Dynamic Mean Variance Asset Allocation: Numerics and Backtests

Peter Forsyth[1]

[1]Cheriton School of Computer Science
University of Waterloo

December 19, 2016
Delft

Outline

1. Dynamic mean variance
   - Embedding result $\Rightarrow$ quadratic target
   - Removal of spurious points
2. HJB PDE
   - Intuitive discretization
   - Semi-Lagrangian timestepping and explicit control
   - Unconditionally stable, monotone and consistent
3. Calibrate to historical market data (1926-2015)
   - Synthetic market: M-V optimal beats constant proportion
   - Backtests using real historical data: M-V optimal even better!
   - Constant proportion beats any deterministic *glide path* strategy (lumpsum investment)[1]

     $\rightarrow$ M-V optimal beats any deterministic glide path strategy

[1] Strategy used in Target Date funds (over $750 billion in US)

Dynamic Mean Variance: Abstract Formulation

Define:

$$
\begin{aligned}
X &= \text{Process} \\
&\quad \frac{dX}{dt} = \text{SDE} \\
x &= (X(t) = x) = \text{State} \\
W(X(t), t) &= \text{total wealth}
\end{aligned}
$$

Control $c(X(t), t)$ is applied to $X(t)$

Define admissible set $\mathcal{Z}$, i.e.

$$
c(x, t) \in \mathcal{Z}(x, t)
$$

Mean and Variance under control $c(X(t), t)$

Let:

$$\underbrace{E_{t,x}^{c(\cdot)}[W(T)]}_{Reward} = \text{Expectation conditional on } (x, t) \text{ under control } c(\cdot)$$

$$\underbrace{\text{Var}_{t,x}^{c(\cdot)}[W(T)]}_{Risk} = \text{Variance conditional on } (x, t) \text{ under control } c(\cdot)$$

Important:

- mean and variance of $W(T)$ are as observed at time t, initial state x.

Basic Problem: Find Pareto Optimal Strategy

We desire to find the investment strategy $c^*(\cdot)$ such that, there exists no other other strategy $c(\cdot)$ such that

$$\underbrace{E_{t,x}^{c(\cdot)}[W_T]}_{Reward\ \ under\ \ strategy\ \ c(\cdot)} \geq \underbrace{E_{t,x}^{c(^*\cdot)}[W_T]}_{Reward\ \ under\ \ strategy\ \ c^*(\cdot)}$$

$$\underbrace{\text{Var}_{t,x}^{c(\cdot)}[W_T]}_{Risk\ \ under\ \ strategy\ \ c(\cdot)} \leq \underbrace{\text{Var}_{t,x}^{c^*(\cdot)}[W_T]}_{Risk\ \ under\ \ strategy\ \ c^*(\cdot)}$$

and at least one of the inequalities is strict.

Scalarization: For $\lambda > 0$, find $c(\cdot)$ which solves

$$\inf_{c(\cdot)} \left\{ \lambda \text{Var}_{t,x}^{c(\cdot)}[W_T] - E_{t,x}^{c(\cdot)}[W_T] \right\}$$

Varying λ traces out the efficient frontier.

Pareto optimal points

Let

$$\mathcal{E} = E_{t,x}^{c(\cdot)}[W_T] \quad ; \quad \mathcal{V} = \mathrm{Var}_{t,x}^{c(\cdot)}[W_T]$$

The *achievable set* $\mathcal{Y}$ is

$$\mathcal{Y} = \{(\mathcal{V}, \mathcal{E}) \,:\, c(\cdot) \in \mathcal{Z}\},$$

Given $\lambda > 0$, define scalarization set [2]

$$\mathcal{S}_\lambda(\mathcal{Y}) = \{(\mathcal{V}, \mathcal{E}) \in \bar{\mathcal{Y}} : \lambda\mathcal{V} - \mathcal{E} = \inf_{(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}} (\lambda\mathcal{V}_* - \mathcal{E}_*\}$$

The efficient frontier $\mathcal{Y}_P$ is

$$\mathcal{Y}_P = \bigcup_{\lambda > 0} \mathcal{S}_\lambda(\mathcal{Y})$$

The efficient frontier is a collection of Pareto points

[2] $\bar{\mathcal{Y}}$ is the closure of $\mathcal{Y}$.

