Expected-Shortfall Envelopes and Hansen–Jagannathan Kernels from Regularized Conditional Quantiles

Johannes Bleher[a,*], Joachim Grammig[b], Johannes Schief[b]

[a] *University of Hohenheim, Schloss Hohenheim, Raum: 4.31/119, 70593 Stuttgart, Germany*

[b] *University of Tübingen, Department of Statistics, Econometrics and Empirical Economics, Mohlstr. 36, 72074 Tübingen, Germany*

Abstract

This paper develops a regularized conditional quantile workflow that maps portfolio tail information into tail-state measures and, when feasible, admissible empirical stochastic discount factors (SDFs). The econometric contribution is to separate the expected-shortfall tail-probability envelope from the no-arbitrage pricing kernel satisfying normalization, pricing, and nonnegativity restrictions. The method estimates a shared portfolio and conditional quantile surface with elastic-net penalties, then uses residual ranks to construct an ES-envelope. This object is a tail probability measure before pricing; only after normalization and projection can it become an admissible nonnegative SDF. We compare it with a state-dependent Hansen–Jagannathan minimum-second-moment benchmark. In monthly Fama–French portfolios and U.S. macro states, the ES-envelope kernel shows the clearest volatility relation, while labor-market evidence is mixed. Non-forecast rolling holdout diagnostics show that raw envelopes and unconstrained HJ fits fail admissibility, whereas soft nonnegative projections sharply reduce pricing errors.

Keywords: Stochastic Discount Factor (SDF); Conditional Asset Pricing; Regularization; Quantile Optimization; Portfolio Selection; Macro-Finance Linkages; Elastic Net; Convex Optimization; Hansen–Jagannathan Distance; State-Dependent Risk Premia

JEL Classification: G12, G11, C61, E44

*Corresponding author. Tel.: +49 711 459 23054

Email addresses: johannes.bleher@uni-hohenheim.de (Johannes Bleher), joachim.grammig@uni-tuebingen.de (Joachim Grammig), johannes.schief@uni-tuebingen.de (Johannes Schief)

1. Introduction

The stochastic discount factor (SDF) is the workhorse object of modern asset pricing, pinning down present values through the fundamental restriction $E[m_{t+1}R_{t+1}] = 1$ and tracing back to consumption-based first principles (e.g., Cochrane, 2005). Empirically, two challenges recur. First, the SDF must satisfy no-arbitrage restrictions in finite samples while remaining stable in high-dimensional return spaces. Second, risk pricing varies with macroeconomic conditions, suggesting the need to incorporate conditioning information without imposing a structural preference model (Hansen and Richard, 1987).

A large literature recovers SDFs (or factor-mimicking portfolios) under mean–variance regularity, typified by the Hansen–Jagannathan (HJ) minimum-variance SDF and the associated frontier (Hansen and Jagannathan, 1991). Conditioning information complicates this exercise because it changes both the attainable payoff space and the finite-sample behavior of HJ bounds (Hansen and Richard, 1987; Ferson and Siegel, 2001, 2003). In parallel, direct portfolio-policy and quantile-risk methods connect allocation to conditioning information and downside risk without first estimating a full high-dimensional return distribution (Brandt et al., 2009; Koenker, 2005; Rockafellar and Uryasev, 2000, 2002; Acerbi and Tasche, 2002; Bonaccolto et al., 2018; Merlo et al., 2021). The recent VaR/ES forecasting literature shows that expected shortfall is not elicitable on its own, whereas the VaR–ES pair is jointly elicitable and admits strictly consistent scoring functions (Fissler and Ziegel, 2016; Taylor, 2019; Patton et al., 2019). A related macro-finance literature uses conditional quantiles to summarize downside vulnerability in macro outcomes (Adrian et al., 2019). Yet, most implementations treat portfolio optimization, tail-risk control, and SDF recovery as separate tasks, or they impose static prices of risk that abstract from macroeconomic states known to forecast returns (e.g., Lettau and Ludvigson, 2001).

We ask whether a regularized conditional quantile fit can extract downside-state information from an investable portfolio and whether that information survives the no-arbitrage restrictions required of a nonnegative SDF, therefore, satisfies the fundamental theorem of finance. To answer this question, we propose a unified, optimization-based workflow built around a *regularized conditional quantile* portfolio. Concretely, we solve a multi-τ quantile program that (i) shares one portfolio across quantiles, (ii) enforces a trapezoidal mean constraint at a target conditioning state, and (iii) applies elastic-net penalties to stabilize portfolio weights and conditional quantile coefficients. The fitted quantile surface then produces a residual ranking of tail states. From this ranking we construct a simple ES-envelope tail measure using the expected-shortfall dual representation and project its normalized version onto the no-arbitrage pricing set when exact pricing is required. We compare the resulting tail-focused kernel with a state-dependent HJ-style minimum-second-moment SDF estimated under the same data and normalization conventions. The reported implementation

is deliberately transparent: the quantile surface is linear in the standardized instruments, and flexibility comes from the quantile grid, regularization, and projection discipline rather than from a black-box SDF model. The identifying logic is correspondingly simple: the test assets and normalization define the finite-sample pricing restrictions, and the HJ criterion and ES-envelope are selection rules whose failures are reported as diagnostics rather than hidden by the estimator.

