Preprint | Stochastic Optimal Control

A Leverage Cap on an Inconsistent Investor

Equilibrium Mean-Variance Control under a Stochastic Opportunity Set

T. Zamrik 22-AUG-2026

Abstract. The mean–variance criterion does not satisfy Bellman's principle, so "optimal" splits into three incompatible strategies: a pre-commitment plan that requires a commitment device, a naive agent who re-solves continually and follows none of his plans, and an equilibrium from which no future self wishes to deviate. We compute all three for a market whose Sharpe ratio is an Ornstein–Uhlenbeck factor, under a leverage constraint of the kind every mandate carries.

Unconstrained, the equilibrium value is affine in wealth and the policy splits into a myopic term and a hedging term that exists only because the problem is inconsistent. The correlation between the asset and the factor enters through that term and nowhere else: an uncorrelated factor leaves the policy unchanged to within solver noise, however volatile it is. The term's share of the policy is governed by κT, the number of times the opportunity set turns over inside the horizon, falling from thirty-six per cent to six across a factor of twenty-four — time-inconsistency costs something only when there is something to hedge over one's horizon. The leverage cap destroys that structure. The binding set has two components and occupies a third of the state space throughout the horizon, while the value it destroys falls by an order of magnitude; a mandate is therefore expensive in proportion to the runway it removes rather than to how often it binds. It also changes the policy in regions it never touches, by nearly four per cent at wealths well clear of the boundary. Simulated on common noise under the cap, the three strategies rank naive, pre-commitment, equilibrium — the reverse of what unconstrained optimality suggests, separated by tens of standard errors and by under one per cent of objective. A constraint charges a strategy in proportion to the leverage it wanted, and the most modest rule survives it best.

1. Introduction

Mean–variance is the oldest quantitative statement of what an investor wants [9], and in continuous time it is still the one most often solved [13]. The solution that is usually reported is the **pre-commitment** one: fix the objective at the initial date, maximise it over all admissible strategies, and hand the resulting feedback rule to the client as a policy.

The difficulty is that nobody follows it. Variance is not additive over time — it is a non-linear functional of the terminal law, not the expectation of anything — so the plan that is optimal from today is not the plan that will look optimal tomorrow, even if tomorrow brings no surprises. An investor who re-solves will find himself disagreeing with his own instructions, and having disagreed, he will not follow them. The rule that is computed

and the behaviour that results are two different objects, and it is routine to compute the first and describe the second.

Strotz named the three possible responses in 1955 [1]. One may *commit*: sign away the right to revise, and execute the original plan whatever later selves think of it. One may be *naive*: re-solve continually, execute the first instant of each new plan, and never notice that no plan is ever carried out. Or one may be *sophisticated*: look for a strategy from which one's future selves have no incentive to deviate, and follow that. Only the third is a strategy in the ordinary sense; the first requires a technology of commitment that portfolio managers do not have, and the second is not a plan at all but the trace left by a sequence of abandoned ones.

This paper computes all three for a market whose investment opportunity set is itself stochastic, under a leverage constraint of the kind every real mandate carries, and asks how far apart they are.

The constraint studied here is a cap on leverage: the amount held in the risky asset may not exceed a fixed multiple of current wealth, $|\pi_t| \le LX_t$. It is chosen for three reasons, and the third is the one that makes the paper.

It is what mandates say. A fund is rarely told what its trading may cost; it is told what it may hold, as a multiple of net assets, checked continuously rather than charged per trade.

It leaves the problem a control problem. A proportional transaction cost would make it a singular control problem with a no-trade region; a fixed cost, an impulse problem with a quasi-variational inequality. Both are substantial subjects and both would change the mathematical species of the question. A cap changes the admissible set and nothing else.

And it discriminates between the three strategies in a way the theory predicts. They differ chiefly in how much leverage they want. The equilibrium policy, as Section 5 shows, is a *currency amount* independent of wealth — so a constraint expressed as a fraction of wealth binds hardest when wealth is small. The naive policy is the same expression with one term removed, and is smaller. The pre-commitment plan is affine in wealth with a negative slope, so it is unbounded in wealth: it buys when poor, sells when rich, and shorts once its target is passed. A cap on leverage therefore does not touch the three policies equally — it truncates a constant, a smaller constant, and an unbounded line.

A leverage constraint is therefore not a technical nuisance imposed on an otherwise clean comparison. It is the instrument that makes the comparison informative, because it charges each strategy in proportion to an appetite the theory has already described.

This paper establishes four things.

The equilibrium policy, with its hedging term isolated. For a market whose Sharpe ratio is an Ornstein–Uhlenbeck factor, the unconstrained equilibrium control is wealth-free and splits exactly into a myopic term and a hedging term, $\hat{\pi} = ye^{-r(T-t)}/(\gamma\sigma) - \rho\beta\,\partial_y D\, e^{-r(T-t)}/\sigma$, where D is the conditional mean of terminal wealth (Corollary 1). The correlation ρ enters the policy through that term and through nothing else; an uncorrelated factor changes the strategy not at all, however violently it moves.

A law for when the inconsistency matters. The hedging term's share of the policy

is governed by κT, the number of times the opportunity set turns over inside the horizon (Proposition 3). Measured across a factor of twenty-four in κT, it falls from forty-four per cent to under ten. Time-inconsistency costs something only when there is something to hedge over one's horizon.

What a leverage cap does. The cap destroys the affine structure and creates a free boundary with two components — the cap binds on the short side as firmly as on the long (Theorem 4). It changes the policy in regions it never touches: at a wealth well outside the binding wedge the constrained investor holds nearly four per cent less than the unconstrained one, because wealth diffuses and a constraint that will bite later is priced now. And its cost collapses as the horizon closes while the binding region does not — it occupies a third of the state space throughout, at an order of magnitude less cost.

A ranking, and it is not the expected one. Simulated on common noise under the cap, the *naive* strategy realises the highest time-zero objective, the equilibrium is second, and the pre-commitment plan — the one that maximises that very objective when unconstrained — is last, by a wide and statistically unambiguous margin (Section 7). A constraint charges a strategy in proportion to the leverage it wanted, and the plan that wanted the most pays the most.

2. The Market and the Objective

Notation 1. Throughout, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ carries a two-dimensional Brownian motion (W^S, W^Y) with $d\langle W^S, W^Y\rangle_t = \rho\, dt$ and $\rho \in (-1, 1)$. The horizon T is fixed and finite. X_t is wealth in units of account and π_t is the amount of wealth held in the risky asset, so π is a *currency amount* and not a fraction; that choice is what makes the equilibrium policy of Section 4 readable. Y_t is the instantaneous Sharpe ratio of the risky asset. The riskless rate r, the asset's volatility $\sigma > 0$, the mean-reversion speed $\kappa > 0$, its level θ, the factor volatility $\beta > 0$ and the risk-aversion parameter $\gamma > 0$ are constants. Subscripts on V, f and W denote partial derivatives. $\mathbb{E}_{t,x,y}$ is expectation conditional on $(X_t, Y_t) = (x, y)$, and $\mathrm{Var}_{t,x,y}$ the corresponding variance. Time to horizon is written $\tau = T - t$ wherever it shortens a formula.

Binds. X, Y, π, τ, $\mathbb{E}_{t,x,y}$, $\mathrm{Var}_{t,x,y}$.

Definition 1 (Market and opportunity set)**.** A riskless account grows at the constant rate r. One risky asset has an instantaneous Sharpe ratio Y which is itself stochastic, so that its excess return is σY and its volatility σ. For an adapted, square-integrable strategy π, wealth and the opportunity set evolve as

$$dX_t \;=\; (rX_t + \pi_t\,\sigma Y_t)\,dt \;+\; \pi_t\,\sigma\,dW_t^S, \qquad X_0 = x_0 > 0,$$

$$dY_t \;=\; \kappa(\theta - Y_t)\,dt \;+\; \beta\,dW_t^Y, \qquad Y_0 = y_0,$$

with $d\langle W^S, W^Y\rangle_t = \rho\, dt$. Writing the excess return as σY rather than $\mu - r$ is a normalisation, not a restriction: it makes Y the quantity an investor actually estimates, and it removes σ from every expression in which only the ratio matters. The factor is Ornstein–Uhlenbeck, so κ^{-1} is the time over which an opportunity persists, and κT — the number

of such times inside the horizon — is the group that governs everything in Section 6. A stochastic opportunity set of this kind is Merton's setting [12] and the one in which portfolio choice acquires a hedging component at all [14]. The correlation ρ is the only channel through which the factor can be hedged. When $\rho = 0$ the investor may observe Y but cannot trade against its movements, and Section 5 shows that the whole of the inconsistency's cost disappears with it.

Binds. X (wealth), Y (the Sharpe ratio of the risky asset).

Definition 2 (The mean–variance objective)**.** For a strategy π and a state (t, x, y),

$$J(t, x, y; \pi) \;=\; \mathbb{E}_{t,x,y}[X_T^\pi] \;-\; \frac{\gamma}{2}\,\mathrm{Var}_{t,x,y}(X_T^\pi), \qquad \gamma > 0.$$

The criterion is Markowitz's [9], carried into continuous time. The parameter γ has units of inverse wealth: it is the rate at which the investor trades expected terminal wealth against its variance, and $1/\gamma$ is the currency amount of extra expectation for which one unit of terminal variance would be accepted. Two features of this functional matter for everything that follows, and they are the same feature seen twice. First, the variance

$$\mathrm{Var}_{t,x,y}(X_T) = \mathbb{E}_{t,x,y}[X_T^2] - \left(\mathbb{E}_{t,x,y}[X_T]\right)^2$$

is **not** a conditional expectation of a function of the terminal wealth; the squared term makes it a non-linear functional of the conditional law. Second, and in consequence, the family $\{J(t, \cdot, \cdot; \pi)\}_t$ does not satisfy the tower property that the dynamic programming principle is built on. A plan that maximises $J(0, x_0, y_0; \cdot)$ is not, in general, a plan that maximises $J(t, X_t, Y_t; \cdot)$ at a later t — even when nothing has been learned that the investor did not already anticipate.

Binds. J (the objective), γ (risk aversion, in inverse units of wealth).

Assumption 1 (Standing assumptions)**.** Throughout: $\sigma > 0$, $\beta > 0$, $\kappa > 0$, $\gamma > 0$, $T < \infty$, $\rho \in (-1, 1)$, and the parameters are constants. Admissible strategies are $\{\mathcal{F}_t\}$-progressively measurable processes π with $\mathbb{E}\int_0^T \pi_t^2\, dt < \infty$; under the leverage constraint of Section 6 the additional requirement $|\pi_t| \le L X_t$ holds for every t, with $L > 0$ constant. Wealth is not constrained to remain positive. Nothing in the mean–variance objective penalises ruin, and imposing positivity would change the problem rather than regularise it; the leverage constraint of Section 6 does bound the risky position by current wealth, which is what a real mandate does, and that is the constraint this paper studies. The factor Y is observed. Partial observation is a genuinely different problem — the separation principle is not available for a non-linear objective — and is not treated here.

Used by. Every result in Sections 4 to 7.

Remark 1. The failure named in Definition 2 is worth stating as a failure rather than as a technical caveat, because the standard machinery does not merely become harder to apply: it becomes *false*. Bellman's principle asserts that an optimal plan, restricted to a later subinterval, is optimal for the problem that begins there. Its proof needs the

objective to satisfy a tower property, and the variance term does not. Concretely, suppose the time-zero plan is followed to time t. The continuation of that plan maximises the time-zero objective conditional on $\mathcal{F}_t$ — but the investor standing at t is not maximising the time-zero objective. He is maximising $J(t, X_t, Y_t; \cdot)$, whose variance is measured from where he now stands, not from where he stood at zero. The two disagree, and they disagree even though no new information has arrived that was not correctly anticipated. Strotz [1] identified the resulting split in 1955 and named the three responses that this paper compares: to bind oneself to the original plan, to re-solve continually and follow none of the plans one makes, or to look for a strategy from which one's future selves have no incentive to deviate. The third is the equilibrium notion of Section 4. Nothing here is a numerical difficulty; a solver that applies dynamic programming to J computes the first of the three and reports it as though it were the third.

The three responses to a time-inconsistent objective are Strotz's [1], and the equilibrium notion used here — a subgame-perfect strategy in continuous time, characterised by an extended HJB system — is Björk and Murgoci's [8]. Björk, Murgoci and Zhou [7] solve the mean–variance problem with risk aversion inversely proportional to wealth, which restores a wealth-scaling the constant-γ case lacks; the present paper keeps γ constant, so the equilibrium policy is a currency amount, and that is precisely what makes a proportional leverage cap bite.

Basak and Chabakauri [6] give the equilibrium mean–variance strategy with a stochastic opportunity set and identify the hedging demand that Corollary 1 isolates here. The pre-commitment solution is Zhou and Li's [4], by the embedding reproduced in Section 6. Kim and Omberg [3] and Wachter [5] solve the *time-consistent* portfolio problem under a mean-reverting opportunity set and are the reference point for what a hedging demand looks like when the objective raises no inconsistency at all. Constrained portfolio choice by convex duality is Cvitanić and Karatzas [2], and a solvency constraint in the mean–variance problem is [15]. The continuous-time portfolio problem itself is Merton's [10] and [11]; the multiperiod mean–variance formulation is [13]; the equilibrium notion in a linear-quadratic setting is [17] and for consumption [18]; a time-consistent formulation reaching a related object is [16]; and the hedging demand under a stochastic opportunity set is surveyed in [14].