Scalarization: intuition[3]

Recall scalarization set:

$$\mathcal{S}_\lambda(\mathcal{Y}) = \{(\mathcal{V},\mathcal{E}) \in \bar{\mathcal{Y}} : \lambda\mathcal{V} - \mathcal{E} = \inf_{(\mathcal{V}_*,\mathcal{E}_*)\in\mathcal{Y}} (\lambda\mathcal{V}_* - \mathcal{E}_*\} \tag{1}$$

Geometric interpretation:

- Consider the straight line (for fixed λ)

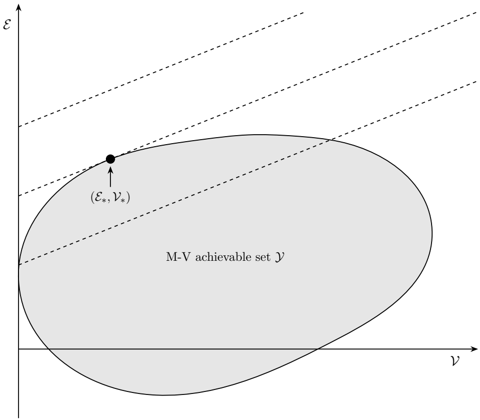
$$\lambda\mathcal{V} - \mathcal{E} = C_1 \tag{2}$$

Points in (1)

- Choose C_1 as small as possible, such that:
  - $\rightarrow$ Intersection of $\mathcal{Y}$ and straight line (2) has at least one point

[3] We may not get all the Pareto points here if $\mathcal{Y}$ is not convex

Intuition



Move dotted lines line $\lambda \mathcal{V} - \mathcal{E} = C_1$ to the left as much as possible (decrease C_1)

Line will touch $\mathcal{Y}$ at Pareto point

Problem

Pareto point

$$\lambda \mathcal{V} - \mathcal{E} = \inf_{(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}} (\lambda \mathcal{V}_* - \mathcal{E}_* \} \tag{3}$$

Problem arises from variance

$$\begin{aligned} \mathcal{V} &= E^c[W(T)^2] - (E^c[W(T)])^2 \\ & \quad (E^c[W(T)])^2 \rightarrow \text{ problem for dynamic programming} \end{aligned}$$

Consider the optimization problem (for fixed γ)

$$\inf_{(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}} \mathcal{V} + \mathcal{E}^2 - \gamma \mathcal{E} \tag{4}$$

Note that

$$\mathcal{V} + \mathcal{E}^2 = E^c[W(T)^2]$$

Minimizing (4) can be done using dynamic programming

Embedded Objective Function Intuition

Examine points $(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}$ such that (for fixed γ)

$$\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E} = \inf_{(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}} \mathcal{V}_* + \mathcal{E}_*^2 - \gamma\mathcal{E}_* \tag{5}$$

Geometric interpretation:

- Consider the parabola

$$\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E} = C_2 \tag{6}$$

Points in (5)

- Choose C_2 as small as possible, such that
  - Intersection of parabola and $\mathcal{Y}$ has at least one point

Rewriting equation (6)

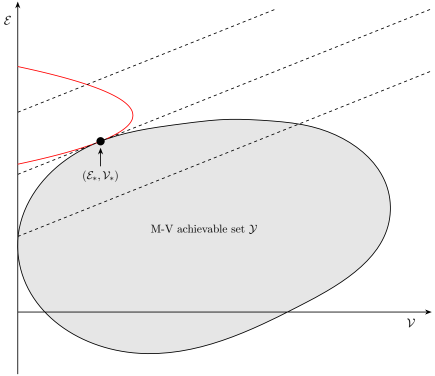
$$\begin{aligned} \mathcal{V} &= -\left(\mathcal{E}^2 - \gamma\mathcal{E}\right) + C_2 = -\left(\mathcal{E} - \gamma/2\right)^2 + \gamma^2/4 + C_2 \\ &= -\left(\mathcal{E} - \gamma/2\right)^2 + C_3. \end{aligned}$$

Parabola faces left, symmetric about line $\mathcal{E} = \gamma/2$

Embedded Pareto Points

Suppose $(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}_P \rightarrow \exists \lambda > 0,\ C_1$, s.t.

$$\lambda \mathcal{V}_* - \mathcal{E}_* = C_1$$



Parabola:
$\mathcal{V} = -(\mathcal{E} - \gamma/2)^2 + C_3$.

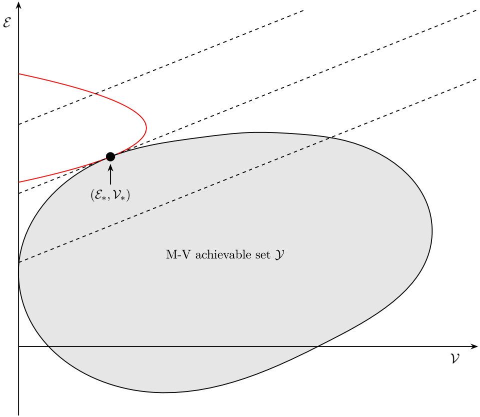
$\exists\ \gamma/2, C_3$, such that we can

Move parabola to left (C_3)

Move parabola up/down ($\gamma/2$)

$\Rightarrow$ intersect line $\lambda\mathcal{V} - \mathcal{E} = C_1$ at a single point $(\mathcal{V}^*, \mathcal{E}^*)$.