Applied to monthly Fama–French portfolios with macroeconomic conditioning, the method provides evidence that the selected pricing kernels move with macro-financial conditions, but the patterns are not uniform across states. The clearest state-dependence result is the positive ES-envelope relation with VIX. The HJ kernel varies more modestly, with the most reliable spread appearing in low term-spread states. Labor-market slack is not a stable positive tail-kernel relation in the lag-corrected sample: terciles show only mild ES elevation in high-unemployment states, while quintile spreads and robustness rows do not support a monotone positive claim. The inflation relation is non-monotone in the current application and is therefore treated as an empirical pattern rather than a maintained claim. Regularization is important because it stabilizes weights and state-price mappings in high dimensions and mitigates multicollinearity, while the multi-quantile fit imposes coherence across tails. Conceptually, our approach bridges mean–variance and tail-risk perspectives while embedding conditioning information directly in the SDF analysis.

Beyond linking SDF recovery to quantile-based portfolio choice, our framework speaks to four adjacent debates. First, the HJ-distance and misspecified-SDF literature asks how to compare pricing kernels when exact asset pricing is not guaranteed (Hansen and Jagannathan, 1997; Gospodinov et al., 2014). Recent nonparametric tests of SDF applicability make the same point from a different angle: pricing failures can identify the boundary of the payoff span on which a class of kernels remains empirically credible (Pezzo et al., 2024). Our rolling design follows that discipline by reporting projection failures, negative affine kernels, and separating-portfolio certificates as economic diagnostics rather than hiding them inside a fit statistic. Second, regularization is now central to empirical asset pricing because high-dimensional test assets and candidate predictors make unconstrained estimates unstable. Large cross-section SDF estimators face this problem directly, while shrinkage-based SDF estimation and penalized factor methods explicitly trade small bias for lower sampling error (Kim and Korajczyk, 2024; Brodie et al., 2009; Freyberger et al., 2020; Kozak et al., 2020; Lettau and Pelger, 2020). Recent work on smart SDFs develops a related no-arbitrage framework with convex pricing constraints and SDF regularization under market frictions (Korsaye et al., 2025). Our elastic-net treatment extends this idea to a joint quantile-portfolio program, stabilizing both the asset weights and the state-dependent quantile surface. Third, flexible SDF estimators increasingly use no-arbitrage restrictions as an estimation criterion while relying on characteristic-based or machine-learning representations of states and payoffs (Kelly et al., 2019; Gu et al., 2020, 2021; Chen et al., 2024). This literature is also part of a broader discussion about when com-

plexity improves return measurement (Kelly et al., 2024). At the same time, recent evidence on characteristics-based SDFs cautions that machine-learning success in return measurement need not imply strong macroeconomic content in the non-market component of the SDF (Rytchkov and Zhong, 2025). Our contribution is complementary: it keeps the state representation deliberately low-dimensional and interpretable, but makes the admissibility and projection restrictions explicit and treats the macro-state patterns as empirical diagnostics rather than structural preference estimates. Fourth, downside-risk and tail-risk-premium papers ask whether bad states carry distinct compensation in the cross-section or in aggregate return predictability (Ang et al., 2006; Bollerslev et al., 2015; Farago and Tedongap, 2018). Quantile and tail-sensitive preferences have also been used directly for portfolio allocation in models of pessimistic or ambiguity-averse investors, most notably via Choquet expected utility and quantile-oriented loss functions (Bassett et al., 2004). A closely related option-implied strand uses state prices of conditional quantiles to summarize how markets price distributional states and to forecast equity premia (Metaxoglou and Smith, 2017a,b). More broadly, option-implied state-price and recovery approaches infer risk aversion or physical probabilities from derivatives (Ait-Sahalia and Lo, 2000; Ross, 2015). Our object differs in both source and restriction: we do not infer state prices from option-implied risk-neutral distributions, recover physical probabilities from derivatives, or estimate a structural downside-risk premium. Instead, we construct an ES-envelope tail measure from portfolio residual ranks and then ask whether projection can make it an admissible nonnegative SDF in the traded test-asset span. By constructing this ES-envelope tail measure and projecting it when pricing is required, we provide a transparent, data-driven kernel that summarizes tail emphasis and can be made compatible with no-arbitrage pricing restrictions when the nonnegative pricing set is feasible.

The closest econometric perspective is the literature that treats asset-pricing models as restrictions to be estimated and evaluated, rather than as calibrated narratives alone. Grammig and Kuechlin (2018) and Soenksen and Grammig (2021) use simulation-based and indirect-inference designs to assess structurally motivated consumption-based pricing kernels, while Grammig et al. (2025) compare theory-based and machine-learning risk-premium measurements on a common forecasting target. We share the emphasis on explicit restrictions, loss functions, and validity diagnostics, but our object is different. Instead of estimating preference parameters or a return-premium forecast, we estimate an empirical tail-state measure and ask whether projection can turn that measure into an admissible SDF in the observed test-asset span. This positioning is important because it makes the ES-envelope a measurement device for state-price mass, not a structural model of preferences.