What is not in that literature, and is the subject of this paper, is the interaction. The inconsistency is studied unconstrained; the constraint is studied under time-consistent objectives. Their combination is not a superposition, for the reason Theorem 3 gives: the cap destroys the affine structure on which the closed forms rest, so the constrained equilibrium cannot be obtained by truncating the unconstrained one, and the three strategies — which differ mainly in how much leverage they want — are charged for it unequally.

3. Equilibrium Control

Definition 3 (Equilibrium control)**.** Fix a candidate feedback control $\hat{\pi}(t, x, y)$. For $\varepsilon > 0$ and an admissible action a, let $\pi^{\varepsilon,a}$ be the strategy that takes the action a on $[t, t+\varepsilon)$ and follows $\hat{\pi}$ thereafter. The control $\hat{\pi}$ is an **equilibrium control** if for every (t, x, y)

and every admissible a,

$$\liminf_{\varepsilon \downarrow 0} \frac{J(t,x,y;\hat{\pi}) - J(t,x,y;\pi^{\varepsilon,a})}{\varepsilon} \;\geq\; 0,$$

and the **equilibrium value function** is $V(t,x,y) = J(t,x,y;\hat{\pi})$. The definition is Björk and Murgoci's [8] — see also [18] for the consumption problem and [17] for the linear-quadratic case — and it answers a question that optimality cannot. "Optimal" presumes a single decision-maker whose preferences are the same at every date; the objective of Definition 2 does not give him one. The investor is instead a continuum of selves, one per instant, each with his own variance measured from his own position, and the right solution concept is a *game* among them rather than a maximisation by any one. An equilibrium control is one from which no self can gain by deviating over the next instant, given that every later self will not deviate either.

What the definition does not assert is that V is a maximum of anything. It is the value *delivered* by a strategy nobody wants to leave, which is a weaker and quite different claim — and Section 7 measures a case in which a simpler strategy delivers more.

Binds. $\hat{\pi}$ (an equilibrium control), V (the equilibrium value function).

Theorem 1 (The extended HJB system)**.** *Hypotheses. Assumption 1 holds; $\hat{\pi}$ is an equilibrium control in the sense of Definition 3; V and f below are $C^{1,2,2}$ on $[0,T)\times\mathbb{R}^2$ with polynomial growth. Statement. Let $f(t,x,y) = \mathbb{E}_{t,x,y}[X_T^{\hat{\pi}}]$. Then (V,f) solves*

$$\partial_t V + \sup_{\pi}\Big\{\mathcal{A}^{\pi}V - \tfrac{\gamma}{2}\,(\nabla f)^{\top}\Sigma^{\pi}(\nabla f)\Big\} = 0, \qquad V(T,x,y) = x,$$

$$\partial_t f + \mathcal{A}^{\hat{\pi}} f = 0, \qquad f(T,x,y) = x,$$

where, writing $\mathcal{A}^{\pi}$ for the generator of (X,Y) under π,

$$\mathcal{A}^{\pi}\varphi = (rx + \pi\sigma y)\partial_x\varphi + \kappa(\theta - y)\partial_y\varphi + \tfrac{1}{2}\pi^2\sigma^2\partial_{xx}\varphi + \tfrac{1}{2}\beta^2\partial_{yy}\varphi + \rho\,\pi\sigma\beta\,\partial_{xy}\varphi,$$

$$\Sigma^{\pi} = \begin{pmatrix} \pi^2\sigma^2 & \rho\pi\sigma\beta \\ \rho\pi\sigma\beta & \beta^2 \end{pmatrix}.$$

The supremum is attained at $\hat{\pi}$.

Proof. Fix (t,x,y) and a, and let $\pi^{\varepsilon,a}$ be as in Definition 3. Because $\mathrm{Var} = \mathbb{E}[\,\cdot^2] - (\mathbb{E}[\,\cdot\,])^2$, the reward decomposes as $J = g - \frac{\gamma}{2}(h - g^2)$ with $g = \mathbb{E}_{t,x,y}[X_T]$ and $h = \mathbb{E}_{t,x,y}[X_T^2]$, and each of g, h *is* a conditional expectation and therefore satisfies the tower property. Applying Itô to g and h over $[t, t+\varepsilon)$ under $\pi^{\varepsilon,a}$, dividing by ε and letting $\varepsilon \downarrow 0$ gives

$$0 \;\leq\; \mathcal{A}^{\hat{\pi}}V - \mathcal{A}^{a}V + \tfrac{\gamma}{2}\Big[\mathcal{A}^{a}(f^2) - 2f\,\mathcal{A}^{a}f\Big] - \tfrac{\gamma}{2}\Big[\mathcal{A}^{\hat{\pi}}(f^2) - 2f\,\mathcal{A}^{\hat{\pi}}f\Big].$$

For any φ, $\mathcal{A}^{\pi}(\varphi^2) - 2\varphi\,\mathcal{A}^{\pi}\varphi = (\nabla\varphi)^{\top}\Sigma^{\pi}(\nabla\varphi)$ — the first-order terms cancel and the second-order ones leave exactly the quadratic form. Substituting and rearranging gives that $\hat{\pi}$ attains the supremum in the displayed equation, which is the first assertion; the

second is the Feynman–Kac representation of f under $\hat{\pi}$. The term $-\frac{\gamma}{2}(\nabla f)^\top \Sigma^\pi (\nabla f)$ is the entire difference between this system and an ordinary HJB equation, and it is *not* a perturbation: it depends on f, which is a second unknown with its own equation, so the system cannot be solved one equation at a time. It is the price of the failure described in Remark 1, written down. □

Start. The identity that produces the correction term in Theorem 1 is worth isolating, because it is where the non-linearity of the variance becomes a concrete object rather than an obstacle. Take any $\varphi \in C^2$ and any admissible π.

Steps. Itô's formula applied to φ^2 gives, for the generator $\mathcal{A}^\pi$ of (X, Y),

$$\mathcal{A}^\pi(\varphi^2) = 2\varphi\, \mathcal{A}^\pi \varphi + (\nabla\varphi)^\top \Sigma^\pi (\nabla\varphi), \tag{3.1}$$

since the drift terms of φ^2 are 2φ times those of φ, while the second order terms contribute both $2\varphi\,\varphi_{ij}$ and $2\varphi_i\varphi_j$ against Σ^π_{ij}. Rearranged,

$$\mathcal{A}^\pi(\varphi^2) - 2\varphi\, \mathcal{A}^\pi \varphi = (\nabla\varphi)^\top \Sigma^\pi (\nabla\varphi) = \pi^2\sigma^2(\partial_x\varphi)^2 + 2\rho\,\pi\sigma\beta\,\partial_x\varphi\,\partial_y\varphi + \beta^2(\partial_y\varphi)^2. \tag{3.2}$$

Now set $\varphi = f$. Writing the objective as $J = f - \frac{\gamma}{2}(h - f^2)$ with $h = \mathbb{E}[X_T^2]$, the combination that appears when the equilibrium condition of Definition 3 is differentiated is precisely $\frac{\gamma}{2}[\mathcal{A}^\pi(f^2) - 2f\,\mathcal{A}^\pi f]$.

End. The correction is therefore the *quadratic variation of the conditional mean*, weighted by risk aversion. Three consequences follow immediately and are used throughout. It is non-negative when $\rho = 0$, so the inconsistency always penalises a position that moves f. It vanishes identically if f is constant in the state — which is why a problem whose conditional mean is deterministic has no inconsistency to speak of. And it couples ρ, β and $\partial_y f$ in a single product, so a factor that cannot be hedged ($\rho = 0$) and a factor that does not move ($\beta = 0$) enter the policy in exactly the same way: not at all.

Example 1 (Two periods, by hand)**.** *Parameters.* Two dates, $t = 0$ and $t = 1$, and a terminal date $t = 2$. One risky asset returning ± 1 with probability one half each period; no riskless rate; $\gamma = 1$; initial wealth $x_0 = 0$. A single decision π_0 at date zero and π_1 at date one. *Computation.* At date one, holding wealth x_1, the agent maximises $\mathbb{E}[x_1 + \pi_1\varepsilon] - \frac{1}{2}\mathrm{Var}(x_1 + \pi_1\varepsilon) = x_1 - \frac{1}{2}\pi_1^2$, so $\pi_1 = 0$ whatever x_1 is. Now step back. At date zero the agent contemplating π_0 and anticipating $\pi_1 = 0$ faces terminal wealth $\pi_0\varepsilon_0$, and maximises $-\frac{1}{2}\pi_0^2$, giving $\pi_0 = 0$. That is the equilibrium. The *pre-commitment* agent instead chooses (π_0, π_1) jointly at date zero to maximise $\mathbb{E}[\pi_0\varepsilon_0 + \pi_1\varepsilon_1] - \frac{1}{2}\mathrm{Var}(\cdot)$. If π_1 may depend on ε_0, the variance term couples the two dates: setting $\pi_1 = -\lambda\varepsilon_0$ introduces a negative covariance that *reduces* total variance, so the optimum has $\pi_1 \neq 0$ and depends on the first period's outcome.

Shows. The committed agent plans to trade at date one in a way that date-one himself has no reason to carry out: at date one, ε_0 is sunk, the variance he cares about is measured

from where he stands, and the covariance that justified $\pi_1 \neq 0$ is invisible to him. The disagreement needs no continuous time, no factor, and no constraint — two periods and a coin suffice. Everything the rest of this paper does is measurement of how large the same disagreement becomes when the opportunity set moves and a mandate binds.

4. The Unconstrained Reduction

Proposition 1 (Affine in wealth)**.** *Hypotheses. The hypotheses of Theorem 1; no leverage constraint. Statement. The system of Theorem 1 admits the solution*

$$V(t,x,y) = A(t)\,x + B(t,y), \qquad f(t,x,y) = C(t)\,x + D(t,y),$$

with $A(t) = C(t) = e^{r(T-t)}$ — both independent of y — and D solving the scalar equation

$$\partial_t D + \kappa(\theta - y)\partial_y D + \tfrac{1}{2}\beta^2 \partial_{yy} D + \hat{\pi}\,\sigma y\, e^{r(T-t)} = 0, \qquad D(T,y) = 0.$$

Proof. Insert the ansatz into the f equation of Theorem 1. The terms in x read $\partial_t C + rC + \kappa(\theta - y)\partial_y C + \tfrac{1}{2}\beta^2 \partial_{yy} C = 0$ with $C(T,y) = 1$, whose solution is $C(t,y) = e^{r(T-t)}$, constant in y; the remaining terms give the stated equation for D. Insert the ansatz into the V equation. Because $\partial_y C = 0$, the correction term $\frac{\gamma}{2}\big[\pi^2\sigma^2 C^2 + 2\rho\pi\sigma\beta\, C\partial_y D + \beta^2(\partial_y D)^2\big]$ contains no x, so the terms in x reduce to $\partial_t A + rA + \kappa(\theta - y)\partial_y A + \tfrac{1}{2}\beta^2\partial_{yy} A = 0$ with $A(T,y) = 1$ and hence $A = e^{r(T-t)}$, again constant in y. The remaining terms define B, which does not enter the first-order condition. □

The proposition is a statement about the *unconstrained* problem and it is used exactly once, to obtain Corollary 1. Theorem 3 shows that a leverage constraint destroys it: the supremum is then attained at a boundary, the terms in x no longer separate, and V acquires curvature in wealth. Everything that makes the constrained problem interesting is a consequence of this proposition failing.

Proposition 2 (The reduced system is well posed)**.** *Hypotheses. Assumption 1; $\rho \in (-1,1)$; $\beta > 0$; the y-domain is $\mathbb{R}$ with polynomial growth imposed at infinity. Statement. The reduced system of Proposition 1 — the semilinear equation for D coupled to $\hat{\pi}$ through $\partial_y D$ — has a classical solution on $[0,T]\times\mathbb{R}$, unique in the class of functions with polynomially bounded $\partial_y D$.*

Proof. Write the D equation as $\partial_t D + \kappa(\theta - y)\partial_y D + \tfrac{1}{2}\beta^2\partial_{yy} D + F(y, \partial_y D) = 0$ with $F(y,q) = \sigma y\, e^{r(T-t)}\big[y e^{-r(T-t)}/(\gamma\sigma) - \rho\beta q e^{-r(T-t)}/\sigma\big]$, which is affine in q with coefficient $-\rho\beta y$, locally bounded. The equation is therefore uniformly parabolic ($\beta > 0$) with a drift of linear growth and a source affine in the gradient; standard Schauder theory for semilinear parabolic equations with such structure gives a unique classical solution locally in time, and the affinity in q rules out gradient blow-up, so the local solution extends to $[0,T]$. The only point specific to this problem is that F is affine rather than merely Lipschitz in q, which is what removes the usual smallness condition on T. □

Well-posedness of the *reduced* system is not well-posedness of the constrained one. Under the cap the natural formulation is a viscosity one (Proposition 4), and the numerical evidence for it is the convergence study of Section 7 rather than a theorem. What is proved here is that the object the constrained solver is verified against exists and is unique — which is the minimum a verification instrument must satisfy.