Tangency Condition



Parabola $\mathcal{V} = -(\mathcal{E} - \gamma/2)^2 + C_3$ tangent to line $\lambda\mathcal{V} - \mathcal{E} = C_1$ at $(\mathcal{V}_*, \mathcal{E}_*)$

$$\left(\frac{\partial\mathcal{E}}{\partial\mathcal{V}}\right)_{parabola} = \lambda \quad ; \quad \lambda = \text{ slope of dotted lines}$$

$$\rightarrow \qquad \gamma/2 = 1/(2\lambda) + \mathcal{E}_*$$

Embedding Result

Theorem 1 ((Li and Ng (2000); Zhou and Li (2000))
If

$$\lambda \mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}} (\lambda \mathcal{V} - \mathcal{E}), \quad (7)$$

then

$$\mathcal{V}_0 + \mathcal{E}_0^2 - \gamma \mathcal{E}_0 = \inf_{(\mathcal{V},\mathcal{E}) \in \mathcal{Y}} (\mathcal{V} + \mathcal{E}^2 - \gamma \mathcal{E}), \quad (8)$$

$$\gamma = \frac{1}{\lambda} + 2\mathcal{E}_0$$

Implication

- We can determine all the Pareto points from (7) by solving problem (8)

Value function

Note:

$$\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E} = E^c_{t,x}[(W(T) - \frac{\gamma}{2})^2] + \frac{\gamma^2}{4},$$

Define value function[4] (ignore $\gamma^2/4$ term when minimizing)

$$V(x,t) \;=\; \inf_{c(\cdot)\in\mathcal{Z}} E^{c(\cdot)}_{t,x}[(W(T) - \gamma/2)^2] \qquad (9)$$

Key Result: Given point $(\mathcal{V}^*, \mathcal{E}^*)$ on the efficient frontier, generated by control $c^*(\cdot)$, then $\exists\gamma$ s.t.

$\rightarrow$ $c^*(\cdot)$ is an optimal control for (9)

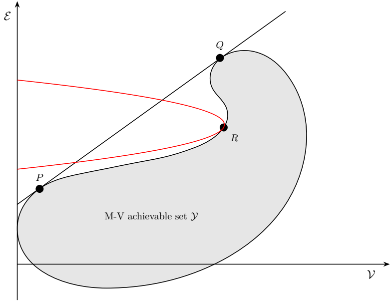
[4]Precommitment MV optimal $\equiv$ quadratic target optimal. Precommitment $\rightarrow$ choose target wealth $\gamma/2$ at time zero

Spurious points

But, converse not necessarily true: i.e. there may be some $\gamma \in (-\infty, +\infty)$ s.t. $c^*(\cdot)$ which solves

$$V(x,t) \quad = \quad \inf_{c(\cdot)\in\mathcal{Z}} E^{c(\cdot)}_{t,x}[(W(T) - \gamma/2)^2] \tag{10}$$

does not correspond to a point on the efficient frontier



Technical Point: Precommitment vs. time consistent

We are solving for the optimal *precommitment* policy

- This is not *time-consistent*, since $\gamma/2$ (the target) depends on the initial state
- However, a way to think about this is as follows
  - At $t = 0$ we determine where we want to be on the efficient frontier. This fixes $\gamma/2$.
  - At $t > 0$, we can think of this policy as the optimal time consistent strategy which minimizes quadratic loss w.r.t. fixed $\gamma/2$.
  - This is intuitive and easy to explain to pension plan investors (Vigna (2014))
    - The target amount $\gamma/2$ is the amount needed to fund retirement

Basic Algorithm

Discretize the parameter γ

$$\gamma \in \Gamma^k = [-|\gamma^k_{\max}|, -|\gamma^k_{\max}| + h_k, \ldots, |\gamma^k_{\max}|] \quad (11)$$

$$h_k \to 0 \; ; \; \gamma^k_{\max} \to \infty \; ; \; k \to \infty \quad (12)$$

- For each γ_i,
  - Determine optimal control $c^*_{\gamma_i}(\cdot)$ by solving the embedded problem (solve HJB equation, store control)
  - Using this control, compute $E^{c^*_{\gamma_i}(\cdot)}_{t,x}[(W_T)]$, $Var^{c^*_{\gamma_i}(\cdot)}_{t,x}[(W_T)]$ via Monte Carlo (one point on the frontier)

Does this converge to *true* efficient frontier as $k \to \infty$?