This paper contributes in three ways. The novelty is not the HJ benchmark itself; it is the discipline that maps a tail probability envelope from the quantile portfolio into the SDF space and makes that tail-focused selection comparable to a conventional HJ selection. The paper therefore follows a moment-based empirical-asset-pricing discipline: define the restrictions, make the loss and selec-

tion rules explicit, and evaluate where the selected object satisfies or violates the restrictions in and out of sample. First, we introduce a joint conditional quantile–portfolio program with elastic-net regularization that yields an investable portfolio, coherent VaR and return-tail ES reports, and a conditional quantile surface on a user-chosen grid. Second, we turn the expected-shortfall residual envelope into a disciplined SDF construction by showing exactly when the raw tail measure is only a tail-state diagnostic and when its projection is an admissible nonnegative pricing kernel satisfying $E[m_{t+1}R^e_{t+1}] = 0$ with either $E[m_{t+1}] = 1$ or $E[m_{t+1}R_{f,t+1}] = 1$. Third, we use Fama–French portfolios and macroeconomic state variables to compare the state allocations of a minimum-second-moment HJ selection and a tail-focused ES-envelope selection under the same normalization and rolling diagnostic design. Our framework complements recent advances in conditional and flexible SDF estimation–including macro-conditioned kernels (Hansen and Richard, 1987; Lettau and Ludvigson, 2001), characteristic-based conditional factor models (Kelly et al., 2019), machine-learning SDFs (Gu et al., 2020; Chen et al., 2024), and factor methods penalizing pricing errors (Lettau and Pelger, 2020)–by providing a transparent, optimization-first bridge from tail-risk control to no-arbitrage SDF recovery.

The remainder of the paper is structured as follows. Section 2 defines the quantile-portfolio program and the two SDF constructions. Section 3 describes the macro-finance application and empirical design. Section 4 reports in-sample, state-dependence, and rolling holdout evidence before the conclusion summarizes what the framework establishes and what remains open.

2. Methodology

2.1. Data Objects and Notation

Let $R_t \in \mathbb{R}^k$ denote the k-vector of *excess* asset returns at time $t = 1, \ldots, T$, and let $X_t \in \mathbb{R}^m$ collect m predetermined state variables (including an intercept). We stack the data into matrices

$$\mathbf{R} := \begin{bmatrix} R_1^\top \\ \vdots \\ R_T^\top \end{bmatrix} \in \mathbb{R}^{T\times k}, \qquad \mathbf{X} := \begin{bmatrix} X_1^\top \\ \vdots \\ X_T^\top \end{bmatrix} \in \mathbb{R}^{T\times m}.$$

Let e_d denote the d-vector of ones; in particular, e_k is used for portfolio budget constraints and e_T for sample normalizations. We seek portfolio weights $w \in \mathbb{R}^k$ (with budget and box constraints) and a conditional quantile surface $q_\tau(x) = x^\top \beta(\tau)$, where $x \in \mathbb{R}^m$ denotes a generic realized conditioning-state vector, including the intercept, and $\beta(\tau) \in \mathbb{R}^m$ varies with $\tau \in (0, 1)$. Let $\mu_R \in \mathbb{R}^k$ denote the expected excess-return vector used in the target-return constraint, and let

$\mu_0 \in \mathbb{R}$ denote the scalar minimum expected excess return required of the selected portfolio. In the baseline implementation we set $\mu_R = \bar{r} := \mathbb{E}_T[R_t]$ and require $\mu_R^\top w \geq \mu_0$.

2.2. Quantile-Portfolio Program and Expected-Shortfall Link

For a single quantile level τ, define residuals $\varepsilon_{\tau,t} := R_t^\top w - X_t^\top \beta(\tau)$ and their positive/negative parts $\varepsilon_{\tau,t}^+, \varepsilon_{\tau,t}^- \geq 0$ such that $\varepsilon_{\tau,t} = \varepsilon_{\tau,t}^+ - \varepsilon_{\tau,t}^-$. The empirical check loss is

$$\mathcal{L}_\tau(w, \beta(\tau)) \;=\; \sum_{t=1}^{T} \big[\, \tau\, \varepsilon_{\tau,t}^+ + (1-\tau)\, \varepsilon_{\tau,t}^- \,\big].$$

Minimizing $\mathcal{L}_\tau$ jointly over $(w, \beta(\tau))$ under linear constraints yields both the conditional τ-quantile $q_\tau(X_t)$ and optimal portfolio weights at that quantile level. Here X_t denotes the observed conditioning state in the sample, while x^* denotes a chosen evaluation state at which the fitted conditional quantile surface is reported. As discussed in Koenker (2005), the check loss is tightly connected to tail risk; we report the implied $\mathrm{VaR}_\tau(x^*) = x^{*\top}\beta(\tau)$ at the evaluation state x^* and the associated return-tail $\mathrm{ES}_\tau^{\mathrm{return}}(x^*)$ via the sample identity

$$\mathrm{ES}_\tau^{\mathrm{return}}(x^*) \;=\; \mathrm{VaR}_\tau(x^*) \;+\; \frac{1}{1-\tau}\,\frac{1}{T}\sum_{t=1}^{T} \varepsilon_{\tau,t}^+.$$