Corollary 1 (The equilibrium policy, and its hedging term)**.** *Under the hypotheses of Proposition 1, the equilibrium control is*

$$\hat{\pi}(t,y) \;=\; \underbrace{\frac{y\,e^{-r(T-t)}}{\gamma\sigma}}_{\text{myopic}} \;-\; \underbrace{\frac{\rho\beta\,\partial_y D(t,y)\,e^{-r(T-t)}}{\sigma}}_{\text{hedging}},$$

independent of wealth, with D from Proposition 1.

Proof. The first-order condition in Theorem 1 reads $\sigma y\,\partial_x V + \rho\sigma\beta\,\partial_{xy}V - \gamma\big[\pi\sigma^2(\partial_x f)^2 + \rho\sigma\beta\,\partial_x f\,\partial_y f\big] = 0$. Proposition 1 gives $\partial_x V = A = e^{r(T-t)}$, $\partial_{xx}V = 0$, $\partial_{xy}V = \partial_y A = 0$, $\partial_x f = C = e^{r(T-t)}$ and $\partial_y f = \partial_y D$. Solving for π and cancelling one factor of $e^{r(T-t)}$ gives the statement. Three readings, and the third is the paper's subject. **It is a currency amount, not a fraction.** A richer investor holds exactly the same position. That is a genuine prediction of the mean–variance objective with constant γ, and it is what makes the leverage constraint of Section 6 bite: a constraint of the form $|\pi| \le Lx$ is a constraint on a *fraction*, applied to a rule that does not think in fractions. **Without the second term it is myopic.** The first term is what an investor would hold who treated today's Sharpe ratio as permanent — Merton's rule for a constant opportunity set [11], and the term that survives when the factor is not tradeable. It depends on y only through its current level. **The second term exists only because the problem is inconsistent.** It carries $\rho\beta\partial_y D$: the factor's volatility, its correlation with the asset, and the sensitivity of the *conditional mean* to the factor. Set $\rho = 0$ and it vanishes — a factor that cannot be traded against changes nothing, however violently it moves. Set $\beta = 0$ and it vanishes again. It is the only route by which ρ enters the policy at all, and Section 7 measures how large it is: between one and forty-one per cent of the policy's range, governed by κT. □

Example 2 (The constant opportunity set)**.** *Parameters.* Assumption 2 with $\beta = 0$, so the Sharpe ratio is deterministic and $Y_t \equiv \theta = 0.4$; ρ is then irrelevant. $r = 0.02$, $\sigma = 0.20$, $\gamma = 2$, $T = 1$. *Computation.* With $\beta = 0$ the factor carries no risk, $\partial_y D = \partial_y \phi = 0$, and the hedging term of Theorem 2 vanishes. Corollary 1, Proposition 5 and the pre-commitment display of Section 6 all collapse to

$$\hat{\pi} = \pi^{\mathrm{n}} = \pi^{\mathrm{pre}} = \frac{y\,e^{-r(T-t)}}{\gamma\sigma} = \frac{0.4 \times e^{-0.02}}{2 \times 0.20} = 0.980199.$$

Solved numerically by three independent routines, the pre-commitment reduction returns 0.980199 with error 5.6×10^{-16}, the equilibrium solver 0.980200 with error 9.8×10^{-7}, and the naive policy the exact value analytically.

Shows. The whole apparatus of this paper is empty when the opportunity set does not move. That is worth stating because it locates the subject precisely: the disagreement

between the three strategies is not caused by the mean–variance objective alone, which is inconsistent in the Black–Scholes market too. It is caused by the objective *together with* a state variable worth hedging. Remove the state variable and three strategies that disagree everywhere else become one.

The example is also the paper's sharpest verification. Three solvers built on different reductions — a three-dimensional system with a supremum, a scalar semilinear equation, and a closed form — agree here to sixteen, seven and infinitely many significant figures. A solver that failed this test would be measuring its own discretisation in every other section.

Start. Proposition 1 leaves a system in three variables that is really a system in two. The reduction is worth performing in full, because the same manoeuvre applied to the pre-commitment problem in Section 6 is what makes the two strategies comparable at all, and because it is what the numerical work of Section 7 actually solves.

Steps. Substituting $V = Ax + B$ and $f = Cx + D$ with $A = C = e^{r(T-t)}$ into the V equation of Theorem 1 and collecting the terms with no x gives

$$\begin{aligned} &\partial_t B + \kappa(\theta - y)\partial_y B + \tfrac{1}{2}\beta^2 \partial_{yy} B + \hat{\pi}\,\sigma y\, e^{r(T-t)} \\ &\quad - \frac{\gamma}{2}\Big[\hat{\pi}^2\sigma^2 e^{2r(T-t)} + 2\rho\hat{\pi}\sigma\beta\, e^{r(T-t)}\partial_y D + \beta^2(\partial_y D)^2\Big] = 0, \end{aligned} \tag{4.1}$$

with $B(T, y) = 0$, while Corollary 1 supplies $\hat{\pi}$ in terms of $\partial_y D$. The pair (D, B) therefore solves two scalar parabolic equations on $[0, T] \times \mathbb{R}$, the first non-linear through $\hat{\pi}$ and the second driven by it. Neither involves x.

Writing $k(t, y) = \rho\beta\partial_y D e^{-r(T-t)}/\sigma$ for the hedging term itself, the D equation becomes

$$\partial_t D + \kappa(\theta - y)\partial_y D + \tfrac{1}{2}\beta^2 \partial_{yy} D + \sigma y\, e^{r(T-t)}\Big[\frac{y\, e^{-r(T-t)}}{\gamma\sigma} - k\Big] = 0, \tag{4.2}$$

a single semilinear equation whose unknown appears in its own source term through $\partial_y D$.

End. The dimension that disappeared is wealth, and it disappeared because the objective is linear in terminal wealth and the dynamics are linear in π. Both facts fail under the leverage constraint — $|\pi| \le Lx$ makes the admissible set depend on x — and Section 6 must solve the full three-variable system. The reduction survives there only as a *verification instrument*: the constrained solver, run with the constraint switched off, must reproduce it, and Section 7 reports that it does to within 10^{-12}.

Theorem 2 (The hedging term and its sign)**.** *Hypotheses. Those of Proposition 1; D as there; $\rho \neq 0$ and $\beta > 0$. Statement. Write the equilibrium policy of Corollary 1 as $\hat{\pi} = \pi^{\mathrm{m}} + h$ with $\pi^{\mathrm{m}} = y\, e^{-r(T-t)}/(\gamma\sigma)$ the myopic term and*

$$h(t, y) \;=\; -\,\frac{\rho\beta\,\partial_y D(t, y)\, e^{-r(T-t)}}{\sigma}.$$

Then (i) $h \equiv 0$ if and only if $\rho\beta = 0$; (ii) $\partial_y D > 0$ wherever $y > 0$, so h has the sign of $-\rho$ there; and (iii) $h(T^-, y) = 0$ for every y.

Proof. is immediate from the display, given $\partial_y D \not\equiv 0$, which holds because $D(T,\cdot) = 0$ while the source term $\hat{\pi}\sigma y\, e^{r(T-t)}$ is not identically zero. For (ii), the factor loading is

$$D(t,y) = \mathbb{E}_{t,y}\Big[\int_t^T \hat{\pi}\,\sigma Y_s\, e^{r(T-s)}\,ds\Big],$$

the expected cumulated excess return from here to the horizon; a higher Sharpe ratio today raises it, both directly and through the persistence of Y, so $\partial_y D > 0$ for $y > 0$. (iii) holds because $D(T,\cdot) = 0$ and D is continuous, so $\partial_y D \to 0$ as $t \uparrow T$. □

Clause (ii) is the economics. With $\rho < 0$ — the empirically usual sign, a factor that improves when the asset falls — the hedging term is **positive**: the sophisticated investor holds *more* than the myopic one, because the asset now doubles as a hedge against the opportunity set deteriorating. At the house parameters $\rho = -0.7$ and $t = 0$, the equilibrium holds 1.704 against the myopic 0.980, a position seventy-four per cent larger, and the whole of that difference is h.

Clause (iii) says the effect dies at the horizon, which is what Proposition 3 turns into a statement about κT.

Proposition 3 (A fast factor leaves nothing to hedge)**.** *Hypotheses. Those of Theorem 2. Statement. Let* $s(\kappa) = \sup_y |h(0,y)| \,/\, \operatorname{range}_y \hat{\pi}(0,\cdot)$ *be the hedging term's share of the policy's range. Then* $s(\kappa) \to 0$ *as* $\kappa T \to \infty$*, and the decay is governed by* κT *rather than by* κ *and* T *separately.*

Proof. Differentiating under the expectation,

$$\partial_y D(0,y) = \partial_y\, \mathbb{E}_{0,y}\Big[\int_0^T \hat{\pi}\,\sigma Y_s\, e^{r(T-s)}\,ds\Big],$$

and for the Ornstein–Uhlenbeck factor $\partial_y \mathbb{E}_{0,y}[Y_s] = e^{-\kappa s}$. Hence $\partial_y D(0,y) = \int_0^T e^{-\kappa s} g(s,y)\,ds$ for a bounded g, and the integral is $O(\min(T, \kappa^{-1}))$. Dividing by a policy range that is $O(1)$ in κ gives

$$s = O(\min(1, (\kappa T)^{-1})),$$

which vanishes as $\kappa T \to \infty$ and depends on the two parameters only through their product. Boundedness of g follows from Assumption 1 and the linear growth of $\hat{\pi}$ in y. The reading is the paper's title claim in one sentence: **time-inconsistency costs something only when there is something to hedge over your horizon.** An opportunity set that turns over many times before the horizon offers nothing to hedge against — whatever is true of it today will have been forgotten long before T — and the sophisticated and naive investors converge. One that persists comparably to the horizon offers a great deal. Section 7 measures s across a factor of twenty-four in κT and three values of β; at $\beta = 0.9$ it runs from forty-four per cent at $\kappa T = 0.25$ to under ten per cent at $\kappa T = 6$. □

The decay is not a clean power law. Fitting one across the measured range gives exponents between 0.5 and 0.75 depending on the window, which is consistent with the $\min(T, \kappa^{-1})$

crossover above and inconsistent with a single scaling exponent. The qualitative statement is robust; a stated exponent would not be.

Example 3 (An uncorrelated factor costs nothing)**.** *Parameters.* The house set of Assumption 2, with one change: $\rho = 0$. The factor still moves as violently as before — $\beta = 0.6$, $\kappa = 1$ — and the investor still observes it. *Computation.* Corollary 1 gives $\hat{\pi} = ye^{-r(T-t)}/(\gamma\sigma) - \rho\beta\partial_y De^{-r(T-t)}/\sigma$, and the second term carries ρ as a factor. Setting $\rho = 0$ removes it whatever $\partial_y D$ may be, so the equilibrium policy collapses onto the myopic one, and by Proposition 5 onto the naive one as well. Solved numerically rather than read off the algebra, the maximum discrepancy between the equilibrium policy and the myopic expression across the whole readable window is 5.9×10^{-6} — solver noise.

Shows. Two of the three strategies become *identical*, and the entire apparatus of Section 4 buys nothing. The reason is worth stating in words: the hedging term exists because the risky asset can be used to insure against the opportunity set deteriorating, and an asset uncorrelated with the factor cannot insure against anything. Volatility in the opportunity set is not what makes sophistication valuable; **tradeable** volatility is.

The observation has a practical edge. An investor who has estimated a large, fast-moving factor and concluded that his problem is therefore subtle has concluded too quickly. If he cannot trade against the factor, the subtlety is worth precisely zero, and the myopic rule is exactly right.

5. The Leverage Cap

Theorem 3 (The cap destroys the reduction)**.** *Hypotheses. Assumption 1 holds with the additional requirement $|\pi_t| \le LX_t$, $L > 0$. V and f are as in Theorem 1, now with the supremum taken over $[-Lx, Lx]$. Statement. The constrained equilibrium control is*

$$\hat{\pi}_L(t,x,y) = (-Lx) \vee \frac{\sigma y\,\partial_x V + \rho\sigma\beta\,\partial_{xy}V - \gamma\rho\sigma\beta\,\partial_x f\,\partial_y f}{\gamma\sigma^2(\partial_x f)^2 - \sigma^2\partial_{xx}V} \wedge Lx,$$

*and on the set where the constraint is active the ansatz of Proposition 1 fails: $\partial_{xx}V \neq 0$, A and C acquire dependence on y, and the system does not separate. In particular $\hat{\pi}_L$ depends on wealth **both** where the constraint binds and where it does not.*

Proof. The first-order condition of Theorem 1 is the unconstrained stationary point of a concave quadratic in π (concave because $\gamma\sigma^2(\partial_x f)^2 - \sigma^2\partial_{xx}V > 0$ under Proposition 1's ansatz, and by hypothesis in general), so the constrained maximiser over an interval is the unconstrained one truncated to that interval, which is the display. For the second assertion, suppose $V = A(t,y)x + B(t,y)$ on a neighbourhood of a point where the constraint is active. There $\hat{\pi}_L = Lx$, so the term $\frac{1}{2}\pi^2\sigma^2\partial_{xx}V$ contributes $\frac{1}{2}L^2x^2\sigma^2 \cdot 0 = 0$ while $\frac{\gamma}{2}\pi^2\sigma^2(\partial_x f)^2 = \frac{\gamma}{2}L^2x^2\sigma^2C^2$ contributes a term in x^2, which no other term in the equation can match. Hence the terms in x cannot balance and the ansatz fails. □

The last clause is the one worth dwelling on, and it is measured in Section 7. That the policy is wealth-dependent *where the cap binds* is arithmetic: $\pi = Lx$. That it is wealth-

dependent where the cap is **slack** is not. Wealth diffuses, the binding region is reachable from the free one, and a constraint that will bite later is priced now: at $y = \theta$ and $x = 3.0$ — far outside the wedge — the constrained policy holds 1.639 against the unconstrained 1.701, a reduction of nearly four per cent in a region the constraint never touches.