Problems

1. Controls which minimize $E_{t,x}^{c(\cdot)}[(W(T) - \gamma/2)^2]$ (from numerical solve)
   - May generate spurious points (e.g. non-convex $\mathcal{Y}$)
2. The control which minimizes

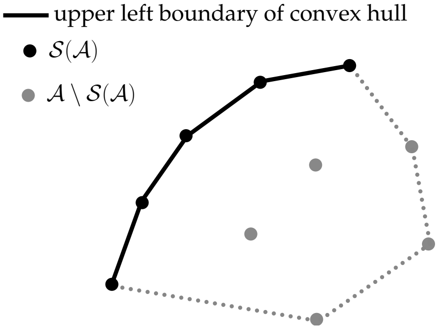
$$E_{t,x}^{c(\cdot)}[(W(T) - \gamma/2)^2] \tag{13}$$

   may not be unique.
   - Numerical HJB solve for fixed $\gamma/2$
     $\rightarrow$ picks out only one control $c^*(\cdot)$
   - Does the control we compute correspond to a point in $\mathcal{Y}_P$?

Convergent Algorithm[5]

For $k = 0, 1, \ldots$

- Solve value function $\forall \gamma_i \in \Gamma^k$
- Generate set of candidate points on the efficient frontier $\mathcal{A}^k$
- Determine upper left convex hull $\mathcal{S}(\mathcal{A}^k)$
- Approximate points on efficient frontier: $\mathcal{A}^k \cap \mathcal{S}(\mathcal{A}^k)$



[5] Tse, Forsyth, Li (2014, SIAM Cont. Opt.); Dang,Forsyth, Li (2016, Numerische Mathematik)

Convergence result

Recall def'n of scalarization set:

$$\mathcal{S}_\lambda(\mathcal{X}) = \{(\mathcal{V}_*, \mathcal{E}_*) \in \overline{\mathcal{X}} : \lambda\mathcal{V}_* - \mathcal{E}_* = \inf_{(\mathcal{V},\mathcal{E})\in\mathcal{X}} \lambda\mathcal{V} - \mathcal{E}\}, \quad (14)$$

Suppose $\mathcal{S}_\lambda(\mathcal{Y}) \neq \emptyset, \lambda > 0$ (i.e. $\mathcal{S}_\lambda(\mathcal{Y})$ are points on the efficient frontier for fixed λ)

Theorem 2

Suppose Γ^k *is* systematically *refined* [6] *as* $k \to \infty$, *and let* $(\mathcal{V}_k, \mathcal{E}_k) \in \mathcal{S}_\lambda(\mathcal{A}^k)$. *Let* $(\mathcal{V}_*, \mathcal{E}_*)$ *be a limit point of* $\{(\mathcal{V}_k, \mathcal{E}_k)\}$. *Then* $(\mathcal{V}_*, \mathcal{E}_*)$ *is on the original efficient frontier.*

Remark 1

All points on the approximate efficient frontier $\mathcal{A}^k \cap \mathcal{S}(\mathcal{A}^k)$ *are valid points on the true efficient frontier as* $k \to \infty$. [7]

---

[6] Any reasonable refinement satisfies this condition

[7] There may some gaps in the approximate frontier if there are 3 or more points on a straight line segment.

Asset allocation: risk free bond, stock index

Risk free bond B

$$dB = rB\ dt$$
$$r = \text{ risk-free rate}$$

Amount in risky stock index S (jump diffusion)

$$dS = (\mu - \rho\kappa)S\ dt + \sigma S\ dZ + (J-1)S\ dq$$

$$\mu = \mathbb{P} \text{ measure drift} \quad ; \quad \sigma = \text{ volatility}$$
$$dZ = \text{ increment of a Wiener process}$$

$$dq = \begin{cases} 0 & \text{with probability } 1 - \rho\ dt \\ 1 & \text{with probability } \rho dt, \end{cases}$$
$$\log J \sim \text{ double exponential.} \quad ; \quad \kappa = E[J-1]$$

Optimal Control

Define:

$$\begin{aligned} X &= (S(t), B(t)) = \text{Process} \\ x &= (S(t) = s, B(t) = b) = (s, b) = \text{State} \\ (s+b) &= \text{total wealth} \end{aligned}$$

Let $(s, b) = (S(t^-), B(t^-))$ be the state of the portfolio the instant before applying a control

The control $c(s, b) = (d, B^+)$ generates a new state

$$\begin{aligned} b &\rightarrow B^+ \\ s &\rightarrow S^+ \\ & S^+ = \underbrace{(s+b)}_{wealth\ \ at\ \ t^-} - B^+ - \underbrace{d}_{withdrawal} \end{aligned}$$

Note: we allow cash withdrawals of an amount $d \geq 0$ at a rebalancing time

Optimal de-risking (free cash flow)