The second term averages the positive residual exceedances above the fitted quantile and rescales them by the tail probability $1-\tau$; it therefore converts the quantile level into the mean return in the corresponding tail. The sign convention matters for the SDF interpretation. The expression above is written in the return convention and uses positive residuals relative to the fitted return quantile. The pricing-kernel construction below uses the loss convention instead: at ES confidence level τ, residual losses are $\ell_t = -r_t$, the lower-tail probability mass is $1-\tau$, and the ES-envelope assigns probability to the lowest $1-\tau$ fraction of residuals. Thus the empirical choice $\tau = 0.95$ corresponds to a five percent residual-loss tail, not to a claim that the raw return quantile alone is an admissible SDF.

2.3. Multi-Quantile Surface and Mean Constraint

Rather than estimating $\beta(\tau)$ at a single τ, we estimate a grid $\{\beta(\tau_j)\}_{j=1}^{J}$ for $\boldsymbol{\tau} = \{\tau_1, \ldots, \tau_J\} \subset (0,1)$ by solving a *joint* program that shares w across quantiles:

$$\begin{aligned}
\min_{w,\{\beta(\tau_j),u_j,v_j\}_{j=1}^{J}} \quad & \sum_{j=1}^{J}\sum_{t=1}^{T}\left[\,\tau_j\,u_{j,t} + (1-\tau_j)\,v_{j,t}\,\right] \\
& + \lambda_w^{(1)}\|w\|_1 + \tfrac{\lambda_w^{(2)}}{2}\|w\|_2^2 \\
& + \sum_{j=1}^{J}\left(\lambda_\beta^{(1)}\|\beta(\tau_j)\|_1 + \tfrac{\lambda_\beta^{(2)}}{2}\|\beta(\tau_j)\|_2^2\right) \qquad (1)
\end{aligned}$$

$$\text{s.t.} \quad \mathbf{R}w - \mathbf{X}\beta(\tau_j) - u_j + v_j = 0, \qquad u_j \geq 0,\ v_j \geq 0, \quad j = 1, \ldots, J, \qquad (2)$$

$$e_k^\top w = 1, \qquad \mu_R^\top w \geq \mu_0, \qquad \underline{w} \leq w \leq \overline{w}, \qquad (3)$$

$$\mathbf{X}\beta(\tau_j) \;\leq\; \mathbf{X}\beta(\tau_{j+1}), \qquad j = 1, \ldots, J-1 \quad \text{(non-crossing)}. \qquad (4)$$

The rationale for a single investable portfolio w is that sharing weights across quantiles avoids rebalancing to different tail-specific portfolios and therefore corresponds more closely to the allocation a practitioner actually holds. For a fixed portfolio, minimizing the summed check losses across a grid of quantile levels $\{\tau_j\}$ fits a conditional quantile surface over several parts of the support rather than at one tail point alone. The portfolio constraints and penalties mean that this is not an unrestricted distributional estimator; it is a regularized, implementable compromise between quantile calibration, tail control, and portfolio feasibility. Hence, the joint program fits one portfolio whose fitted quantile surface is disciplined across the quantile grid while preserving the restrictions needed for an investable allocation.

The non-crossing condition is not needed for a single quantile, but it is useful in the multi-quantile program because it prevents the fitted conditional quantiles from being ordered incoherently at the observed conditioning states.

The reported implementation imposes the non-crossing constraints (4) on the estimation sample, so the fitted quantile surface is monotone over the chosen grid at the observed conditioning states; it is not a global monotonicity guarantee for every possible covariate value outside the observed design. To link the quantile surface to the *mean* return at a specific state x^*, we add a grid-level mean-link calibration constraint

$$\sum_{j=1}^{J-1} \tfrac{1}{2}\left(x^{*\top}\beta(\tau_j) + x^{*\top}\beta(\tau_{j+1})\right)\Delta\tau_j \;\geq\; \mu_0, \qquad \Delta\tau_j := \tau_{j+1} - \tau_j, \qquad (5)$$

which integrates the fitted quantiles over the chosen grid and stabilizes the tail fit against the target mean. With a sufficiently dense grid spanning the full unit interval, the trapezoidal expression approximates $\int_0^1 q_\tau(x^*)\,d\tau$. In the reported finite-grid implementation, and especially in the compact baseline grid, endpoints outside the grid are not modeled; the restriction is therefore not a full conditional-mean identity over unobserved quantile support.

If only ℓ_1 penalties are present (no ℓ_2) and no quadratic pieces are active, (1)–(3) reduces to a linear program. We also allow an optional cardinality module, which turns the problem into a mixed-integer linear program. When ℓ_2 penalties are used, the continuous formulation is solved as a convex quadratic program.

