Least-Squares Monte Carlo (LSMC) Guided Approximate Dynamic Portfolio Optimization

Forward Simulation, Backward Optimization, and Results

Ethan Baratz

West Chester University

April 2026

Why This Matters

- **Standard models react only to current risk and return**
- **The real investor goal** is to maximize lifetime economic value while avoiding catastrophic losses that permanently impair recovery
- **Key Insight:** Use forward-looking simulations to detect when the future looks dangerous *before* losses materialize
- **Our Approach (LSMC):** Learn a state-dependent portfolio policy that adapts based on both current wealth *and* predicted future outcomes

State vs Control:

$$S_t = (\underbrace{W_t}_{\substack{\text{Wealth}\\ \text{at time } t}}, \underbrace{\pi_{t-1}}_{\substack{\text{Portfolio}\\ \text{at } t-1}}, \underbrace{t}_{\text{Time}}) \longrightarrow U_t = (\underbrace{\pi_t}_{\substack{\text{Portfolio}\\ \text{at } t}}, \underbrace{c_t}_{\text{Consumption}})$$

- S_t: state variables (information available at time t)
- U_t: control variables (decisions made at time t)

The Objective Function

At each time step, we solve a **regularized mean–variance program** with tail-risk control:

$$\pi_t^* = \arg\min_{x \in \mathbb{R}^n} \Big\{ \underbrace{\tfrac{1}{2} x^\top \Sigma_t x}_{\text{variance risk}} - \underbrace{(\mu_t - r\mathbf{1})^\top x}_{\text{excess return}} + \underbrace{\delta \|x\|_2^2}_{\text{diversification}} + \underbrace{\kappa \|x - \pi_{t-1}\|_1}_{\text{transaction costs}} + \underbrace{\Lambda\Big(\alpha + \tfrac{1}{M(1-q)} \textstyle\sum_m u_m\Big)}_{\text{CVaR penalty}} \Big\}$$

Preferences [1, 3]

CRRA utility: $u(c) = \dfrac{c^{1-\gamma}}{1-\gamma}$, Discount factor: $\beta_d = e^{-\rho \Delta t}$

- Fully invested: $\sum_i x_i = 1$, Long-only: $x_i \in [0, 1]$

Risk, Return, and Diversification

Variance Risk & Excess Return [2]

$$\frac{1}{2}x^\top \Sigma_t x - (\mu_t - r\mathbf{1})^\top x$$

- The variance term penalizes portfolio risk using the conditional covariance matrix, capturing diversification across assets
- The return term rewards holding assets with excess expected return above the risk-free rate r

Diversification (Ridge) Penalty [4]

$$\delta \|x\|_2^2$$

- Shrinks weights toward zero, preventing extreme concentration or leverage
- Acts as a ridge penalty that smooths allocations and improves numerical stability

Transaction Costs & Tail Risk (CVaR)

Transaction Cost Penalty [9]

$$\kappa \|x - \pi_{t-1}^{(m)}\|_1$$

- Penalizes changes from the previous portfolio allocation
- The ℓ_1 norm reflects proportional trading costs; encourages sparse, realistic rebalancing

CVaR Tail-Risk Penalty [5]

$$\alpha + \frac{1}{M(1-q)} \sum_{m=1}^{M} u_m$$

- Implements Conditional Value-at-Risk (CVaR) at quantile q
- Focuses the optimizer explicitly on *worst-case* outcomes across Monte Carlo paths
- Scaled by the adaptive penalty $\Lambda_{k,t}$ (discussed next)

Adaptive Risk Penalty: $\Lambda_{k,t}$

Key idea: The CVaR penalty is *modulated* by the state of each wealth bin.

Wealth Ratio

$$\xi_{k,t} = \frac{\widetilde{W}_{k,t}}{\widetilde{W}_{\mathrm{ref},t}}$$

Compares bin wealth to the median across all paths. $\xi > 1$: richer; $\xi < 1$: poorer.

Penalty Construction

$$\lambda_{\text{base}} = 5\, e^{-0.2(\xi_{k,t}-1)}, \qquad \lambda_{\Phi} = e^{-1.5\, \widetilde{\Phi}_{k,t}}$$

$$\Lambda_{k,t} = \min\big(20,\, \max\big(1,\ \lambda_{\text{base}} \cdot \lambda_{\Phi}\big)\big)$$

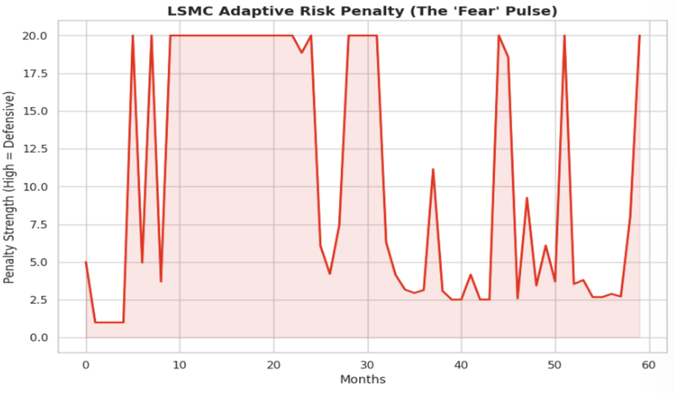
- **Wealth effect:** poorer bins face higher penalties
- **Future outlook effect:** worse continuation values amplify the penalty further