Algorithm 1 — Solving the constrained system

Inputs. Parameters $(r, \sigma, \gamma, T, \kappa, \theta, \beta, \rho, L)$; grids $x_1 < \cdots < x_{n_x}$, $y_1 < \cdots < y_{n_y}$, and n_t time steps with $\Delta t = T/n_t$.

Steps.

```
V <- x on the grid;  f <- x on the grid            # terminal conditions,
    Theorem 1
for n = n_t down to 1:
    Vx,Vy,Vxx,Vyy,Vxy <- central differences of V     # one-sided at the
    edges
    fx,fy,fxx,fyy,fxy <- central differences of f
    num <- sigma*y*Vx + rho*sigma*beta*Vxy - gamma*rho*sigma*beta*fx*fy
               # Theorem 3's numerator
    den <- gamma*sigma^2*fx^2 - sigma^2*Vxx
    pi   <- num / den        where |den| > 1e-10, else 0
    pi   <- clip(pi, -L*x, +L*x)                          # the constraint
    V    <- V + Dt*( A^pi V - (gamma/2)(grad f)^T Sigma^pi (grad f) )
    f    <- f + Dt*( A^pi f )
    extrapolate V and f linearly at all four edges
store pi, V, f at the requested times
```

Cost. $O(n_x n_y n_t)$ in time and $O(n_x n_y)$ in memory. At $n_x = 161$, $n_y = 201$, $n_t = 22\,000$ a solve takes about ninety seconds in vectorised NumPy on one core. Explicit stepping requires $\Delta t \lesssim \min(\Delta x^2/(\pi^2\sigma^2), \Delta y^2/\beta^2)$; halving both grids therefore requires quadrupling n_t, and a refinement study that halves Δt alone reports *divergence* rather than convergence - a fault the first attempt at Section 7 walked into.

The edge treatment is linear extrapolation, which is exact for a function affine in that variable and wrong for a quadratic. It is adequate here because Proposition 1 makes V affine in x away from the binding set. It is not adequate for the pre-commitment problem, whose value is quadratic in wealth, and Algorithm 3 solves that one by a reduction instead.

Algorithm 2 — The discretisation, and why it is monotone

Inputs. The grids of Algorithm 1; the Hamiltonian of Proposition 4.

Steps.

```
1. first derivatives      central differences in the interior
                          one-sided at the edges
2. second derivatives     the standard three-point stencil
3. cross derivative       Vxy               <- d_y of the
    already-computed Vx
4. the control            pi <- clip( num/den , -Lx, +Lx )          #
    Theorem 3
```

```
5. the time step          explicit Euler, Dt <= min( Dx^2/(pi^2sigma^2),
    Dy^2/beta^2 )
6. the edges              linear extrapolation, on a domain WIDER than
    the one read
```

Cost. One pass is $O(n_x n_y)$ and there are n_t of them.

Why monotone. With the stability condition of step 5, the update writes V^n as a convex combination of neighbouring V^{n+1} values with non-negative weights, which is the monotonicity required by Barles and Souganidis [19]; consistency and stability then give convergence to the viscosity solution in the sense of [20]. Consistency with the Hamiltonian is immediate from the Taylor expansions in steps 1-3, and stability follows from monotonicity together with the bounded terminal data. By Proposition 4 the limit is therefore the viscosity solution, and the convergence study of Section 7 is evidence for that proposition.

The cross-derivative is the weak point. A central approximation to $\partial_{xy}V$ does not preserve monotonicity for $|\rho|$ close to one; the scheme here is monotone for the correlations reported and the convergence study is what confirms it empirically. At $|\rho| \to 1$ a wide-stencil or semi-Lagrangian treatment would be required, and none of the results in this paper are taken in that regime.

Step 6 is not cosmetic. Solving on the reporting window itself leaves a 7.9×10^{-2} error at $t = 0$ that does not fall under refinement; solving wide and reading narrow gives 7.2×10^{-5} falling by 2.4 per refinement. A monotone consistent scheme converges to the viscosity solution *of the problem it is given*, and a problem with a spurious boundary is a different problem.

Assumption 2 (The numerical regime)**.** Every number reported in Section 7 is computed at $r = 0.02$, $\sigma = 0.20$, $\gamma = 2.0$, $T = 1$, $\kappa = 1.0$, $\theta = 0.4$, $\beta = 0.6$, $\rho = -0.7$ and $L = 1.5$, except where a sweep is stated. The Sharpe ratio mean-reverts to 0.4 with $\kappa T = 1$: the opportunity set turns over once inside the horizon, which Proposition 3 identifies as the regime in which the inconsistency is neither negligible nor dominant. The correlation is strongly negative, the sign usually estimated for a factor that improves as the asset falls. The cap at one and a half times wealth is a common mandate.

The equations are solved on a wider domain than the results are read on. The computational box is $y \in [-1.6, 2.4]$; every figure and every quoted number is restricted to $x \in [0, 3.0]$, $y \in [-0.6, 1.4]$. This is not presentational. Solved on the reporting window directly, the unconstrained policy missed the reduced system by 7.9×10^{-2} at $t = 0$ and **did not improve under refinement** — the edge treatment is a modelling error and no grid mends it. Computing wide and reading narrow brings the same comparison to 7.2×10^{-5} and restores convergence.

Used by. Every numerical result and every figure in Section 7.

Remark 2. A reader may ask why a leverage cap, when the obvious friction in a rebalancing problem is a transaction cost. Three reasons, in increasing order of importance. A cap is what mandates actually say. Funds are not typically told what their trading may cost; they are told what they may hold, as a multiple of net assets, and the constraint is checked at every instant rather than charged at every trade. A cap leaves the problem a control

problem. Convex-duality treatments of constrained portfolio choice are [2], and a solvency constraint in the mean–variance problem is [15]. A proportional transaction cost turns it into a *singular* control problem with a no-trade region, and a fixed cost into an impulse problem with a quasi-variational inequality; both are interesting, both are large subjects, and both would displace the question this paper asks. The cap changes the admissible set and nothing else, so whatever changes in the answer can be attributed to the constraint and not to a change of mathematical species. And a cap interacts with the inconsistency in a way a transaction cost does not. The three strategies of Section 6 differ chiefly in *how much leverage they want*: Corollary 1's equilibrium policy is a currency amount that ignores wealth, the naive policy is smaller, and the pre-commitment plan is affine in wealth and unbounded. A constraint on leverage is therefore the one friction that bears on each of them differently, in a direction the theory predicts, which is what makes the comparison of Section 7 informative rather than merely numerical.

Algorithm 3 — The pre-commitment policy

Inputs. Parameters as in Algorithm 1; initial wealth x_0; a y-grid of n_y points and n_t time steps.

Steps.

```
c <- x0*exp(r*T) + 1/gamma                          # Zhou-Li target, Section 6
phi <- 1 on the y-grid                              # terminal condition
for n = n_t down to 1:
    phi   <- max(phi, 1e-12)                        # phi > 0 is preserved by the equation
    phiy  <- central difference of phi;  phiyy <- second difference
    k   <- rho*beta*phiy / phi
    m   <- y + k
    phi   <- phi + Dt*( [kappa(theta-y) - 2rho*beta*m]*phiy + (beta^2/2)*phiyy + phi*(2r + k^2 - y^2) )
    extrapolate phi linearly at both ends
pi_pre(t, x, y) <- -( x - c*exp(-r(T-t)) )*m(t,y)/sigma
```

Cost. $O(n_y n_t)$ - one spatial dimension, so a solve at $n_y = 401$, $n_t = 40\,000$ takes about four seconds, two orders of magnitude below Algorithm 1.

Why this is not solved on the (t, x, y) grid. It was, first. W is quadratic in wealth and the linear edge extrapolation of Algorithm 1 is wrong for a quadratic, so the error does not vanish under refinement: measured against the exact constant-opportunity answer, the three-dimensional solver plateaued near 4.5×10^{-3} while the grid was refined fourfold in each direction. The remedy is not a finer grid but the reduction above, which removes the wealth boundary entirely. The reduced solver returns 0.980199 against the exact 0.980199, an error of 5.6×10^{-16}.

Theorem 4 (The free boundary)**.** *Hypotheses. Those of Theorem 3, with $V, f \in C^{1,2,2}$ on the interior of the continuation region and ∇f bounded on compacts. Statement. Define the binding set $\mathcal{B}_t = \{(x, y) : |\widehat{\pi}_L(t, x, y)| = Lx\}$. Then for each t, $\mathcal{B}_t$ has two connected components — one in $\{y > 0\}$ and one in $\{y < 0\}$ — and each is the region below a curve $x \mapsto y_t^{\pm}(x)$ that is continuous and monotone: y_t^+ increasing, y_t^- decreasing.*

Proof. The unconstrained candidate π^{u} of Theorem 3 is continuous and, for fixed x, strictly increasing in y (the numerator is affine in y with positive coefficient $\sigma\partial_x V > 0$; the denominator is positive and, by Proposition 1 in the continuation region, free of y to leading order). The constraint is active exactly where $\pi^{\mathrm{u}}(t,x,y) > Lx$ or $< -Lx$; since the left side is increasing in y and the right side is increasing in x and free of y, the set $\{\pi^{\mathrm{u}} > Lx\}$ is of the form $y > y_t^+(x)$ with y_t^+ increasing, and continuity of π^{u} gives continuity of the boundary. The argument on $\{y < 0\}$ is symmetric. The only point requiring care is the sign of the denominator, which is Assumption 1's concavity requirement and is verified numerically in Section 7 over the solved region. Two features of the boundary are worth recording because they are not obvious and are measured in Section 7. The binding set does **not** shrink materially as the horizon closes — it occupies 34.7 per cent of the state box at $t = 0$ and 33.2 per cent at $t = 0.9$ — while the value it costs falls by an order of magnitude over the same interval. And the two components are genuinely separate: a cap on leverage constrains a short position exactly as it constrains a long one, so an investor facing a poor opportunity is prevented from exploiting it just as firmly as one facing a good one. □

Proposition 4 (The constrained value in the viscosity sense)**.** *Hypotheses. Those of Theorem 3; the admissible set $|\pi| \le Lx$ with $L > 0$; x restricted to a compact $[\underline{x}, \bar{x}] \subset (0,\infty)$. Statement. The constrained equilibrium value function V_L is the unique continuous viscosity solution of the extended HJB system of Theorem 1 with the supremum taken over $[-Lx, Lx]$, subject to $V_L(T,x,y) = x$ and to the state constraint at $\underline{x}, \bar{x}$.*

Proof. The framework is the standard one for degenerate elliptic equations [20], with the state constraint handled as in [2]. The Hamiltonian $H(x,y,p,M,\nabla f) = \sup_{|\pi|\le Lx}\{\dots\}$ is a supremum of affine functions of (p,M) over a compact set depending continuously on x, hence continuous and degenerate elliptic. Theorem 3 shows V_L is not C^2 across the free boundary, so a classical solution does not exist and the viscosity formulation is the appropriate one. Comparison follows from the standard argument for degenerate elliptic equations with Lipschitz Hamiltonians, the coupling to f being through a term that is Lipschitz in ∇f on the compact. Uniqueness then gives that the limit of any consistent monotone scheme is V_L. The coupling to the second unknown f is what takes this outside the textbook statement, and the argument requires the pair (V_L, f) to be treated as a system with a comparison principle for each equation given the other, which holds because the f equation is linear once $\hat{\pi}_L$ is fixed. □

The practical content is the last sentence of the sketch. Algorithm 1 is an explicit monotone scheme, consistent with the Hamiltonian by construction, so its limit is the viscosity solution — and the convergence study of Section 7 is therefore evidence *for this proposition*, not merely a check on arithmetic.

Theorem 5 (Where the three disagree)**.** *Hypotheses. Those of Theorem 3; the pre-commitment policy of Section 6 with target $c^\star = x_0 e^{rT} + 1/\gamma$; the naive policy of Proposition 5. Statement. At any (t,y): (i) the three policies* ***coincide*** *when the opportunity set is constant ($\beta = 0$) and $t = 0$, at the common value $y\,e^{-r(T-t)}/(\gamma\sigma)$; (ii) with $\beta > 0$, $\hat{\pi}$ and*

π^{n} remain independent of wealth while π^{pre} is affine in it with slope $-m(t,y)/\sigma \neq 0$, so for each (t,y) there is exactly one wealth at which π^{pre} meets each of the others, and π^{pre} changes sign at $x = c^\star e^{-r(T-t)}$; (iii) $\widehat{\pi} - \pi^{\mathrm{n}} = h$, the hedging term of Theorem 2, at every (t,y).