Let

$$F(t) = \frac{\gamma}{2} e^{-r(T-t)}$$
$$= \text{discounted target wealth}$$

Proposition 1 (Dang and Forsyth (2016))

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

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

We will refer to the amount withdrawn as a *free cash flow*. [8]

---

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

Constraints on the strategy

The investor can continue trading only if solvent

$$\underbrace{W(s,b) = s + b > 0}_{Solvency\ \ condition}. \tag{15}$$

In the event of bankruptcy, the investor must liquidate [9]

$$S^+ = 0 \quad ; \quad B^+ = W(s,b) \quad ; \quad \text{if } \underbrace{W(s,b) \leq 0}_{bankruptcy} \ .$$

Leverage is also constrained

$$\frac{S^+}{W^+} \quad \leq \quad q_{\max}$$

$$W^+ = S^+ + B^+ = \text{ Total Wealth}$$

[9] The *No Donald Trump* trading condition.

HJB PIDE

Find optimal control $c(\cdot) \Rightarrow$ solve for value function

$$V(x,t) = \inf_{c \in \mathcal{Z}} \left\{ E^{c}_{t,x}[(W(T) - \gamma/2)^2] \right\} ,$$

Define:

$$\begin{aligned} \mathcal{L}V &\equiv \frac{\sigma^2 s^2}{2} V_{ss} + (\mu - \rho\kappa) s V_s - \rho V \ , \\ \mathcal{J}V &\equiv \int_0^\infty p(\xi) V(\xi s, b, \tau) \ d\xi \\ & \quad p(\xi) = \text{ jump size density } ; \ \rho = \text{ jump intensity} \end{aligned}$$

and the intervention operator $\mathcal{M}(c) \ V(s,b,t)$

$$\mathcal{M}(c) \ V(s,b,t) \ = \ V(S^+(s,b,c), B^+(s,b,c), t)$$

HJB PIDE II

Value function, control $c(\cdot) \Rightarrow$ solve impulse control HJB equation

$$\max\left[V_t + \mathcal{L}V + rbV_b + \mathcal{J}V, V - \inf_{c \in \mathcal{Z}} (\mathcal{M}(c)\ V) \right] = 0$$

Discretize computational domain $(s, b) \in [0, \infty) \times (-\infty, +\infty)$

$$\{s_1, s_2, \ldots, s_{i_{\max}}\} \quad ; \quad \{b_1, \ldots, b_{j_{\max}}\}$$

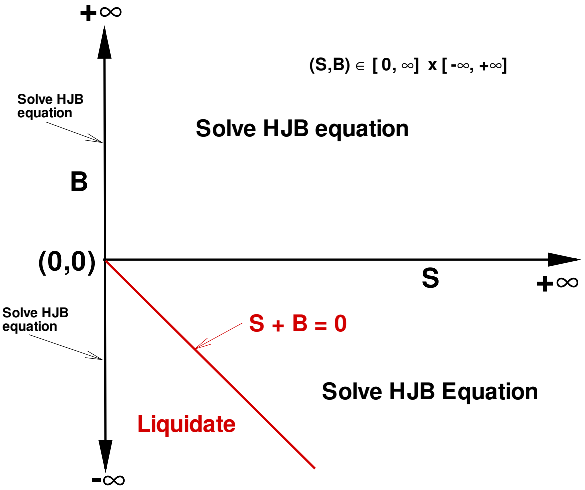
Constant timesteps, discretize control

$$\Delta\tau = \tau^{n+1} - \tau^n \quad ; \quad B^+ \in \{b_1, \ldots, b_{j_{\max}}\}$$

Discretization parameter h

$$\max_i (s_{i+1} - s_i) = \max_j (b_{j+1} - b_j) = \max_n (\tau^{n+1} - \tau^n) = O(h)$$

Computational Domain[10]



[10] If $\mu > r$ it is never optimal to short S

Intuitive Derivation of Discretization

Consider a set of discrete rebalancing times $\{t_1, t_2, \dots\}$
Define

$$t_m^+ = t_m + \epsilon \quad ; \quad t_m^- = t_m - \epsilon \quad ; \quad \epsilon \to 0^+ \tag{16}$$

At $t = t_m^+$, $s = S(t)$ and $b = B(t)$

Step $[t_m^+, t_{m+1}^-]$ (bond amount constant)

- The value function $V(s, b, t)$ evolves according to the PIDE

$$V_t + \overbrace{\mathcal{L}V}^{No \;\; rbV_b \;\; term} + \overbrace{\mathcal{J}V}^{Jump \;\; term} = 0,$$

Evolution over $[t_{m+1}^-, t_{m+1}^+]$

Step $[t_{m+1}^-, t_{m+1}]$ (Stock amount constant)