2.4. Regularization and Asset/Factor Selection

We allow distinct penalties on portfolio weights and on the quantile coefficients:

$$\lambda_w^{(1)}\|w\|_1 + \tfrac{\lambda_w^{(2)}}{2}\|w\|_2^2 \quad \text{and} \quad \lambda_\beta^{(1)}\|\beta(\tau_j)\|_1 + \tfrac{\lambda_\beta^{(2)}}{2}\|\beta(\tau_j)\|_2^2.$$

This interpretation follows the elastic-net argument of Zou and Hastie (2005): the ℓ_1 component induces sparsity, while the ℓ_2 component stabilizes selection among correlated regressors by shrinking related coefficients jointly. In the portfolio-weight block, the same quadratic shrinkage also discourages concentrated positions and is therefore related to regularized Markowitz portfolio construction (Brodie et al., 2009).

The same framework can also impose discrete portfolio restrictions when the problem is kept linear. For example, a long-only portfolio with a minimum number of active positions can be written with binary selection variables $z_i \in \{0, 1\}$:

$$0 \le w_i \le U_i z_i, \quad w_i \ge L_i z_i, \quad \sum_{i=1}^{k} z_i \ge K_{\min}.$$

Here $z_i = 0$ forces asset i out of the portfolio, while $z_i = 1$ allows a positive weight between the lot size L_i and the upper bound U_i. The last inequality requires at least $K_{\min}$ selected assets. This optional cardinality module is used only in the linear program/mixed integer linear porgram (LP/MILP) case, because adding binary variables to a formulation with active ℓ_2 penalties would turn the problem into a mixed-integer quadratic program (MIQP). Linear group restrictions of the form $a^\top w \in [\underline{b}, \overline{b}]$ can be added in the same linear case to bound exposures to regions, sectors, or other asset groups.

2.5. Recovery of the Empirical SDF: Two Complementary Constructions

We report two complementary empirical stochastic discount factors (SDFs) from our estimation routine. The pricing moments identify a set of admissible state-price vectors rather than a unique kernel. Our two constructions therefore choose different representatives from the same empirical discipline: the HJ object is the minimum-second-moment representative, while the ES-envelope object starts from the tail probability measure implied by the quantile residuals and then selects the closest admissible pricing kernel. This distinction matters for interpretation. State dependence in the reported kernels is not a claim that the test assets alone identify a unique marginal utility process; it is a statement about how two economically motivated selections allocate value across the observed macro states.

Maintained restrictions for the finite-sample pricing set.

The projection results below are conditional on the empirical pricing design rather than on a complete-market structural model.

Assumption 1 (Finite-sample projection design)**.** *The observed months are treated as the finite state space for the sample projection. The marketed payoff span is the span of the reported test assets and the normalization payoff; unspanned Arrow state claims are not assumed to be traded. The normalization vector ρ is strictly positive, with $\rho = e_T$ for mean-one normalization and $\rho = R_f$ for risk-free normalization. An exact reported SDF must be nonnegative and must satisfy the selected absolute normalization and excess-return pricing equalities in the relevant sample window. When this nonnegative pricing set is empty, a soft projection remains a penalized approximation rather than an exact SDF.*

This assumption is deliberately finite-sample. It lets us ask whether the observed test-asset span admits a nonnegative state-price vector in a given window, but it does not assume that every state-contingent payoff is traded or that the sample projection point-identifies marginal utility. It also clarifies the arbitrage language below: a separating payoff is a sample certificate for the observed test-asset span, while a negative component of an unconstrained kernel is only a constructive arbitrage when the corresponding nonnegative payoff is marketed or replicated.

(A) Minimum-second-moment (HJ) SDF with conditioning.

Given excess returns, we select a linear, state-dependent SDF by solving a Hansen–Jagannathan minimum-second-moment program. The benchmark follows the HJ and conditioning-information asset-pricing tradition (Hansen and Jagannathan, 1991; Hansen and Richard, 1987; Cochrane,

2005). The pricing restriction used here is unconditional in the empirical test-asset span: the sample moments impose $T^{-1}\mathbf{R}^\top m = 0_k$ and $T^{-1}\rho^\top m = 1$, where ρ is the selected positive normalization payoff. The empirical application uses $\rho = R_f$; the mean-one normalization is the special case $\rho = e_T$. The word "state-dependent" refers to the representation $m_{t+1} = X_t^\top c - R_t^\top \gamma$, which lets the selected kernel vary with lagged macro instruments, not to a claim that we identify the full conditional moment restriction $E[m_{t+1}R^e_{t+1} \mid X_t] = 0$ at every state. This distinction keeps the interpretation tied to the data: the test assets and instruments select an admissible empirical kernel inside the observed span, while the state coefficients summarize how that selected kernel covaries with macro conditions. The program is

$$\begin{aligned}
\min_{m\in\mathbb{R}^T,\, c\in\mathbb{R}^m,\, \gamma\in\mathbb{R}^k} \quad & \sum_{t=1}^{T} m_{t+1}^2 + \lambda_\gamma^{(1)}\|\gamma\|_1 + \tfrac{\lambda_\gamma^{(2)}}{2}\|\gamma\|_2^2 \\
& \quad + \lambda_c^{(1)}\|c\|_1 + \tfrac{\lambda_c^{(2)}}{2}\|c\|_2^2 && (6)\\
\text{s.t.} \quad & m \;=\; \mathbf{X}c \;-\; \mathbf{R}\gamma, && (7)\\
& \tfrac{1}{T}\rho^\top m \;=\; 1, \qquad \tfrac{1}{T}\mathbf{R}^\top m \;=\; 0_k, && (8)\\
& m \;\geq\; 0. && (9)
\end{aligned}$$