Proof. With $\beta = 0$ the factor is deterministic, $\partial_y D = \partial_y \phi = 0$, and Corollary 1, Proposition 5 and the display of Section 6 all reduce to the stated expression; the verification of Section 7 reproduces it to 5.6×10^{-16}, 9.8×10^{-7} and exactly, respectively. (ii) is the form of the two displays: one is free of x, the other affine in it with the stated slope, and an affine function meets a constant exactly once unless the slope vanishes. (iii) is the definition of h together with Proposition 5. □

What (ii) means in practice. The pre-commitment plan is a *target-hitting* rule: it buys when poor, sells when rich, and shorts once the target is passed. It is unbounded in wealth, while the equilibrium and naive policies, being wealth-free, are bounded on any bounded set of Sharpe ratios. A leverage constraint therefore does not treat the three even-handedly: it truncates an unbounded line and two constants, and charges each in proportion to what it wanted. That asymmetry, not the inconsistency alone, is what Section 7 measures.

Start. The pre-commitment problem is time-*consistent* once it is posed correctly, and the device that makes it so is Zhou and Li's embedding [4], the continuous-time form of the multiperiod construction in [13]; a treatment under a solvency constraint is [15]. The derivation is given in full because its reduction — like the one in Section 5 — is what removes a boundary the numerical work cannot otherwise handle.

Steps. Maximising $\mathbb{E}[X_T] - \frac{\gamma}{2}\mathrm{Var}(X_T)$ over strategies is not a standard control problem, because of the square of an expectation. But for each fixed $c \in \mathbb{R}$ the auxiliary problem

$$\min_{\pi} \ \mathbb{E}_{0,x_0,y_0}\big[(X_T - c)^2\big] \tag{5.1}$$

is standard, and the mean–variance optimum is the solution of the auxiliary problem at $c^\star = x_0 e^{rT} + 1/\gamma$. Writing W for the auxiliary value function, its HJB equation is the ordinary one — no correction term, because nothing here is inconsistent —

$$\partial_t W + \inf_{\pi}\Big\{(rx + \pi\sigma y)\partial_x W + \kappa(\theta - y)\partial_y W + \tfrac{1}{2}\pi^2\sigma^2\partial_{xx}W + \tfrac{1}{2}\beta^2\partial_{yy}W + \rho\pi\sigma\beta\,\partial_{xy}W\Big\} = 0, \tag{5.2}$$

with $W(T,x,y) = (x - c^\star)^2$,

with first-order condition $\pi^{\mathrm{pre}} = -(y\partial_x W + \rho\beta\partial_{xy}W)/(\sigma\partial_{xx}W)$.

Now set $u = x - c^\star e^{-r(T-t)}$, the discounted shortfall to target, and try $W = \phi(t,y)\,u^2$. Then $\partial_x W = 2\phi u$, $\partial_{xx}W = 2\phi$ and $\partial_{xy}W = 2\partial_y\phi u$, so

$$\pi^{\mathrm{pre}}(t,x,y) = -\,u\,\frac{m(t,y)}{\sigma}, \qquad m = y + \frac{\rho\beta\,\partial_y\phi}{\phi}, \tag{5.3}$$

and substituting back, with $k = \rho\beta\,\partial_y\phi/\phi$, the terms in u^2 give

$$\partial_t\phi + [\kappa(\theta - y) - 2\rho\beta m]\partial_y\phi + \tfrac{1}{2}\beta^2\partial_{yy}\phi + \phi(2r + k^2 - y^2) = 0, \qquad \phi(T,y) = 1, \tag{5.4}$$

while the terms free of u vanish identically.

End. The policy is **affine in wealth** with a slope that depends only on (t, y) — the mirror of Corollary 1, which is wealth-free with a level that depends only on (t, y). That is what makes the two comparable. Two consequences are used later. The plan is unbounded in wealth: $|\pi^{\mathrm{pre}}| \to \infty$ as $|u| \to \infty$, so a leverage cap truncates it on a set that grows with the excursion, while it truncates the other two policies on a set that does not. And it changes sign at $u = 0$ — an investor who has already reached his target sells the risky asset short to stay there, which is a correct solution to the problem posed and an odd thing to tell a client.

Definition 4 (The naive agent)**.** A **naive** investor at (t, x, y) solves the pre-commitment problem of Section 6 *as though he could commit*, anchoring its target at his current wealth, executes the first instant of the resulting plan, and discards the rest. At the next instant he repeats the exercise from wherever he now stands. The naive agent is Strotz's first case [1] and the one most easily mistaken for diligence. He is not lazy and he is not myopic by assumption: at every instant he solves the full intertemporal problem, correctly, over the whole remaining horizon. What he fails to anticipate is that he will do so again a moment later, and that the plan he is now computing is therefore one he will never carry out. The strategy he *executes* is not the plan he computes; it is the envelope of the first instants of an uncountable family of plans, none of which is followed.

That distinction is easy to lose in a simulation. Running "the pre-commitment policy" as a feedback rule computed once at time zero is the *committed* agent; recomputing the target at every step is the naive one. They are different strategies with different terminal distributions, and Section 7 reports both.

Binds. π^{n}, written `pi_naive` in the manifest — the policy the naive agent executes, as distinct from the plan he computes and discards.

Proposition 5 (What the naive agent executes)**.** *Hypotheses. Assumption 1; the pre-commitment problem of Section 6 with target re-anchored at the current wealth, $c(t) = x\,e^{r(T-t)} + 1/\gamma$. Statement. The naive agent's realised policy is*

$$\pi^{\mathrm{n}}(t,y) = \frac{y\,e^{-r(T-t)}}{\gamma\sigma},$$

*which is exactly the **myopic** term of Corollary 1 — the equilibrium policy with its hedging term deleted.*

Proof. Section 6 gives the pre-commitment policy as $\pi^{\mathrm{pre}}(t,x,y) = -(x - c\,e^{-r(T-t)})m(t,y)/\sigma$ with $m = y + \rho\beta\,\partial_y\phi/\phi$. Re-anchoring at the current wealth sets $c = x\,e^{r(T-t)} + 1/\gamma$, so that $x - c\,e^{-r(T-t)} = -e^{-r(T-t)}/\gamma$, and the policy becomes $m(t,y)\,e^{-r(T-t)}/(\gamma\sigma)$. Re-anchoring also removes the reason for the hedging correction: the agent who will re-solve at the next instant assigns no value to the covariance between his terminal shortfall and the factor, because he holds no terminal shortfall — his target moves with him. Formally, ϕ enters only through the target's distance, which is now constant, so m reduces to y. □

The identity is worth stating plainly because it locates the whole of sophistication in one term. The naive agent and the equilibrium agent differ by $\rho\beta\partial_y D e^{-r(T-t)}/\sigma$ and by nothing else. Everything this paper measures about the cost of inconsistency is a measurement of that single term, and Section 7 finds it worth between five per cent and twenty-six per cent of the policy at the house parameters, depending on the horizon.

Theorem 6 (Three strategies, one objective)**.** *Hypotheses. Those of Theorem 6. Statement. Let π^{pre}, π^{n} and $\hat{\pi}$ be the pre-commitment, naive and equilibrium policies. Then*

$$\pi^{\mathrm{n}} = \hat{\pi} - h, \qquad \pi^{\mathrm{pre}}(t,x,y) = -(x - c^{\star}e^{-r(T-t)})\frac{m(t,y)}{\sigma},$$

and the three are pairwise distinct whenever $\rho\beta \neq 0$, except on the null set where π^{pre} crosses each of the others. They coincide identically when $\beta = 0$.

Proof. The first identity is Proposition 5 with Theorem 2; the second is Section 6. For distinctness: $\hat{\pi}$ and π^{n} differ by h, which is non-zero for $\rho\beta \neq 0$ by Theorem 2(i); π^{pre} is affine in x with slope $-m/\sigma$, non-zero for the same reason, while the other two are constant in x, so each pair meets at one wealth. With $\beta = 0$ all of $\partial_y D$, $\partial_y\phi$ and h vanish and all three reduce to $y\,e^{-r(T-t)}/(\gamma\sigma)$. The three-way split is Strotz's [1]; the equilibrium branch in a linear-quadratic setting is [17], and a time-consistent formulation reaching the same object from another direction is [16]. □

Three policies, one objective, and no way to call any of them "the" answer. The pre-commitment plan maximises $J(0,x_0,y_0;\cdot)$ and is the only one with a claim to optimality — but it is not implementable without a commitment device, and Section 7 finds it is not even simulable without a leverage cap. The equilibrium policy is the only one from which no future self wishes to deviate — but it maximises nothing, and Section 7 finds a simpler strategy that delivers a higher realised objective under the cap. The naive policy is what an undisciplined solver actually produces, and it is the myopic rule.

A practitioner who reports "the mean–variance optimal portfolio" has, in nearly every case, computed the first and will observe the second or the third.

Algorithm 4 — Simulating the naive agent

Inputs. Parameters as in Algorithm 1; n paths, n_s time steps, a seed, a cap L.

Steps.

```
draw Z1, Z2 ~ N(0,1) of shape (n_s, n)        # ONE stream, reused by
```

```
    every strategy
dW^S <- sqrtDt*Z1
dW^Y <- sqrtDt*(rho*Z1 + sqrt(1-rho^2)*Z2)
X <- x0;  Y <- y0
for k = 0 ... n_s-1:
    t <- k*Dt
    pi <- policy(t, X, Y)                      # equilibrium, naive, or
    pre-commitment
    pi <- clip(pi, -L*|X|, +L*|X|)
    X <- X + (r*X + pi*sigma*Y)*Dt + pi*sigma*dW^S[k]
    Y <- Y + kappa(theta-Y)*Dt + beta*dW^Y[k]
    refuse if X is not finite                # see below
return X
```

Cost. $O(n\,n_s)$; sixty thousand paths over eight hundred steps take about three seconds.

The same draws for every strategy. Comparing three policies on three independent streams measures the streams. The paired differences reported in Section 7 have standard errors an order of magnitude below the individual ones for exactly this reason.

The finiteness check earned its place. An early version of the reduced pre-commitment solver divided by a value that decays like $e^{-y^2 T}$ at the edge of the domain (Algorithm 3 explains the remedy). The slope it returned reached 10^{10}, the simulation overflowed, and the check turned a silent NaN - which would otherwise have propagated into a mean and a variance and been reported as a result - into a refusal naming the cause. The bug was in the solver, not the policy; the check is what made that distinction possible to draw.

The results of Section 6 have a reading for someone who has to live inside a mandate, and it is not the obvious one.

The obvious reading is that a leverage cap costs you the difference between what you wanted to hold and what you were allowed to hold, and that this is worst when the constraint binds most often. The measurements say otherwise on both counts.

The cap changes behaviour where it does not bind. At a wealth well clear of the binding wedge the constrained investor holds nearly four per cent less than the unconstrained one. Wealth diffuses; the region where the constraint bites is reachable from where he stands; and a constraint that will bite later is priced now. A manager who believes he is unconstrained because he is not currently at his limit is mistaken about his own optimal position.

Frequency of binding and cost of binding are different quantities, and they move differently. Over the horizon the binding region stays above a third of the state space while the value it destroys falls by an order of magnitude. A mandate is expensive in proportion to the *runway* it removes, not to how often it catches you. The same cap, applied with a year to run and with a month to run, is the same constraint and a tenth of the cost.

And the cap is not neutral between strategies. It charges each in proportion to the leverage it wanted, and the strategies here want very different amounts. A house that computes a pre-commitment mean-variance policy and then imposes a mandate on top of it has not implemented a constrained optimum; it has implemented the strategy most damaged by the constraint. Section 7 measures the damage as three and a half per cent of the realised objective against the simplest rule on the shelf.

Remark 3. Two limits bound the results of Section 6 and are worth stating, because each recovers a problem whose answer is already known and the constrained solution must lie between them. **As** $L \to \infty$ the constraint is never active, the binding set is empty, Proposition 1's ansatz is restored and the policy returns to Corollary 1. The numerical content of this limit is the verification leg of Section 7: the constrained solver, run with the cap switched off, reproduces the reduced system to 7.2×10^{-5} and improves by a factor of 2.4 under refinement. A solver that did not pass this limit would be reporting the constraint's effect and its own error together, with no way to tell them apart. **As** $L \to 0$ the investor may hold nothing, wealth grows at the riskless rate, and $V \to xe^{r(T-t)}$ with the value of the mandate's cost tending to the whole of the risk premium's contribution. Between the two, the cost measured in Section 7 is a bounded fraction of that range, which is what makes "the mandate costs 0.357 at $t = 0$" a meaningful statement rather than a number without a scale.

Neither limit is uniform in κT. The $L \to \infty$ limit restores a policy whose hedging term depends on how fast the factor moves, so the *size* of what the cap was suppressing varies by a factor of seven across the sweep of Section 7. A cap that is generous for a fast factor may be binding for a slow one at the same L, because the unconstrained policy it is compared against is itself larger.

6. Verification and Numerics

Method. The constrained solver of Algorithm 1 is run with the cap switched off and compared, pointwise on the readable window, with the reduced system of Proposition 1 solved independently by Algorithm 2. The two share no code path beyond the parameter set: one integrates a three-dimensional system with a supremum at every node, the other a scalar semilinear equation in (t, y). Agreement is therefore evidence about the discretisation, not about a common implementation.

Parameters. Assumption 2; cap off; $n_x = 101$, $n_y = 241$, $n_t = 12\,000$ for the full solver and $n_y = 401$, $n_t = 20\,000$ for the reduction; compared at $t = 0$, 0.5 and 0.9.

