Pre-Commitment Mean Variance: How Robust Is It?

Peter Forsyth[1]

[1]Cheriton School of Computer Science
University of Waterloo

Thursday November 17
2:00 PM - 4:00 PM
Capital Ballroom H
Austin

Motivation

Conventional Defined Benefit (DB) pension plans are disappearing.

Most employees now belong to Defined Contribution (DC) plans

- Employee and employer contribute yearly to tax advantaged account
- Employee chooses from list of approved investments
- Employee has to manage portfolio
    - No guarantee of value on retirement
    - Employee bears all the risk

Typical investment horizon

- $\simeq$ 30 years accumulation
- $\simeq$ 20 years decumulation (retirement)

DC investment cycle $\simeq$ 50 years

$\rightarrow$ DC plan members are truly long term investors

Basic Asset Allocation Strategy

The most important decision for a DC investor

- How much to allocate to bonds and how much to stocks[1]?

Benjamin Graham rule for the *defensive* investor [2]

- Invest a fraction p of portfolio in stocks, $(1-p)$ in bonds
- Rebalance your portfolio back to that fraction yearly

Target Date Fund (TDF)[3]

- Typical *glide path*

$$\text{Fraction invested in equities} = p(t) = \frac{110 - \text{ your age at } t}{100}.$$

- Intuition: take more risk when younger, less as retirement looms

---

[1] This would usually be an index ETF.

[2] B. Graham, "The Intelligent Investor". Graham suggests $p = 0.5$.

[3] Over \$750 billion invested in TDFs in the US, end of 2015

Asset allocation: risk free bond, stock index

Risk free bond B

$$dB_t = rB_t \, dt$$

$$r = \text{ risk-free rate}$$

Amount in risky stock index S (jump diffusion)

$$\frac{dS_t}{S_{t^-}} = (\mu - \rho\kappa) \, dt + \sigma \, dZ + (J-1) \, dq$$

$$\mu = \mathbb{P} \text{ measure drift} \quad ; \quad \sigma = \text{ volatility}$$

$$dZ = \text{ increment of a Wiener process}$$

$$dq = \begin{cases} 0 & \text{with probability } 1 - \rho \, dt \\ 1 & \text{with probability } \rho dt, \end{cases}$$

$$\log J \sim \text{ double exponential.} \quad ; \quad \kappa = E[J-1]$$

Optimal Control

Define:

$$\begin{aligned} W_t &= S_t + B_t = \text{total wealth} \\ p &= \frac{S_t}{W_t} = \text{fraction in equities} \Leftarrow \text{control} \end{aligned}$$

Three cases:

$$\begin{aligned} p = const. \quad &; \quad \text{constant proportion} \\ p = p(t) \quad &; \quad \text{deterministic (glide path)} \\ p = p(W_t, t) \quad &; \quad \text{adaptive (feedback)} \end{aligned}$$

Discrete rebalancing, contributions at times $t_i, i = 0, 1, \ldots$

- Assume here that $t_{i+1} - t_i = 1$ year
- $W_{t_{i+}} = W_{t_{i-}} + q_i$; $q_i =$ regular contribution.
- Control applied only at discrete times, $p(W_{t_i^+}, t_i) \equiv p_i$

Optimal Deterministic Mean Variance (Glide Path) Policy

Given in terms of terminal wealth at time $t = T$, W_T

$$\min_{\{p_0, p_1, \ldots, p_{T-1}\}} \quad \text{Var}(W_T) = E[(W_T)^2] - d^2$$

$$\text{subject to} \quad \begin{cases} E[W_T] = d \\ p_i = p_i(t_i) \quad \Leftarrow \text{ deterministic constraint} \\ 0 \le p_i \le 1 \end{cases} .$$

Closed form expression for $E[W_T], \text{Var}(W_T)$ as function of p_i

- Use standard optimization method to solve for optimal controls p_i

Optimal Adaptive Mean Variance Policy

$$\min_{\{p_0, p_1, \ldots, p_{M-1}\}} \quad \text{Var}(W_T) = E[W_T^2] - d^2$$

$$\text{subject to} \quad \begin{cases} E[W_T] = d \\ p_i = p_i(W_{t_i^+}, t_i) \Leftarrow \text{ Feedback form} \\ 0 \le p_i \le 1 \end{cases} .$$

Embedding approach: transform this problem to one which can be solved by Dynamic Programming (Zhou and Li(2000)), (Li and Ng (2000)).