where (7) specifies the linear SDF $m_{t+1} = X_t^\top c - R_t^\top \gamma$, and (8) imposes the selected normalization and no-arbitrage pricing restrictions for excess returns. Because an admissible pricing kernel should be nonnegative, the reported specification also imposes $m_{t+1} \geq 0$. The objective (6) minimizes the second moment of m (the Hansen–Jagannathan criterion under the selected pricing restrictions), with optional elastic-net penalties on the prices of risk γ and state-price coefficients c for stability and sparsity. When the normalization fixes $E[m]$, this is the usual minimum-variance HJ problem up to a constant; with the gross risk-free normalization used empirically, we report $\mathrm{Var}(m)$ as a dispersion diagnostic alongside the optimized second moment. This construction is well established; see also Gospodinov et al. (2014) for econometric analysis. We report $(m_{t+1})_{t\leq T}$, γ, c, the selected normalization error $|T^{-1}\rho^\top m - 1|$, risky-asset pricing residuals $T^{-1}\mathbf{R}^\top m$, and $\mathrm{Var}(m)$.

Proposition 1 (Negative fitted kernels and arbitrage interpretation)**.** *Let $\mathcal{M} \subseteq \mathbb{R}^T$ denote the finite-sample payoff span, where each coordinate $s = 1, \ldots, T$ corresponds to one sample state. Let a candidate kernel m define the linear price functional*

$$q_m(y) = \frac{1}{T}\sum_{s=1}^{T} m_s y_s, \qquad y \in \mathcal{M}.$$

If there exists a marketed payoff $y \in \mathcal{M}$ with $y \geq 0$, $y \neq 0$, and $q_m(y) < 0$, then m prices an arbitrage. Here $y \geq 0$ is componentwise, and $y \neq 0$ means that the payoff is strictly positive in at least one sample state; requiring $y > 0$ in every state would be stronger than the usual arbitrage

condition. In particular, if the Arrow payoff e_s for a state with $m_s < 0$ is traded, or is replicated by the test-asset span, then the negative component implies such an arbitrage. If that state payoff is not in $\mathcal{M}$, a negative fitted component is not by itself a constructive arbitrage in the observed test assets. It is nevertheless inadmissible as a stochastic discount factor on an expanded state-by-state payoff space; we therefore use negative unconstrained kernels only as diagnostics and report the nonnegative constrained or projected kernel.

Proof. The first claim follows directly from the definition of an arbitrage. If $q_m(y) < 0$, an investor can buy the marketed payoff y and receives $-q_m(y) > 0$ in cash at the trade date because the price is negative. Since $y \geq 0$ and $y \neq 0$, the future payoff is never negative and is strictly positive in at least one state. The trade therefore has positive initial cash flow and no downside payoff risk, which is an arbitrage. If an Arrow payoff e_s is in $\mathcal{M}$, then $q_m(e_s) = m_s/T$, so $m_s < 0$ gives a negative price for a nonnegative payoff. If no marketed payoff isolates, or replicates exposure to, the state with negative fitted mass, the sign of that component alone does not identify a trade in the observed payoff span. It only shows that the candidate price functional is not positive on the full state-by-state cone. □

Karush–Kuhn–Tucker (KKT) conditions, link to the HJ frontier, and the role of penalties.

Consider the constrained problem in (6)–(9) with excess returns and constraints

$$T^{-1}\rho^\top m = 1, \qquad T^{-1}\mathbf{R}^\top m = 0_k, \qquad m = \mathbf{X}c - \mathbf{R}\gamma, \qquad m \geq 0.$$

Let $\alpha \in \mathbb{R}$, $\zeta \in \mathbb{R}^k$ be the multipliers for the two pricing/normalization constraints, $\eta \in \mathbb{R}^T$ for the defining constraint, and $\nu \in \mathbb{R}^T_+$ for the nonnegativity inequalities written as $-m \leq 0$. The Lagrangian is

$$\begin{aligned}\mathcal{L}(m, c, \gamma) = \sum_{t=1}^{T} m_{t+1}^2 + \lambda_\gamma^{(1)}\|\gamma\|_1 + \tfrac{\lambda_\gamma^{(2)}}{2}\|\gamma\|_2^2 + \lambda_c^{(1)}\|c\|_1 + \tfrac{\lambda_c^{(2)}}{2}\|c\|_2^2 \\ + \alpha\Big(\tfrac{1}{T}\rho^\top m - 1\Big) + \zeta^\top\Big(\tfrac{1}{T}\mathbf{R}^\top m\Big) + \eta^\top(m - \mathbf{X}c + \mathbf{R}\gamma) - \nu^\top m.\end{aligned}$$