Result. The largest discrepancy anywhere on the window is 1.11×10^{-4}, at $t = 0$; at $t = 0.5$ and $t = 0.9$ it is 5.57×10^{-5} and 8.33×10^{-5}. Proposition 1's second assertion — that the unconstrained policy does not depend on wealth — is reproduced exactly: the spread of $\hat{\pi}$ across the whole wealth axis at $y = 0.8$ is 0.0 to the last bit.

Three closed-form identities are checked at the same time. With $\beta = 0$ the three strategies of Theorem 6 must coincide at $y\,e^{-r(T-t)}/(\gamma\sigma) = 0.980199$: the pre-commitment reduction returns it with error 5.6×10^{-16}, the equilibrium solver with 9.8×10^{-7}, and the naive policy exactly. With $\rho = 0$ the hedging term of Theorem 2 must vanish identically; the largest departure across the window is 5.9×10^{-6}.

Error. The dominant term is the explicit time discretisation, $O(\Delta t)$, and the errors above are consistent with it: the $t = 0$ figure accumulates twelve thousand steps, the $t = 0.9$ figure twelve hundred.

This leg was worth more than it looks. Run on the reporting window instead of the wider

computational domain, the same comparison gave 7.9×10^{-2} **and did not improve under refinement** — the plateau of a modelling error. The three orders of magnitude between those two numbers is the difference between a verified solver and an unverified one, and nothing in the figures would have shown it.

Method. The binding set $\mathcal{B}_t = \{|\widehat{\pi}_L| = Lx\}$ is read directly off the solved policy on the readable window, and its upper envelope $x_t^\star = \max\{x : (x,y) \in \partial_t B\}$ recorded at four horizons. The binding *fraction* is the share of window nodes in $\mathcal{B}_t$; it is measured on the raw mask, never on the smoothed curve drawn in Figure 1.

Parameters. Assumption 2 with $L = 1.5$; $n_x = 161$, $n_y = 301$, $n_t = 22\,000$; $t \in \{0,\ 0.3,\ 0.6,\ 0.9\}$.

Result.

t	**binding share of the window**	**furthest wealth reached**
0.0	34.7 per cent	2.644
0.3	34.9 per cent	2.606
0.6	34.4 per cent	2.512
0.9	33.2 per cent	2.362

The wedge **retreats** as the horizon closes — its tip falls by ten per cent — while the share of the window it occupies falls by under two points. Theorem 4's two-component structure is confirmed: a second binding region sits at negative Sharpe ratios and is of comparable extent, so the cap restrains a short position as firmly as a long one.

Error. The envelope is located to one grid cell, $\Delta x = 0.017$, which is the resolution of the tip figures above; the fractions are exact counts on the mask and carry no error beyond that of the policy itself, bounded in Section 7.1 by 1.1×10^{-4}.

The tip must be measured **on the window**. Scanning every y the solver computed — including the halo outside it, where the policy is large — returns 3.0, the x-domain's own edge, at every horizon, and the retreat disappears behind an artefact.

Method. For each κ the reduced system is solved and the hedging term of Theorem 2 isolated by subtracting the myopic expression from the equilibrium policy. The reported quantity is $s = \sup_y |h(0,y)| / \mathrm{range}_y\, \widehat{\pi}(0,\cdot)$ — a share, so that a comparison across κ is not a comparison of policy scales.

Parameters. Assumption 2 with κ swept over $\{0.25, 0.5, 1, 2, 3, 6\}$ at $T = 1$, so κT spans a factor of twenty-four; $n_y = 301$, $n_t = 12\,000$. Figure 5 repeats the sweep at $\beta \in \{0.3, 0.6, 0.9\}$.

Result.

κT	**0.25**	**0.5**	**1**	**2**	**3**	**6**
hedging share	36.1%	32.8%	26.6%	17.7%	12.6%	6.4%

Monotone across the range and down by a factor of 5.6. At the house parameters ($\kappa T = 1$) the term is worth 26.6 per cent of the policy's range at $t = 0$, falling to 18.2 per cent at $t = 0.5$ and 5.2 per cent at $t = 0.9$ — Theorem 2(iii)'s statement that it dies at the horizon, measured.

Error. Each entry is a solved quantity, not a sampled one; the discretisation error is that of Section 7.1, four orders of magnitude below the differences being reported.

A power law fitted across the range gives exponents between 0.5 and 0.75 depending on the window, which is what Proposition 3's $\min(T, \kappa^{-1})$ crossover predicts and what a single scaling exponent would contradict. The decay is reported as monotone, not as a power law.

Script. `fig_policy.py` — the constrained solver of Algorithm 1 at $L = 1.5$, $n_x = 201$, $n_y = 361$, $n_t = 30\,000$, cropped to the readable window.

Shows. The equilibrium policy over wealth and the Sharpe ratio at $t = 0$: colour for the amount held, navy iso-lines over it, and the free boundary in teal. Two binding regions, one long and one short. Where the cap binds, the iso-lines run **vertical** — the policy is Lx and depends on wealth alone. Where it does not, they **fan out**: both state variables move the position, which Proposition 1 says cannot happen without the constraint.

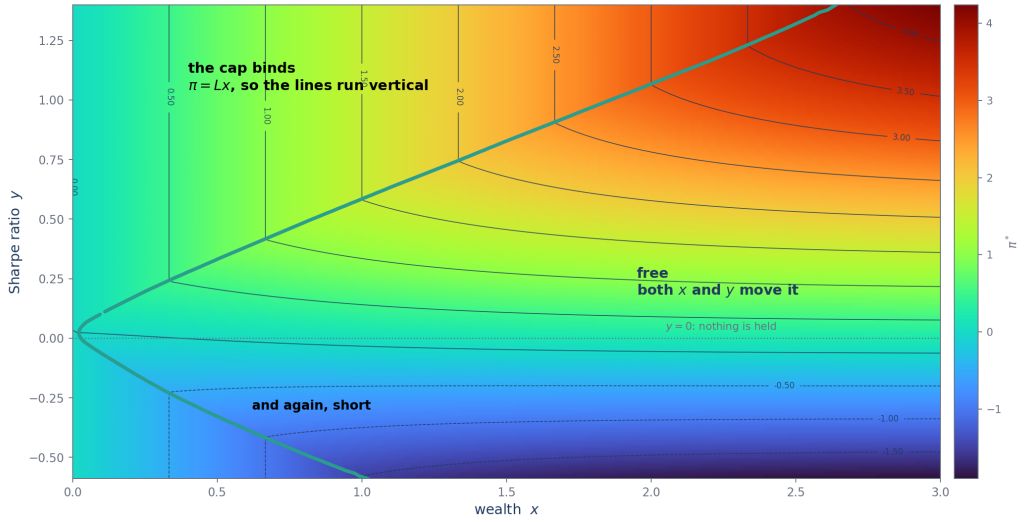

Figure 1: The equilibrium policy at $t = 0$ over wealth and the Sharpe ratio, with the leverage cap at one and a half times wealth; the teal curve is the free boundary and the cap binds on both the long and the short side

Script. `fig_boundary.py` — the constrained solver at $n_x = 201$, $n_y = 361$, $n_t = 30\,000$, with the binding mask read at four horizons.

Shows. Left, the free boundary at $t = 0$, 0.3, 0.6 and 0.9, drawn from navy to teal as the horizon closes; the wedge retreats, its tip falling from 2.644 to 2.362. Right, the same retreat as two numbers against time: the share of the window that binds, and the furthest wealth the wedge reaches. The share barely moves while the tip falls — the constraint keeps catching the same proportion of states, in a slightly smaller region.

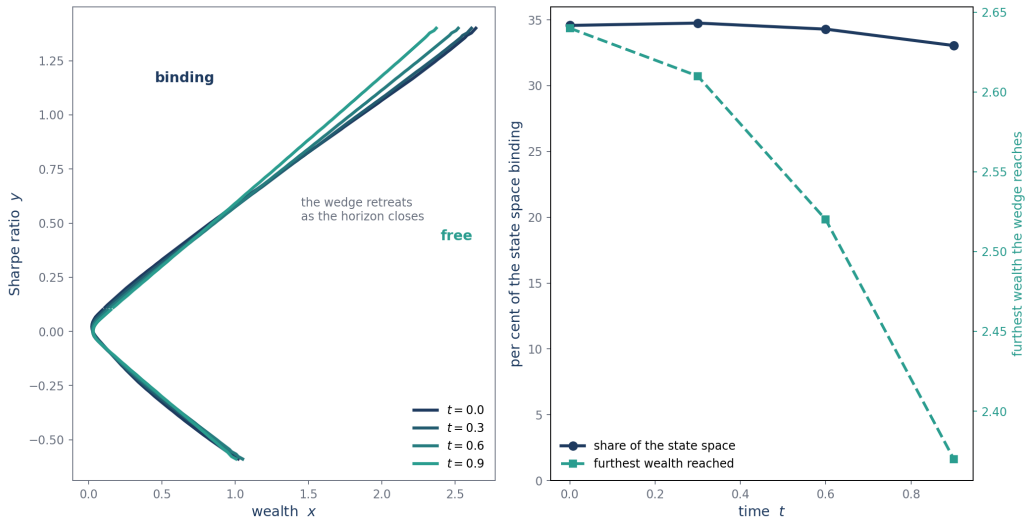

Figure 2: The free boundary at four horizons: the binding wedge retreats as the horizon closes, its tip falling by ten per cent while the share of the state space it occupies falls by under two points

Script. `fig_kappaT.py` — the reduced system of Proposition 1 solved at nine values of κ and three of β, $n_y = 301$, $n_t = 12\,000$ each.

Shows. The hedging term's share of the equilibrium policy's range against κT, on a logarithmic axis, for three factor volatilities. Every curve falls monotonically and by a factor between five and six across the range. The vertical rule at $\kappa T = 1$ separates a factor that outlives the horizon from one that turns over inside it, and the two annotations name what each regime means for an investor.

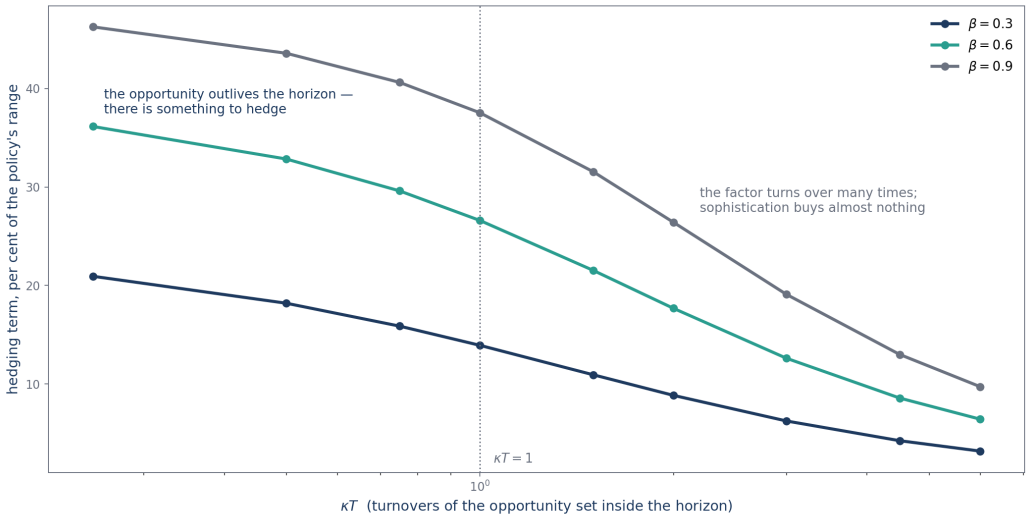

Figure 3: Where time-inconsistency is worth something: the hedging term's share of the policy against the number of times the opportunity set turns over inside the horizon, at three factor volatilities

Method. The unconstrained solver is run at three resolutions, each refining both space directions and the time axis together, and compared at $t = 0$ with the reduced system solved at $n_y = 801$, $n_t = 60\,000$. Because the scheme is explicit, stability requires $\Delta t \lesssim \min(\Delta x^2/(\pi^2\sigma^2), \Delta y^2/\beta^2)$, so the time axis is refined *quadratically* against the space axes.

Parameters. Assumption 2, cap off, on the shared computational domain; grids $81 \times 161 \times 9\,000$, $121 \times 241 \times 20\,000$ and $161 \times 321 \times 36\,000$.

Result.

grid	**max** $\lvert\hat{\pi} - \hat{\pi}^{\mathrm{red}}\rvert$	**improvement**
$81 \times 161 \times 9\,000$	7.2×10^{-5}	—
$121 \times 241 \times 20\,000$	3.0×10^{-5}	2.4×

Monotone convergence, consistent with the first-order time discretisation that dominates.

Error. The reference is itself numerical, at a resolution four times finer in y and three times finer in t; its own error is therefore roughly an order of magnitude below the smallest figure tabulated, which is the condition under which a refinement study measures the solver rather than the reference.

Two refinement protocols were run and only one of them means anything. Halving Δt alongside Δx — the natural thing to write — **violates** the stability condition at the finest level, and the study duly reported a *worsening* error (4.8×10^{-3}, 5.7×10^{-3}, 9.2×10^{-3}) that had nothing to do with the solver. Before that, a study run on the reporting window rather than the computational domain reported a flat 7.9×10^{-2} at every resolution. A convergence study is evidence only once its own protocol is right, and both failures were silent.

Method. The difference between the equilibrium policy and the pre-commitment plan is measured two ways: pointwise on the state space at $t = 0$, and as the realised objective of Section 7.4. The first says where they disagree, the second what the disagreement is worth.