- Pay interest earned in $[t_m^+, t_{m+1}^-]$

$$V(s, b, t_{m+1}^-) = V(s, be^{r\Delta t}, t_{m+1}) \quad ; \quad \text{by no-arbitrage}$$
$$\Delta t = t_{m+1} - t_m$$

Step $[t_{m+1}, t_{m+1}^+]$

- Optimal rebalance

$$V(s, b, t_{m+1}) = \overbrace{\min_c V(S^+(s, b, c), B^+(s, b, c), t_{m+1}^+)}^{rebalance}$$

Backwards time: discrete solution

Now, we write these steps down in backwards time $\tau = T - t$

- Define $V_{i,j}^n \equiv$ discrete solution $V_h(s_i, b_j, \tau^n)$

$$\widetilde{V}_{i,j}^n = \overbrace{\min_{c \in \mathcal{Z}_h} V_h(S^+(s_i, b_j e^{r\Delta\tau}, c), B^+(s_i, b_j e^{r\Delta\tau}, c), \tau^n)}^{\textit{Optimization step with } \tau^n \textit{ data}}$$

$$\underbrace{\frac{V_{i,j}^{n+1}}{\Delta\tau} - \mathcal{L}_h V_{i,j}^{n+1} - \mathcal{J}_h V_{i,j}^{n+1} = \frac{\widetilde{V}_{i,j}^n}{\Delta\tau}}_{\textit{Linear time advance}}$$

Formally: Semi-Lagrangian timestepping and explicit impulse control

Discretization Properties

1. Positive coefficient method used to discretize $\mathcal{P}$,
2. Jump term: fixed point iteration + FFT for dense matrix-vector product
3. Linear interpolation used to approximate V_h at off grid points (needed for optimal control)

Assume strong comparison property holds:

- Consistent, ℓ_∞ stable, monotone
  - $\hookrightarrow$ Convergence to viscosity solution

Example Asset Allocation: Constant Proportions

According to Benjamin Graham[11], *defensive* investors should

- Pick a fraction p of wealth to invest in a diversified equity fund (e.g. $p = 1/2$).
- Invest $(1 - p)$ in bonds
- Rebalance to maintain this asset mix
  - $\rightarrow$ i.e. a constant proportion strategy

How does this strategy compare with standard target date funds, which follow a deterministic glide path over time T?

Typical deterministic glide path strategy[12]

$$p(t) = \frac{\left(110 - \text{ your age }\right)}{100}$$

[11] Benjamin Graham, *The Intelligent Investor*

[12] This used to be (100 − *your age*) but people are living longer

Lumpsum Investment: ineffectiveness of glide paths

Consider any *deterministic* glide path strategy $p(t)$

$$p(t) \quad = \quad \text{fraction of wealth invested in equities}$$

Define a constant weight strategy p^* where

$$\begin{aligned} p^* &= \frac{1}{T} \int_0^T p(s)\, ds \\ &= \text{time average fraction in equities} \end{aligned}$$

Let W denote total wealth. We can prove (GBM + jumps) [13]

$$\overbrace{E[W(T)]}^{constant\ \ weight} = \overbrace{E[W(T)]}^{glide\ \ path} \quad ; \quad \overbrace{Var[W(T)]}^{constant\ \ weight} \leq \overbrace{Var[W(T)]}^{glide\ \ path} \qquad (17)$$

Backtests on historical data and MC simulations[14] indicates (17) holds in general → constant proportion beats deterministic glide path

[13] Graf (2016), Forsyth and Vetzal (2016)

[14] Basu et al (2011), Arnott et al (2013), Esch and Michaud (2014)

Technical Point: lumpsum vs. periodic contributions

Constant proportion beats *any* deterministic glide path for a lumpsum investment.

- This is not true for periodic contributions (accumulation) or withdrawals (decumulation)
- However, for $T > 20$ years
  - Numerical tests show that the *optimal* deterministic MV glide path strategy is only slightly better than a constant proportions strategy.
  - In practice, deterministic glide path strategies are devised using heuristics.