First-order (KKT) conditions are:

$$\partial_m: \quad 2m + \tfrac{\alpha}{T}\rho + \tfrac{1}{T}\mathbf{R}\zeta + \eta - \nu = 0, \qquad \partial_c: \quad -\mathbf{X}^\top\eta + \lambda_c^{(2)}c + \lambda_c^{(1)}s_c = 0,$$

$$\partial_\gamma: \quad \mathbf{R}^\top\eta + \lambda_\gamma^{(2)}\gamma + \lambda_\gamma^{(1)}s_\gamma = 0,$$

$$m = \mathbf{X}c - \mathbf{R}\gamma, \qquad \tfrac{1}{T}\rho^\top m = 1, \qquad \tfrac{1}{T}\mathbf{R}^\top m = 0_k, \qquad m \geq 0, \quad \nu \geq 0, \quad \nu \circ m = 0.$$

Here $s_c \in \partial\|c\|_1$ and $s_\gamma \in \partial\|\gamma\|_1$ are subgradient vectors for the ℓ_1 penalties. Componentwise, $(s_c)_j = \text{sign}(c_j)$ when $c_j \neq 0$ and $(s_c)_j \in [-1, 1]$ when $c_j = 0$; s_γ is defined analogously for γ. For observations where the nonnegativity constraint is inactive, eliminating η using ∂_m shows that m lies in the affine span of $\{\rho, \mathbf{R}, \mathbf{X}\}$ with shrinkage on (c, γ) governed by $(\lambda^{(1)}, \lambda^{(2)})$; at boundary observations, the complementary-slackness term ν sets the kernel to zero. When the penalties are zero, $\mathbf{X}$ is absent, $\rho = e_T$, and the inequality is omitted or inactive, this collapses to the classic Hansen–Jagannathan minimum-variance SDF: minimize $\text{Var}(m)$ subject to $E[m] = 1$ and $E[mR] = 0$, yielding the HJ frontier in $(\text{Var}(m), \|E[mR]\|)$ space. With $\rho = R_f$, the same second-moment program imposes the gross risk-free pricing normalization used in the empirical application. With ℓ_2 penalties, c and γ are shrunk toward zero (a bias that generally changes the selected second moment and reported dispersion but can reduce sampling error); ℓ_1 induces sparsity in state prices and priced directions without affecting feasibility, since the pricing equalities are imposed as hard constraints in-sample. In practice, penalty parameters can be chosen by holdout or rolling windows that evaluate holdout pricing errors and normalization errors while preserving exact hard constraints on the estimation sample.

(B) ES-envelope tail measure and projected SDF from residual ranks

In addition, we construct a tail-focused probability measure using the optimality conditions of quantile/expected-shortfall programs and then turn it into a pricing kernel only after normalization and, when needed, projection. Expected shortfall and conditional value-at-risk are equivalent risk functionals in this setting; we use ES-envelope as the paper-facing label for the tail probability measure. The derivation below defines the residual-loss tail, maps it into ES dual weights, and then separates the raw tail measure from the pricing-feasible projected SDF.

Derivation and normalization.

Let $r_t := R_t^\top \hat{w} - X_t^\top \hat{\beta}(\tau)$ denote the residuals from the joint quantile fit at level $\tau \in (0, 1)$ and define the lower-tail loss as $\ell_t = -r_t$. The Rockafellar–Uryasev (RU) empirical ES program for the loss sequence $\{\ell_t\}_{t=1}^T$ is The idea is to choose a loss threshold α and then average only the losses that exceed this threshold. The slack variable z_t records the positive excess loss above α: if $\ell_t \leq \alpha$, then $z_t = 0$; if $\ell_t > \alpha$, then $z_t = \ell_t - \alpha$ at the optimum. The objective therefore adds the threshold to the average excess loss in the upper $1 - \tau$ loss tail. Minimizing over α selects the empirical tail threshold and yields the empirical expected shortfall.

$$\text{ES}_\tau \;=\; \min_{\alpha,\, z \in \mathbb{R}^T} \Big\{ \alpha + \frac{1}{(1-\tau)T} \sum_{t=1}^{T} z_t \;\;:\;\; z_t \geq \ell_t - \alpha,\; z_t \geq 0 \;\; (t = 1, \ldots, T) \Big\}.$$

The corresponding linear-programming dual (Rockafellar and Uryasev, 2000, 2002; Acerbi and Tasche, 2002) is The dual rewrites expected shortfall as a worst-case weighted average of realized losses. The weights p_{t+1} form a probability distribution over sample states, while the upper bound $1/\{(1-\tau)T\}$ prevents the entire probability mass from being placed on one loss observation. At the optimum, these weights concentrate on the largest losses, so the dual makes explicit which sample states receive tail probability mass.

$$\mathrm{ES}_\tau = \max_{p\in\mathbb{R}^T} \Big\{ \sum_{t=1}^{T} p_{t+1}\ell_t \;\; : \;\; 0 \le p_{t+1} \le \frac{1}{(1-\tau)T},\; \sum_{t=1}^{T} p_{t+1} = 1 \Big\}.$$