Backward Pass: From Simulations to Decisions

1. **Forward Monte Carlo:** Simulate M wealth paths $\{W_t^{(m)}\}$ under a baseline policy
2. **Backward Regression:** Working backwards from T, estimate continuation values $\Phi_t^{(m)}$ via ridge regression on features (log-wealth, volatility, drawdown, etc.)
3. **Wealth Binning:** Group paths into 5 percentile bins; each bin gets a representative wealth $\widetilde{W}_{k,t}$ and continuation value $\widetilde{\Phi}_{k,t}$
4. **Policy Optimization:** For each bin, solve the quadratic program with adaptive $\Lambda_{k,t}$ and dynamic consumption $c^*_{k,t}$
5. **Wealth Update:**

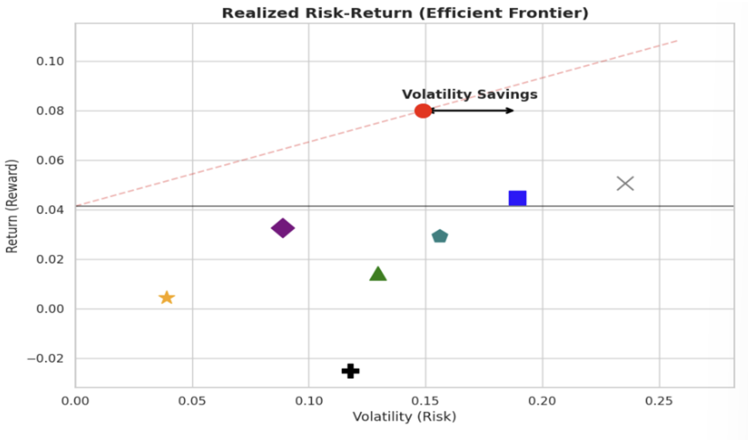
$$W_{t+1}^{-} = W_t^{+} + r(W_t^{+} - \mathbf{1}^\top \theta_t)\Delta t + \theta_t^\top \Delta R_t - c_t W_t^{+} \Delta t$$

The "Fear Pulse" — Λ Over Time (Jan. 2008– Dec. 2012)



- Penalty spikes to 20 during market stress, forcing the model into a defensive posture *before* losses fully materialize
- Penalty decreases in calm periods, allowing the optimizer to seek returns

Results: Realized Risk–Return (Efficient Frontier)



LSMC (red circle) achieves highest return *above* the efficient frontier line, with significantly lower volatility than comparable models

Backtest Performance (Bear Markets 2000 & 2008)

Model	Final Wealth	Std Dev	Sharpe
LSMC	**$1,743.69**	$744.22	**0.403**
Mean-Variance	$1,666.15	$890.03	0.275
Regime Switching HMM	$1,594.55	$929.21	0.204
Buy-and-Hold	$1,573.06	$468.16	0.287
Equal-Weight Baseline	$1,396.96	$368.70	0.205

LSMC leads in final wealth and Sharpe ratio. Max drawdown: 25.1% vs. 39–41% for MV and HMM.

Tail Risk & Risk-Adjusted Alpha

Tail Risk Metrics (Bear Markets 2000 & 2008)

Model	Exp. Shortfall (ES)	Max Drawdown
LSMC	**$1,136.74**	**25.14%**
Mean-Variance	$998.27	39.10%
Regime Switching HMM	$725.20	41.11%
Buy-and-Hold	$913.48	40.48%

Risk-Adjusted Alpha & Capture

Model	Beta	Alpha	Treynor	Down Cap
LSMC	−0.041	6.62%	**1.254**	−55.8%
Mean-Variance	−0.005	6.06%	0.845	−55.3%
Regime Switching HMM	−0.056	4.25%	0.587	−52.0%
Buy-and-Hold	0.024	6.95%	1.965	−44.8%

Conclusion

What We Built

A state-dependent portfolio policy that jointly optimizes risk, return, and tail exposure using LSMC.

Key Contributions:

- **Adaptive CVaR Penalty ($\Lambda_{k,t}$):** Dynamically modulates tail-risk based on each path's wealth relative to peers and its continuation value, $\widetilde{\Phi}_{k,t}$
- **Empirical Outperformance:** LSMC achieves highest ES ($\$1,136.74$) and lowest max drawdown (25.1%) across two major bear markets

Future Work

Replace the ridge regression continuation value estimator with a **neural network** allowing for higher dimensional state spaces and richer policy representations across wealth bins.

J. G. Andréasson and P. V. Shevchenko, "A bias-corrected Least-Squares Monte Carlo for solving multi-period utility models," *Journal of Computational Finance*, 2021.

H. Markowitz, "Portfolio Selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

R. C. Merton, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, vol. 3, no. 4, pp. 373–413, 1971.

S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004, Sections 4.4 and 6.3.2.

R. T. Rockafellar and S. Uryasev, "Optimization of Conditional Value-at-Risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.

References II

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent Measures of Risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

J. Mossin, "Optimal Multiperiod Portfolio Policies," *The Journal of Business*, vol. 41, no. 2, pp. 215–229, 1968.

R. F. Engle, O. Ledoit, and M. Wolf, "Large Dynamic Covariance Matrices," *Journal of Business & Economic Statistics*, 2019.

X. Mei and F. J. Nogales, "Portfolio Selection with Proportional Transaction Costs and Predictability," *Journal of Banking & Finance*, vol. 94, pp. 131–151, 2018.