Monte Carlo Simulation Results

- Inflation-adjusted equity: jump diffusion[15] model estimated using CRSP[16] total return index and CPI data (1926 to 2015)
- Inflation-adjusted bonds: average real 3M T-bills (1926 to 2015)

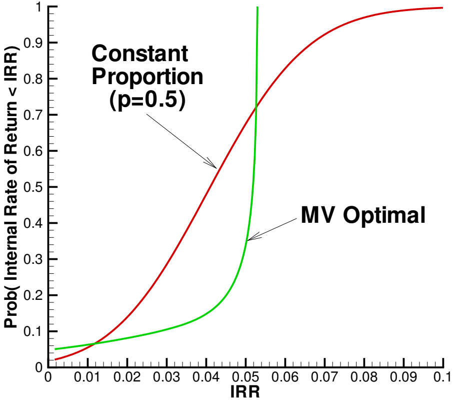
Strategy	Expected Value	Standard Deviation	Prob(W(T)) < 300	Prob(W(T)) < 400
Constant Proportion $p = 0.5$	417	**299**	**0.41**	**0.60**
M-V Optimal Control	417	**117**	**0.13**	**0.22**

Table : Investment horizon $T = 30$ years. Initial investment $W(0) = 100$. Optimal de-risking; no trading if insolvent; maximum leverage $= 1.5$, rebalancing once/year.

Standard deviation reduced by 250%, shortfall probability reduced by 3×

[15] Jump size had double exponential distribution (Kou, 2002)
[16] Capitalization weighted index of all stocks traded on major US exchanges.

Cumulative Distribution Function: IRR[17]



$E[W(T)] = 417$ same for both strategies

Optimal policy: Contrarian: when market goes down $\rightarrow$ increase stock allocation; when market goes up $\rightarrow$ decrease stock allocation

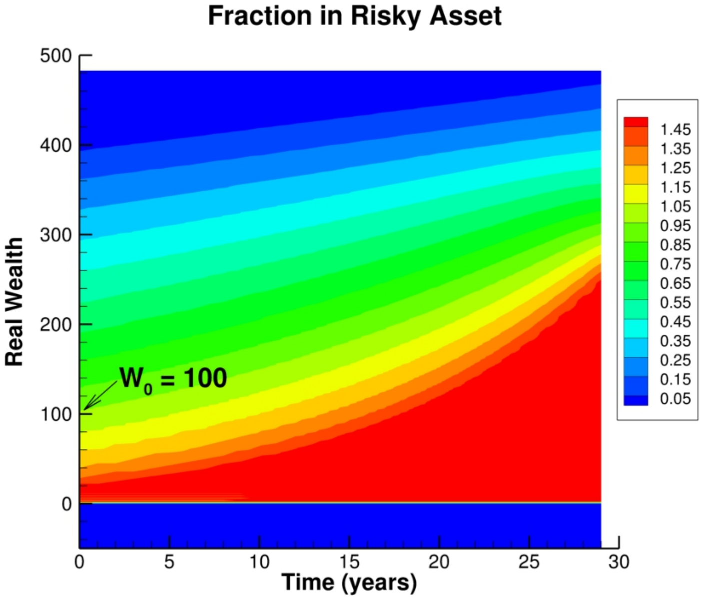
Optimal allocation gives up gains $\gg$ target in order to reduce variance and probability of shortfall.

Investor must pre-commit to target wealth

MV optimal beats constant proportion, consequently it also beats any deterministic glide path!

[17] Internal rate of return (i.e. effective rate of return) $= \log(W(T)/W(0))/T$

Strategy Heat Map



Red: maximum leverage

Blue: 100% bond

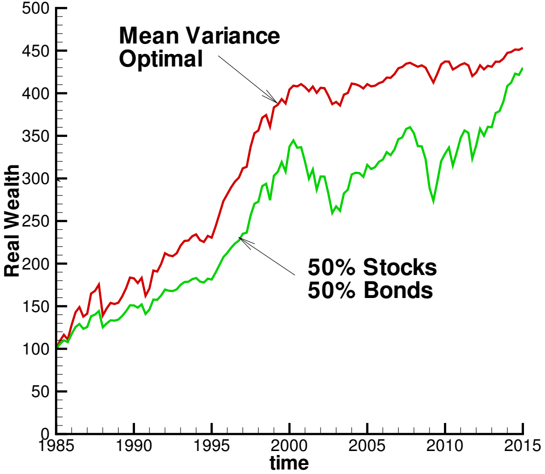
$E[W_T] = 417$

Back Testing

M-V optimal performance on historical data

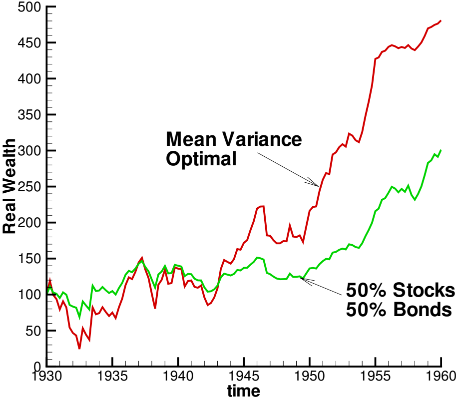
- Compute and store strategy based on estimated parameters for entire historical period (January 1, 1926 - December 31, 2014).
- $E[W(T)]$ same as for constant proportion strategy ($p = .5$), for this set of average parameters.
- Select starting date
- Compare:
  - Optimal MV strategy (based on average parameters, not tuned to this period)
  - Constant proportion strategy

Back Test, Real Returns: Jan 1, 1985 - Dec 31, 2014[18]



[18] $W(1985) = 100$. Maximum leverage 1.5. Optimal MV strategy computed using parameters for 1926-2015 period. Yearly rebalancing.