Embedding Approach

For any control $\{p_0^*, p_1^*, \ldots, p_{M-1}^*\}$ which solves the original adaptive MV problem

$\exists$ a scaler W^* such that $\{p_0^*, p_1^*, \ldots, p_{M-1}^*\}$ solves

$$\begin{aligned} \min_{\{p_0, p_1, \ldots, p_{M-1}\}} \quad & E[(W^* - W_T)^2] \\ \text{subject to} \quad & \begin{cases} p_i = p_i(W_i, t_i) \\ 0 \leq p_i \leq 1 \end{cases} . \end{aligned} \tag{1}$$

Equation (1) is now amenable to solution by dynamic programming

Determine control by numerically solving an HJB equation (Dang and Forsyth (2014)) [4] [5]

---

[4] See D-M Dang later this afternoon

[5] $E[W_T] = d$ enforced by Newton iteration.

Pre-commitment policy

We are solving for the *pre-commitment* solution

$\rightarrow$ Not time consistent

Time inconsistency arises since W^* depends on initial state

But we can interpret this solution in the following way

- At $t = 0$, we determine a target W^* by examining the efficient frontier
- $\forall t > 0$, strategy is the time consistent solution of minimizing the expected quadratic loss w.r.t. fixed target W^* (Vigna (2014))

Setting a fixed target for real wealth W^* at $t = 0$ allows for optimal de-risking.

$\rightarrow$ Perhaps this is a good idea when saving for retirement

Optimal de-risking (surplus cash flow)

Let

$$
\begin{aligned}
F(t) &= W^* e^{-r(T-t)} \\
&= \text{discounted target wealth}
\end{aligned}
$$

Proposition 1 (Dang and Forsyth (2016))

If $W_t > F(t)$, $t \in [0, T]$, an optimal MV strategy is

- *Withdraw cash $(W_t - F(t))$ from the portfolio*
- *Invest the remaining amount $F(t)$ in the risk-free asset.*

We will refer to the amount withdrawn as a *surplus cash flow*. [6]

---

[6] See also: Ehrbar, *J. Econ. Theory* (1990); Cui, Li, Wang, Zhu *Mathematical Finance* (2012); Bauerle, Grether *Mathematical Methods of Operations Research* (2015).

Historical Data: 1926:1-2015:12

CRSP[7] monthly returns cap weighted stock index (US).

- Total return (including all dividends and distributions) for all stocks trading on all major exchanges, *1926:1 - 2015:12*

CRSP 3-month T-bills, and CPI index

Construct real (inflation adjusted) stock and bond indices from this data.

Fit double exponential jump diffusion model to the real stock index data

[7] Centre for Research in Securities Pricing

Fitting Data

Use threshold technique (parameter α) Cont and Mancini (2011). and Maximum Likelihood (MLE)

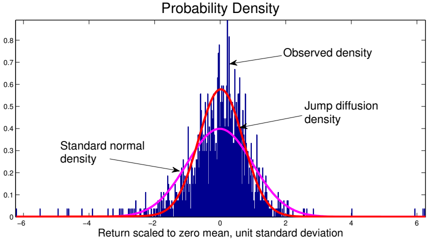


Figure : Probability density of log returns for real CRSP capitalization weighted index. Monthly data, *1926:1-2015:12*, scaled to unit standard deviation and zero mean. Standard normal density and fitted double exponential (threshold, $\alpha = 3$) also shown.

