target date retirement funds. *Journal of Portfolio Management* 37(2), 83–96.
- Benartzi, S. and R. H. Thaler (2001). Naive diversification strategies in defined contribution savings plans. *American Economic Review* 91, 79–98.
- Benartzi, S. and R. H. Thaler (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives* 21(3), 81–104.
- Benartzi, S. and R. H. Thaler (2013). Behavioral economics and the retirement savings crisis. *Science* 339(6124), 1152–1153.
- Björk, T., A. Murgoci, and X. Y. Zhou (2014). Mean variance portfolio optimization with state dependent risk aversion. *Mathematical Finance* 24, 1–24.
- Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2004). For better or for worse: default effects and 401(k) savings behavior. In D. A. Wise (Ed.), *Perspectives in the Economics of Aging*, pp. 81–121. Chicago: University of Chicago Press.
- Clewlow, L. and C. Strickland (2000). *Energy Derivatives: Pricing and Risk Management*. London: Lacima Group.
- Cogneau, P. and V. Zakalmouline (2013). Block bootstrap methods and the choice of stocks for the long run. *Quantitative Finance* 13, 1443–1457.
- Cont, R. and C. Mancini (2011). Nonparametric tests for pathwise properties of semimartingales. *Bernoulli* 17, 781–813.
- Cui, X., J. Gao, X. Li, and D. Li (2014). Optimal multi-period mean variance policy under no-shorting constraint. *European Journal of Operational Research* 234, 459–468.
- Dang, D.-M. and P. A. Forsyth (2014). Continuous time mean-variance optimal portfolio allocation under jump diffusion: a numerical impulse control approach. *Numerical Methods for Partial Differential Equations* 30, 664–698.
- Dang, D.-M. and P. A. Forsyth (2016). Better than pre-commitment mean-variance portfolio allocation strategies: a semi-self-financing Hamilton-Jacobi-Bellman equation approach. *European Journal of Operational Research* 250, 827–841.
- Dang, D.-M., P. A. Forsyth, and Y. Li (2016). Convergence of the embedded mean-variance optimal points with discrete sampling. *Numerische Mathematik* 132, 272–302.
- Dang, D.-M., P. A. Forsyth, and K. R. Vetzal (2017). The 4% strategy revisited: a pre-commitment optimal mean-variance approach to wealth management. *Quantitative Finance* 17, 335–351.
- Dichtl, H., W. Drobetz, and M. Wambach (2016). Testing rebalancing strategies for stock-bond portfolios across different asset allocations. *Applied Economics* 48, 772–788.

- Esch, D. N. and R. O. Michaud (2014). The false promise of target date funds. Working paper, New Frontier Advisors, LLC.
- Forsyth, P. and G. Labahn (2017). ϵ -Monotone Fourier methods for optimal stochastic control in finance. Working paper, Cheriton School of Computer Science, University of Waterloo.
- Forsyth, P. A. and K. R. Vetzal (2017). Dynamic mean variance asset allocation: Tests for robustness. *International Journal of Financial Engineering* 4, 1750021:1–1750021:37. DOI: 10.1142/S2424786317500219.
- Graf, S. (2017). Life-cycle funds: Much ado about nothing? *European Journal of Finance* 23, 974–998.
- He, X. and Z. Jiang (2017). Dynamic mean-risk asset allocation an myopic strategies: a universal portfolio rule. Working paper, Chinese University of Hong Kong.
- Homer, S. and R. Sylla (2005). *A History of Interest Rates*. New York: Wiley.
- Honore, P. (1998). Pitfalls in estimating jump diffusion models. Working paper, Center for Analytical Finance, University of Aarhus.
- ICI (2017). *Investment Company Institute 2017 Investment Company Fact Book*. Available at www.icifactbook.org.
- Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science* 48, 1086–1101.
- Kou, S. G. and H. Wang (2004). Option pricing under a double exponential jump diffusion model. *Management Science* 50, 1178–1192.
- Li, D. and W.-L. Ng (2000). Optimal dynamic portfolio selection: multiperiod mean-variance formulation. *Mathematical Finance* 10, 387–406.
- Lioui, A. and P. Poncet (2016). Understanding mean variance asset allocation. *European Journal of Operational Research* 254, 320–337.
- Ma, K. and P. A. Forsyth (2016). Numerical solution of the Hamilton-Jacobi-Bellman formulation for continuous time mean variance asset allocation under stochastic volatility. *Journal of Computational Finance* 20(1), 1–37.
- Madrian, B. C. and D. F. Shea (2001). The power of suggestion: inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics* 116, 1149–1525.
- Mancini, C. (2009). Non-parametric threshold estimation models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36, 270–296.
- Menoncin, F. and E. Vigna (2017). Mean-variance target based optimisation for defined contribution pension schemes in a stochastic framework. *Insurance: Mathematics and Economics* 76, 172–184.
- Nocedal, J. and S. Wright (2006). *Numerical Optimization*. New York: Springer Verlag. Springer Series in Operations Research.
- Patton, A., D. Politis, and H. White (2009). Correction to: automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 28, 372–375.

- Politis, D. and J. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.
- Politis, D. and H. White (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 23, 53–70.
- Poterba, J. M., J. Rauh, S. F. Venti, and D. A. Wise (2009). Life-cycle asset allocation strategies and the distribution of 401(k) retirement wealth. In D. A. Wise (Ed.), *Developments in the Economics of Aging*, pp. 15–50. Chicago: University of Chicago Press.
- Ramezani, C. A. and Y. Zeng (2007). Maximum likelihood estimation of the double exponential jump-diffusion process. *Annals of Finance* 3, 487–507.
- Shimizu, Y. (2013). Threshold estimation for stochastic differential equations with jumps. *Proceedings of the 59th ISI World Statistics Conference*, Hong Kong.
- Tse, S. T., P. A. Forsyth, and Y. Li (2014). Preservation of scalarization optimal points in the embedding technique for continuous time mean variance optimization. *SIAM Journal on Control and Optimization* 52, 1527–1546.
- Vanguard (2017). How America saves 2017. Available at <https://pressroom.vanguard.com/nonindexed/How-America-Saves-2017.pdf>.
- Vigna, E. (2017). Tail optimality and preferences consistency for intertemporal optimization problems. Working paper no. 502 , Collegio Carlo Alberto, Università Degli Studi di Torino.
- Wang, J. and P. A. Forsyth (2010). Numerical solution of the Hamilton-Jacobi-Bellman formulation for continuous time mean variance asset allocation. *Journal of Economic Dynamics and Control* 34, 207–230.
- Wang, J. and P. A. Forsyth (2011). Continuous time mean variance asset allocation: A time-consistent strategy. *European Journal of Operational Research* 209, 184–201.
- Wang, J. and P. A. Forsyth (2012). Comparison of mean variance like strategies for optimal asset allocation problems. *International Journal of Theoretical and Applied Finance* 15:2. DOI: 10.1142/S0219024912500148.
- Zhou, X. Y. and D. Li (2000). Continuous-time mean-variance portfolio selection: a stochastic LQ framework. *Applied Mathematics and Optimization* 42, 19–33.