Back Test, Real Returns: Jan 1, 1930 - Dec 31, 1959[19]



Note *Falling Knife* effect in 1932

Can we fix this: regime switching plus machine learning?

[19] $W(1930) = 100$. Maximum leverage 1.5. Optimal MV strategy computed using parameters for 1926-2015 period. Yearly rebalancing.

Bootstrap Resampling: 1926-2015

More Scientific Test: Resampling

Use real historical data, monthly returns

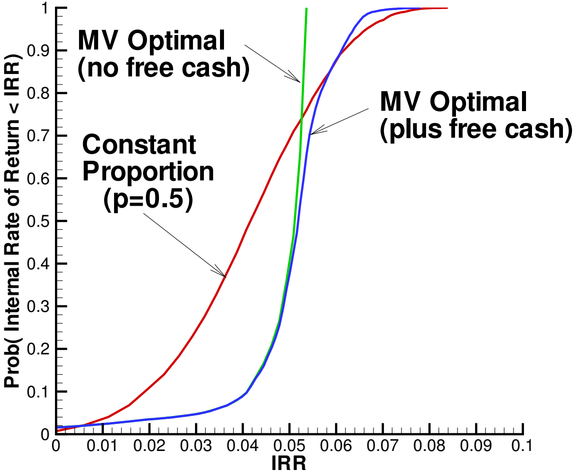
- Randomly draw 30 years of returns (with replacement) from historical returns (blocksize 10 years)
- 10,000 simulations, each block starts at random month

Strategy	Expected Value	Standard Deviation	$Pr(W(T)) < 300$	Expected Free Cash
Constant Proportion $p = 0.5$	385	**183**	**0.38**	0.0
M-V Optimal Control	431	**84**	**0.07**	40

Table : $T = 30$ years. $W(0) = 100$. Yearly rebalancing. Optimal de-risking ; no trading if insolvent; maximum leverage $= 1.5$.

Performs even better on actual historical data than on synthetic market data!

Resampled Cumulative Distribution Function: IRR[20]



[20]Internal rate of return, (i.e. effective rate of return) $= \log(W(T)/W(0))/T$

Technical Point: Bootstrap resampling

Data is *wrapped around* to avoid end effects (i.e. 1930s appears more often).

To minimize blocksize end effects, blocksize is selected randomly from a geometric distribution

Time series	Optimal Expected Block size (months)
Real 90-day T-bills	50.1
Real 10 year treasury	4.7
Real CRSP index (cap weight)	1.8
Real CRSP index (equal weight)	10.4

Table : Optimal expected blocksize $1/p$ where the blocksize is distributed according to a geometric distribution $Pr(b = k) = (1 - p)^{k-1}p$. The algorithm in (Politis et al (2009)) is used.

Technical Point: Bootstrap resampling vs rolling quarters

A common backtest is to use *rolling quarters*

- This amounts to starting the investment at each historical quarter, and then seeing how it performed over the next 30 years.
- Summary statistics of probability of failure are then quoted
- However, there are not enough 30 year rolling blocks to get reasonable samples $\rightarrow$ investment results are too good!

Strategy	Expected Value	$Pr(W(T)) < 300$	$Pr(W(T)) < 400$
Constant Proportion $p = 0.5$	351	**0.31**	**0.73**
M-V Optimal Control	453	**0.0**	**0.07**

Table : $T = 30$ years. $W(0) = 100$. Yearly rebalancing. Optimal de-risking ; no trading if insolvent; maximum leverage $= 1.5$. Rolling quarters, 30 year blocksize, data is wrapped-around.

Conclusions

- M-V strategy is very robust
  - Insensitive to calibration ambiguity
  - MC tests: insensitive to random perturbations of synthetic market SDE parameters
  - Stochastic volatility: typical parameters, insignificant for long term investors
  - 10 year treasuries (instead of 3-M) similar results
  - Good results on historical backtests
- Similar results for accumulation, decumulation
- M-V beats constant proportion, i.e. probability of shortfall $2 - 3\times$ smaller
  - $\rightarrow$ Constant proportion beats any deterministic glide path
- M-V optimal equivalent to minimizing quadratic loss w.r.t. wealth target
  - Optimal strategy is M-V optimal **and** quadratic loss optimal
- More sophisticated models
  - Regime switching? (machine learning approach being investigated)