Threshold method: jump detected if scaled $|\log(\textit{return})| > \alpha$

MLE seems quite unstable (but we do it anyway)

Example DC Investor

Investment horizon	30 years
Market parameters	Threshold($\alpha = 3$)
Real risk-free rate r	0.00827 (average *1926:1-2015:12*)
Initial investment W_0	0.0
Real investment each year	10.0
Rebalancing interval	1 year

Table : Cash is injected at $t = 0, 1, \ldots, 29$ years.

Note: All results are for real (inflation adjusted) returns.

Numerical Results: Synthetic Market[9]

Determine optimal controls such that[8]

$$\overbrace{E[W_T]}^{constant\ \ proportion} = \overbrace{E[W_T]}^{optimal\ \ deterministic} = \overbrace{E[W_T]}^{optimal\ \ adaptive}$$

			Probability of Shortfall	
Strategy	$E[W_T]$	$std[W_T]$	$W_T < 500$	$W_T < 600$
Const. Prop. ($p = 0.5$)	706	**349**	.28	.45
Optimal Deterministic	706	**341**	.27	.45
Optimal Adaptive	706	**154**	**.12**	**.17**

$\Rightarrow$ Large reduction in probabilities of shortfall for adaptive strategy.

[8] For a lumpsum investment, we can prove that optimal deterministic strategy **is** the constant proportion strategy.

[9] By synthetic market, we mean the market which follows the SDEs with constant fitted parameters

Cumulative Distribution Function, optimal control [10]

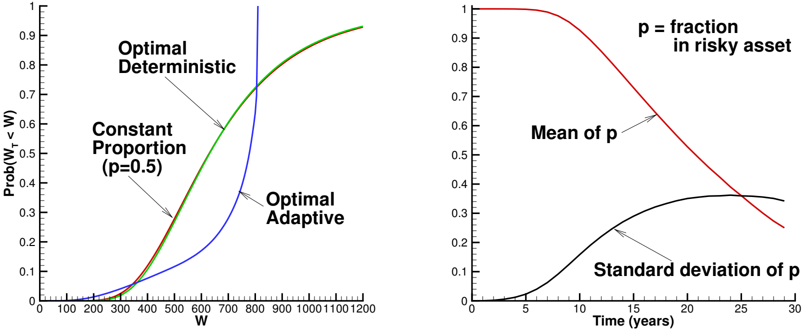


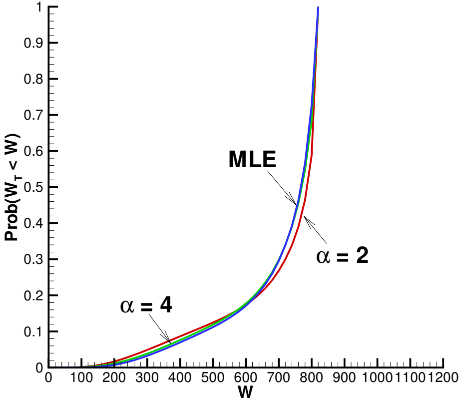
Figure : Left: cumulative distribution functions. Right: mean and standard deviation of optimal adaptive MV control.

[10] Optimal adaptive policy gives up large gains to reduce downside

Robustness to misspecified parameters (adaptive)

We compute the optimal adaptive control using the Threshold ($\alpha = 3$).

We then carry out MC simulations, but the MC simulations are carried out using different parameters ($\alpha = 2, 4$, MLE)



⇐ **Cumulative distribution function**

Optimal control computed using (threshold, $\alpha = 3$).

Monte Carlo simulations carried out using cases: threshold ($\alpha = 2$); threshold ($\alpha = 4$); MLE.