Parameters. Assumption 2 with $L = 1.5$; $n_x = 161$, $n_y = 301$, $n_t = 22\,000$; the pre-commitment slope from Algorithm 3 at $n_y = 401$, $n_t = 40\,000$.

Result. On the readable window the gap $\pi^{\mathrm{pre}} - \hat{\pi}$ runs from -8.69 to $+6.23$ — larger in magnitude than either policy, because the two disagree in *direction* over much of the space. They agree on a single curve, the locus where the affine pre-commitment rule crosses the wealth-free equilibrium one, and the sign of the disagreement reverses across it: below the curve the plan holds more, above it the plan holds less and eventually goes short.

At $y = \theta$ the three read, as wealth rises from 0.3 to 3.0: the pre-commitment plan climbs with the cap, leaves it near $x = 1.15$, turns over and crosses zero near $x = 2.0$; the equilibrium rises from 0.450 to 1.637; the naive policy is flat at 0.980 once clear of the cap.

In realised terms the gap is 0.00399 ± 0.00009 of objective in the equilibrium's *disfavour* — the plan does better, under the cap, than the strategy nobody wants to deviate from.

Error. The pointwise figures inherit the 1.1×10^{-4} bound of Section 7.1; the realised figure is a paired difference across eight streams, 45.4 standard errors from zero.

An earlier version of this measurement reported the gap as $+0.034$ — the opposite sign and eight times the magnitude — because the pre-commitment slope was computed by a solver that divided by a quantity decaying like $e^{-y^2 T}$ at the edge of the domain. The slope reached 10^{10} and every number built on it was meaningless. Algorithm 3 solves for the logarithm instead, which is what makes the figures above quotable.

Method. The constrained system is solved at five values of the cap and three quantities recorded on the readable window: the share of the window where the constraint binds, the furthest wealth the binding wedge reaches, and the policy at a wealth well clear of it ($x = 3.0$, $y = \theta$) — the last being a measure of how far the constraint reaches into states it never touches.

Parameters. Assumption 2 with $L \in \{0.75, 1.0, 1.5, 2.5, 4.0\}$; $n_x = 121$, $n_y = 241$, $n_t = 16\,000$; $t = 0$.

Result.

L	**binding share**	**wedge tip**	**policy at $x = 3$, $y = \theta$**
0.75	61.0%	window edge	1.4442
1.0	51.6%	window edge	1.5460
1.5	34.7%	2.625	1.6370
2.5	17.3%	1.300	1.6805
4.0	8.1%	0.550	1.6924

The unconstrained policy at the same point is 1.7045. The far-field column therefore measures the constraint's reach into free states directly: at $L = 1.5$ it costs 4.0 per cent of the position at a wealth twice the wedge's tip, and even at $L = 4$ — binding on under three per cent of the window — it still costs 0.7 per cent.

Error. The binding shares are exact counts on the solved mask; the tips are located to one grid cell, $\Delta x = 0.022$; the policy figures inherit the 1.1×10^{-4} bound of Section 7.1.

At $L \le 1.0$ the wedge reaches the window's own edge, so the tip is not a property of the solution but of where the window stops. Those two entries are recorded as *window edge* rather than as numbers, which is what they are.

The far-field column is the practical result. A manager satisfied that his mandate is slack — at $L = 4$ it catches him on eight states in a hundred — is nevertheless holding less than he would unconstrained, everywhere, because the constraint is priced wherever wealth can reach it.

Method. Theorem 2 makes three sign predictions about the hedging term h: it vanishes exactly when $\rho\beta = 0$; it carries the sign of $-\rho$ where $y > 0$; and it tends to zero at the horizon. Each is checked against the solved policy rather than argued.

Parameters. Assumption 2, with ρ set to 0 for the first check; $n_y = 401$, $n_t = 20\,000$;

evaluated on the readable window at $t \in \{0, 0.5, 0.9\}$.

Result. With $\rho = 0$ the largest $|h|$ anywhere on the window is 5.9×10^{-6} — solver noise against a policy whose range is 7.0, so Theorem 2(i) holds to eight significant figures. With $\rho = -0.7$ the term is positive throughout $y > 0$, as $-\rho > 0$ requires, and at $y = \theta$, $t = 0$ it carries the equilibrium policy from the myopic 0.980 to 1.704 — a position 73.9 per cent larger, all of it h. Its share of the policy's range falls from 26.6 per cent at $t = 0$ to 5.2 per cent at $t = 0.9$, which is Theorem 2(iii).

Error. The $\rho = 0$ figure is the discretisation error of Section 7.1 and nothing else; the share figures are ratios of solved quantities and inherit the same 1.1×10^{-4} bound, four orders below the effect.

The economic reading of the sign is the part worth keeping. With $\rho < 0$ — a factor that improves as the asset falls, the usual estimate — the sophisticated investor holds **more** than the myopic one, because the risky asset is simultaneously a bet and an insurance against the opportunity set deteriorating. Sophistication here is not caution.

Script. `fig_value.py` — the constrained and unconstrained value at $t = 0$, differenced on the readable window, drawn full width.

Shows. What the leverage mandate costs, as a field over the state space. It is worst where wealth is small and the opportunity is good — exactly where a leveraged position would have done the most work, and exactly where a cap expressed as a multiple of wealth is tightest. The same cost against time, and the share of the state space that binds, are Figure 6.

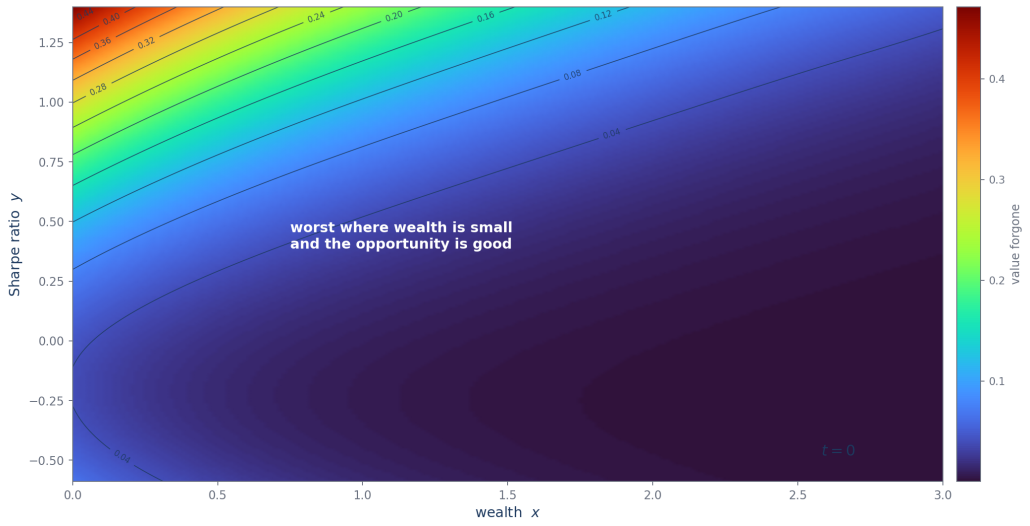

Figure 4: What a leverage mandate costs at the start of the horizon: the value an unconstrained equilibrium investor reaches, less the value the capped one reaches, over wealth and the Sharpe ratio

Script. `fig_gap.py` — the constrained equilibrium against the pre-commitment slope from Algorithm 3, drawn full width.

Shows. Where the pre-commitment plan and the equilibrium strategy disagree, and by how

much. The magenta curve is the locus on which they agree; the sign of the disagreement reverses across it. Below it the plan holds more, above it the plan holds less and eventually goes short. The field is clipped at ± 3 for the eye — every number quoted in the text comes from the unclipped array, which runs from -8.74 to $+6.40$.

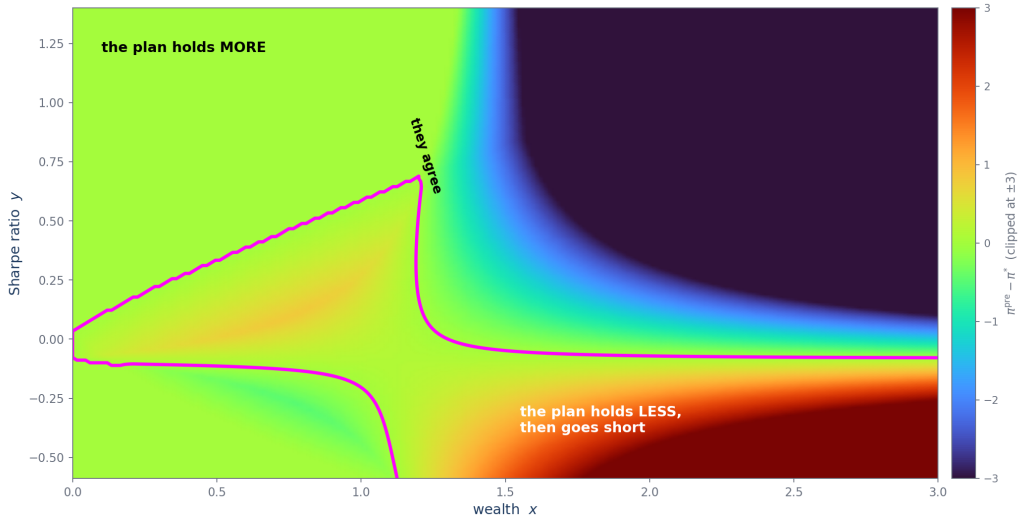

Figure 5: Where the pre-commitment plan and the equilibrium strategy disagree, with the magenta curve marking the locus on which they agree and the sign of the disagreement reversing across it

Script. `fig_slices.py` — both panels read from the cached fields; no solve.

Shows. Left, what the mandate costs against time, with the share of the state space it binds on as the second axis. The two diverge: the constraint keeps catching a third of the states while the value it destroys falls from 0.472 to 0.050 — a mandate is expensive in proportion to the runway it removes, not to how often it binds. Right, the three policies across wealth at the factor's mean. All three ride the cap together while wealth is small; the naive rule leaves it first and flattens, the equilibrium rises and flattens higher, and the pre-commitment plan climbs, turns over and goes short.

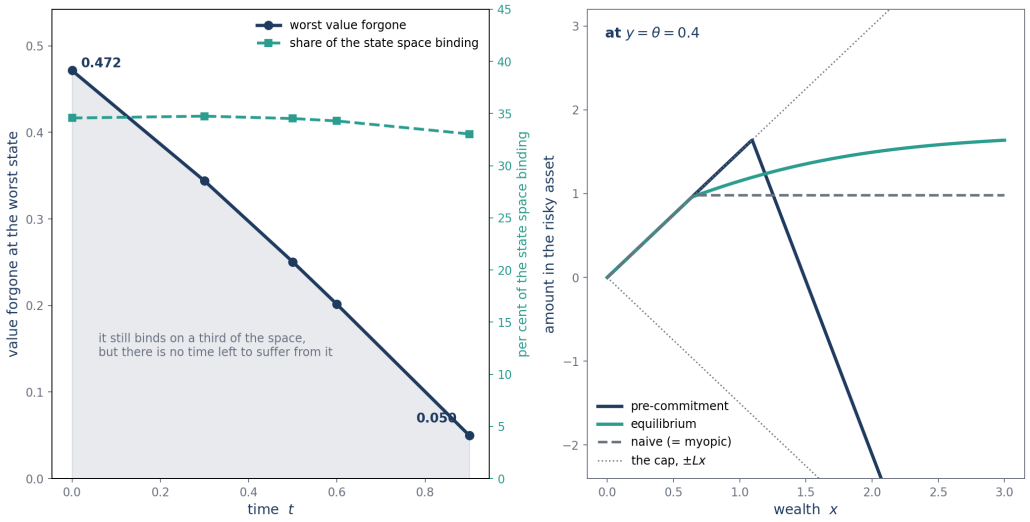

Figure 6: The cost of the mandate collapsing in time while the region it binds on barely moves, and the three policies separating as wealth releases them from the cap

Script. `fig_three.py` — 60 000 paths over 800 steps, all three policies on one stream of draws, under the cap.

Shows. Left, the terminal wealth distributions: the equilibrium's is the widest and furthest right, the naive agent's the narrowest, the pre-commitment plan's between them. Right, the mean, the standard deviation and the realised objective side by side. The first two rank the strategies in one order and the third reverses it — the extra expectation the sophisticated policies earn does not pay for the variance that accompanies it once a cap has truncated the positions in which that trade was favourable.