Controls computed using the wrong parameters

Robustness to misspecified parameters II

Compute optimal MV control using Threshold ($\alpha = 3$)

- Carry out MC simulations, use Threshold ($\alpha = 3$) parameters, **but** we lower the drift rate: $\mu_{MC} = \mu_{\alpha=3} - 200\ bps$
- i.e. we compute control using the wrong drift (too large)

			Probability of Shortfall	
Strategy	$E[W_T]$	$std[W_T]$	$W_T < 500$	$W_T < 600$
Const. Prop. ($p = 0.5$)	580	274	.46	.64
Optimal Deterministic	577	267	.46	.65
Optimal Adaptive	**632**	**197**	**.24**	**.32**

$\Rightarrow$ Performance degrades compared to good estimate of parameters, but still much better than deterministic strategies.

Bootstrap resampling

We have assumed i.i.d. returns and simple models

- Jump diffusion for equities
- Constant real short term interest rate

How do these strategies perform on historical data?

To take into account

- Actual real equity returns
- Randomly varying interest rates

We will use *block bootstrap* resampling of the historical data.

Block bootstrap II

Compute and store optimal strategies for fitted parameters *1926:1-2015:12* (optimal in synthetic market).

- Control is a table of optimal equity fractions, as a function of time-to-go and real wealth

Divide time horizon T into k blocks

$$T = kb \;\; ; \;\; b = \text{ blocksize}$$

A single bootstrap path

- Select k blocks at random (with replacement) from historical data[11]
- Concatenate blocks to form a single path

Some tweaks

- Stationary block bootstrap: block size is randomly sampled from geometric distribution
- Data is *wrapped around* to avoid end effects

[11] Same block of data for stocks and bonds

Bootstrap III

Using blocks takes into account serial data dependence, non-constant interest rates

Econometric criteria

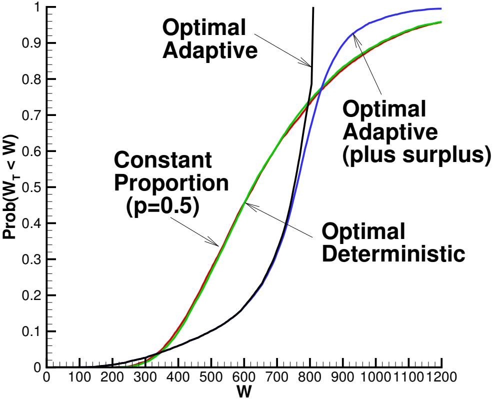
- Blocksize $\simeq$ 2 months for stock index (deflated)
- Blocksize $\simeq$ 5 years for bond index (deflated)

Experimented with blocksizes $.25 - 5.0$ years, results qualitatively similar $\rightarrow$ show results for expected blocksize $= 2$ years

			Probability of Shortfall	
Strategy	$E[W_T]$	$std[W_T]$	$W_T < 500$	$W_T < 600$
Const. Prop. ($p = .5$)	677	264	.27	.45
Optimal Deterministic	676	257	.26	.45
Optimal Adaptive	**700**	**137**	**.10**	**.17**

Table : 10,000 bootstrap samples. $E[W_T] = 706$ in synthetic market.

Bootstrap resampling: cumulative distribution function



Bootstrap resampled results, using historical data from *1926:1 - 2015:12*

Expected blocksize = 2 years.

CTE(95%) Adaptive MV: 279

CTE(95%) Deterministic: 316

We give up some large gains (optimal adaptive) to reduce probability of shortfall.
CTE paths: bootstrap samples 1930*s* many times in a single path.

Conclusions

- Optimal deterministic glide path strategies
  - Negligibly better than constant proportion
  - Most TDFs[12] in the market use deterministic strategies!
- Optimal adaptive MV strategies
  - Robust to parameter misspecification
  - Give up large gains to reduce downside risk (probability of shortfall) compared to deterministic strategies
  - Good performance on bootstrap resampled tests (control computed using average historical parameters)
- CTE(95%) for adaptive MV slightly slightly worse than CTE(95%) for deterministic strategy
  - This effect is smaller for bootstrap resampling than for synthetic markets
  - Occurs when markets trend downward for 30 years

[12] Over $750 billion in TDFs in US at end of 2015