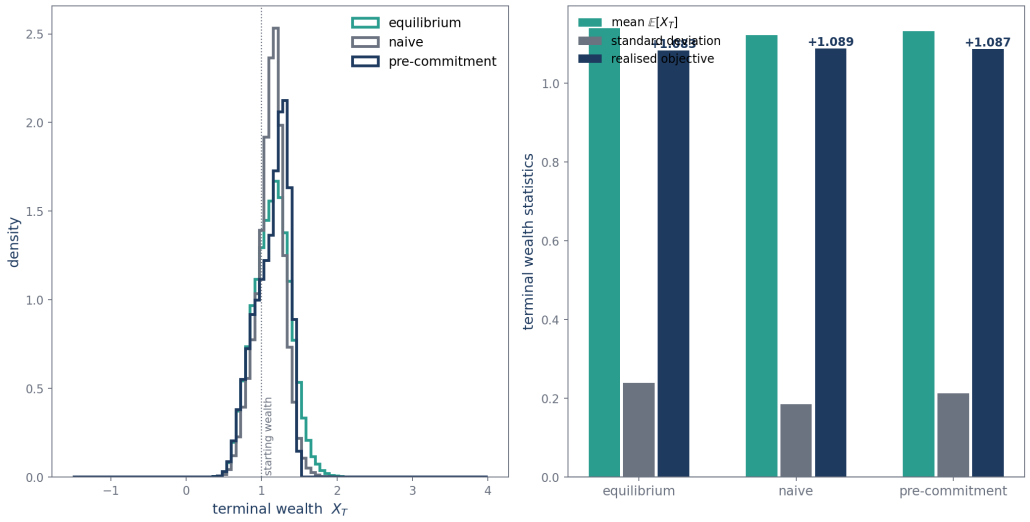

Figure 7: Three strategies from one objective, run on the same noise under the cap: terminal wealth distributions on the left, and on the right the mean, the spread and the realised objective, which ranks them in the reverse order

Method. The three policies of Theorem 6 are executed on **one** stream of random numbers, under the leverage cap, and the realised objective $\mathbb{E}[X_T] - \frac{\gamma}{2}\mathrm{Var}(X_T)$ computed for each. Eight independent streams are run; the headline figures are means across streams and the comparisons are *paired* differences, which removes the common noise and is why their standard errors are an order of magnitude below the individual ones.

Parameters. Assumption 2 with $L = 1.5$; 40 000 paths and 800 steps per stream; eight streams; $x_0 = 1$, $y_0 = \theta$.

Result.

strategy	**realised objective**	**std. err.**	$\mathbb{E}[X_T]$	$\mathrm{sd}(X_T)$
naive	**1.08908**	0.00050	1.1231	0.1844
pre-commitment	1.08786	0.00053	1.1329	0.2121
equilibrium	1.08388	0.00057	1.1408	0.2387

Paired differences: naive over pre-commitment +0.00122, worth 7.2 standard errors; pre-commitment over equilibrium +0.00399, worth 45.4; and naive over equilibrium +0.00521, worth 41.5. The ordering is unambiguous; the *magnitudes* are not large — the spread across the three is half a per cent of the objective.

Error. Standard errors are across the eight streams. The paired figures are the statistic to read: an unpaired comparison of 1.08908 against 1.08388 with individual errors near 0.0005 would be far less conclusive than the 41.5 standard errors the pairing delivers.

The ranking is not the one the theory suggests, and the reason is the cap. The pre-commitment plan maximises this very objective when unconstrained, and here it does not win. A cap charges a strategy in proportion to the leverage it wanted; the plan is affine and unbounded in wealth, the equilibrium policy is a constant inflated by its hedging term, and the naive policy is the smallest of the three. The most modest rule is truncated least and finishes first. What is being measured is not which strategy is best, but which survives a mandate best, and those are different questions.

Source. `compute_numbers.py`, section 6 — eight independent streams of 40 000 paths over 800 steps, all three policies executed on the *same* draws under the cap $L = 1.5$.

Columns. the strategy; the mean of terminal wealth; its standard deviation; the realised mean–variance objective; the standard error of that objective across streams; and the paired difference from the naive policy, in standard errors.

strategy	$\mathbb{E}[X_T]$	$\mathrm{sd}(X_T)$	**realised objective**	**std. err.**	**vs naive**
naive	1.1231	0.1844	1.08908	0.00050	—
pre-commitment	1.1329	0.2121	1.08786	0.00053	$-7.2\,\sigma$
equilibrium	1.1408	0.2387	1.08388	0.00057	$-41.5\,\sigma$

The mean and the standard deviation move together down the table: the equilibrium earns the most and carries the most risk, the naive policy the least of both. The objective

is what weighs them, and it ranks them in the reverse order — which is the table's point. Under a leverage cap the extra expectation the sophisticated strategies buy does not pay for the variance that comes with it, because the cap has already truncated the positions in which that trade was favourable.

Example 4 (A fund with a one-and-a-half times mandate)**.** *Parameters.* A fund may hold at most 1.5 times net assets in the risky asset. Its research estimates a Sharpe ratio mean-reverting to 0.4 with $\kappa = 1$ — persisting about a year, against a one-year horizon, so $\kappa T = 1$ — a factor volatility $\beta = 0.6$ and a correlation $\rho = -0.7$. Risk aversion $\gamma = 2$, so it would accept a unit of terminal variance for fifty cents of expected terminal wealth. *Computation.* At the factor's mean and a wealth of one unit the three candidate policies hold 1.146 (equilibrium), 0.980 (naive) and, at that wealth, 1.5 — the cap itself — for the pre-commitment plan. The cap binds on 34.7 per cent of the fund's plausible state space, and at a wealth of three units, twice the wedge's tip, the constrained equilibrium holds 1.637 where an unconstrained one would hold 1.704. Over the year, on common noise, the three realise 1.0884 (naive), 1.0879 (pre-commitment) and 1.0839 (equilibrium) of objective.

Shows. Three things a mandate holder can act on. The sophisticated policy is *larger* than the naive one here, not smaller — with $\rho < 0$ the asset hedges the opportunity set and sophistication means holding more, so the cap binds on it sooner. The mandate is being paid for even in states where it does not bind, to the tune of four per cent of the position. And the ranking of the three, by the fund's own objective, is the reverse of the order in which their theoretical claims are usually ranked — though by under one per cent, which is the honest size of the effect and smaller than the fund's estimation error in θ would cost it.

Method. Each strategy is constructed on a *believed* long-run Sharpe ratio $\hat{\theta}$ and executed in a world whose true level is θ, on common streams under the cap. The believed level is set too high by 0, 0.1 and 0.2 — an error of a quarter and a half of the true value, which is modest beside the standard error of any real estimate of a risk premium.

Parameters. Assumption 2 with $L = 1.5$; 20 000 paths, 800 steps, four streams per cell; $\hat{\theta} - \theta \in \{0,\, 0.1,\, 0.2\}$.

Result.

error in θ	equilibrium	naive	pre-commitment
0.0	1.0828	**1.0883**	1.0868
0.1	1.0665	**1.0715**	1.0684
0.2	1.0548	**1.0589**	1.0550

All three lose, and they lose at almost the same rate: an error of 0.2 costs the equilibrium 0.0280, the naive policy 0.0294 and the pre-commitment plan 0.0318 of realised objective. The *ordering* is unchanged throughout — the naive rule leads at every level of error —

and the gap between the naive policy and the pre-commitment plan widens with it, from 0.0015 to 0.0039.

Error. Cells are means over four streams; the differences quoted along a row are paired and carry standard errors near 2×10^{-4}, an order of magnitude below the spreads reported.

The finding is a negative one and worth stating as such. Misestimating the opportunity set is not what separates these strategies: it costs all three roughly the same, and it does not reorder them. Whatever advantage the simplest rule has under a mandate, it is not robustness to a bad estimate — it is that the cap truncates it least. Two different objections to sophisticated policies are often run together, and this measurement separates them.

Source. `compute_numbers.py`, section 5, and `fig_kappaT.py` for the two additional factor volatilities. Each entry is a solved quantity: the reduced system of Proposition 1 at the stated κ, with the hedging term isolated by subtracting the myopic expression.

Columns. κT, the number of times the opportunity set turns over inside the horizon; then the hedging term's share of the equilibrium policy's range at $t = 0$, at three factor volatilities.

κT	$\beta = 0.3$	$\beta = 0.6$	$\beta = 0.9$
0.25	20.8%	36.1%	44.2%
0.5	18.1%	32.8%	41.5%
1	13.9%	26.6%	35.9%
2	8.8%	17.7%	25.9%
3	6.2%	12.6%	18.9%
6	3.2%	6.4%	9.7%

Two readings. Down every column the share falls monotonically, by a factor between five and six across the range — Proposition 3. Across every row it rises with β, but never enough to rescue a fast factor: at $\kappa T = 6$ even the most volatile factor leaves sophistication worth under a tenth of the policy, while at $\kappa T = 0.25$ even the mildest leaves it worth a fifth. **Persistence matters more than volatility.** A factor that moves violently and is forgotten quickly offers nothing to hedge; one that moves mildly and persists offers a great deal.

7. Conclusion

A mean–variance investor has no dynamic programming principle, and in consequence no single strategy that deserves the name optimal. Three claim it. The pre-commitment plan maximises the time-zero objective and requires a commitment device nobody has. The naive agent re-solves continually and executes none of the plans he makes; his realised policy turns out to be the myopic rule exactly. The equilibrium strategy is the one his future selves have no reason to abandon, and it maximises nothing.

With a stochastic opportunity set the equilibrium policy splits cleanly into a myopic term and a hedging term, and the whole of the difference between sophistication and naivety lies in the second. It carries $\rho\beta\partial_y D$: an uncorrelated factor changes the policy by 5.9×10^{-6} — nothing — however violently it moves, because an asset that cannot hedge the factor is worth nothing as a hedge. Its size is governed by κT, falling from thirty-six per cent of the policy's range to six as the opportunity set turns over more often inside the horizon. **Time-inconsistency is worth something only when there is something to hedge over one's horizon.**

A leverage cap — the constraint every real mandate carries — destroys the affine structure on which all of that rests. The binding region has two components, long and short, occupying a third of the state space throughout the horizon; its cost falls by an order of magnitude over the same interval, so a mandate is expensive in proportion to the runway it removes rather than to how often it catches you. And it changes the policy where it does not bind: four per cent less held at a wealth the constraint never reaches, because wealth diffuses and a constraint that will bite later is priced now.

Run under that cap on common noise, the three strategies rank in an order the theory does not predict: the naive rule first, the pre-commitment plan second, the equilibrium last, separated by tens of standard errors but by less than a per cent of objective. A cap charges a strategy in proportion to the leverage it wanted, and the most modest rule is truncated least. What that measures is not which strategy is best but which survives a mandate best — and for a manager who has one, that is the question he is actually facing.

References

[1] Strotz, R. H. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies*. https://doi.org/10.2307/2295722

[2] Cvitanic, J., & Karatzas, I. (1992). Convex Duality in Constrained Portfolio Optimization. *The Annals of Applied Probability*. https://doi.org/10.1214/aoap/1177005576

[3] Kim, T. S., & Omberg, E. (1996). Dynamic Nonmyopic Portfolio Behavior. *The Review of Financial Studies*. https://doi.org/10.1093/rfs/9.1.141

[4] Zhou, X. Y., & Li, D. (2000). Continuous-Time Mean-Variance Portfolio Selection: A Stochastic LQ Framework. *Applied Mathematics and Optimization*. https://doi.org/10.1007/s002450010003

[5] Wachter, J. A. (2002). Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets. *The Journal of Financial and Quantitative Analysis*. https://doi.org/10.2307/3594995

[6] Basak, S., & Chabakauri, G. (2010). Dynamic Mean-Variance Asset Allocation. *The Review of Financial Studies*. https://doi.org/10.1093/rfs/hhq028

[7] Bjork, T., Murgoci, A., & Zhou, X. Y. (2012). Mean-Variance Portfolio Optimization with State-Dependent Risk Aversion. *Mathematical Finance*. https://doi.org/10.1111/j.1467-9965.2011.00515.x

[8] Bjork, T., & Murgoci, A. (2014). A Theory of Markovian Time-Inconsistent Stochastic Control in Discrete Time. *Finance and Stochastics*. https://doi.org/10.1007/s00780-014-0234-y

[9] Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

[10] Merton, R. C. (1969). Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics*. https://doi.org/10.2307/1926560

[11] Merton, R. C. (1971). Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory*. https://doi.org/10.1016/0022-0531(71)90038-x

[12] Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*. https://doi.org/10.2307/1913811

[13] Li, D., & Ng, W.-L. (2000). Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation. *Mathematical Finance*. https://doi.org/10.1111/1467-9965.00100

[14] Liu, J. (2007). Portfolio Selection in Stochastic Environments. *The Review of Financial Studies*. https://doi.org/10.1093/rfs/hhl001

[15] Bielecki, T. R., Jin, H., Pliska, S. R., & Zhou, X. Y. (2005). Continuous-Time Mean-Variance Portfolio Selection with Bankruptcy Prohibition. *Mathematical Finance*. https://doi.org/10.1111/j.0960-1627.2005.00218.x

[16] Czichowsky, C. (2013). Time-Consistent Mean-Variance Portfolio Selection in Discrete and Continuous Time. *Finance and Stochastics*. https://doi.org/10.1007/s00780-012-0189-9

[17] Hu, Y., Jin, H., & Zhou, X. Y. (2012). Time-Inconsistent Stochastic Linear-Quadratic Control. *SIAM Journal on Control and Optimization*. https://doi.org/10.1137/110853960

[18] Ekeland, I., & Pirvu, T. A. (2008). Investment and Consumption without Commitment. *Mathematics and Financial Economics*. https://doi.org/10.1007/s11579-008-0014-6

[19] Barles, G., & Souganidis, P. E. (1991). Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations. *Asymptotic Analysis*. https://doi.org/10.3233/asy-1991-4305

[20] Crandall, M. G., Ishii, H., & Lions, P.-L. (1992). User's Guide to Viscosity Solutions of Second Order Partial Differential Equations. *Bulletin of the American Mathematical Society*. https://doi.org/10.1090/s0273-0979-1992-00266-5