At an optimal solution, the dual variables p_{t+1} saturate the upper bound on the largest losses, equivalently the lowest residuals, and are zero otherwise except possibly at the boundary. Let $h = (1-\tau)T$, set $\bar{p} = 1/h$, and let $a = r_{(\lceil h \rceil)}$ denote the residual cutoff for the lower tail. Define $B = \{t : r_t < a\}$ as the set of observations strictly inside the residual-loss tail and $C = \{t : r_t = a\}$ as the set of observations tied at the cutoff. We use the canonical threshold convention that residuals in B receive the cap $\bar{p}$, while the remaining probability mass is split equally across the tied observations in C. This yields the *ES–envelope measure*

$$p_{t+1}^{\mathrm{env}} \;=\; \bar{p}\,\mathbb{1}\{r_t < a\} \;+\; \frac{1-\bar{p}|B|}{|C|}\,\mathbb{1}\{r_t = a\},$$

which is nonnegative by construction, sums to one, and respects the ES dual cap. To use this tail measure as a stochastic discount factor under a selected positive normalization payoff ρ, we define

$$m_{t+1}^{\mathrm{raw}}(\rho) \;=\; \frac{T p_{t+1}^{\mathrm{env}}}{\rho_t}, \qquad \text{so that} \qquad \frac{1}{T}\rho^\top m^{\mathrm{raw}}(\rho) = 1.$$

Mean-one normalization is the special case $\rho = e_T$; the empirical application uses the gross risk-free payoff, $\rho = R_f$. The raw envelope is then projected, if necessary, onto the selected nonnegative pricing set

$$\mathcal{A}_+(\rho) \;=\; \Big\{ m \in \mathbb{R}_+^T : \tfrac{1}{T}\mathbf{R}^\top m = 0_k,\; \tfrac{1}{T}\rho^\top m = 1 \Big\}.$$

For $\rho = e_T$, this reduces to $\mathcal{A}_+ = \{m \in \mathbb{R}_+^T : T^{-1}\mathbf{R}^\top m = 0_k,\; T^{-1}e_T^\top m = 1\}$, and feasibility is equivalent to $0_k \in \mathrm{conv}\{R_1, \ldots, R_T\}$, that is, if and only if the origin lies in the convex hull of the observed excess-return vectors. More generally, under a positive normalization vector ρ_t such as a gross risk-free return, feasibility of $\mathcal{A}_+(\rho)$ is equivalent to $0_k \in \mathrm{conv}\{R_1/\rho_1, \ldots, R_T/\rho_T\}$. The data therefore identify the admissible set $\mathcal{A}_+(\rho)$, and the projection below identifies the closest element of that set to the selected-normalized tail envelope. Here $\Pi_{\mathcal{A}_+(\rho)}$ denotes the Euclidean projection operator onto $\mathcal{A}_+(\rho)$. Because this set includes the inequality constraint $m \ge 0$, the map from m^{raw} to m^{proj} is generally not a linear matrix projection; it is a convex, piecewise-affine projection obtained by solving a quadratic program. With this normalization in place, the reported

L^2 projection $m^{\text{proj}} = \Pi_{\mathcal{A}_+(\rho)}(m^{\text{raw}}(\rho))$ solves

$$\min_{m \in \mathcal{A}_+(\rho)} \|m - m^{\text{raw}}(\rho)\|_2^2.$$

If the nonnegativity restriction is omitted, the affine projection onto

$$\mathcal{A}(\rho) \;=\; \left\{ m \in \mathbb{R}^T : \tfrac{1}{T}\mathbf{R}^\top m = 0_k,\ \tfrac{1}{T}\rho^\top m = 1 \right\}$$

has the closed form

$$m^{\text{aff}} \;=\; m^{\text{raw}}(\rho) - \mathbf{R}a - \rho\, b,$$

where

$$\begin{bmatrix} \frac{1}{T}\mathbf{R}^\top\mathbf{R} & \frac{1}{T}\mathbf{R}^\top\rho \\ \frac{1}{T}\rho^\top\mathbf{R} & \frac{1}{T}\rho^\top\rho \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \;=\; \begin{bmatrix} \frac{1}{T}\mathbf{R}^\top m^{\text{raw}}(\rho) \\ \frac{1}{T}\rho^\top m^{\text{raw}}(\rho) - 1 \end{bmatrix}.$$

This linear system is the normal equation for the equality-only projection: it chooses (a, b) so that m^{aff} satisfies the risky-asset pricing moments and the selected normalization. It does not impose nonnegativity. The nonnegative projection m^{proj} instead solves the quadratic program over $\mathcal{A}_+(\rho)$. Because $m^{\text{raw}}(\rho)$ arises from the ES dual on residuals r_t rather than from the asset-pricing moment conditions, *pricing errors need not be zero for* $m^{raw}(\rho)$. They are zero *by construction* only after projection: we report the nonnegative m^{proj} in tables/figures when exact pricing ($\frac{1}{T}\mathbf{R}^\top m = 0$) is required.

Proposition 2 (Tail-envelope SDF and pricing projection)**.** *Fix $\hat{w}$, $\hat{\beta}(\tau)$, the residuals $r_t = R_t^\top \hat{w} - X_t^\top \hat{\beta}(\tau)$, lower-tail losses $\ell_t = -r_t$, and a strictly positive normalization vector ρ. The ES dual defines a tail probability vector p^{env} that solves the empirical expected-shortfall dual and concentrates mass on residual tail states. The raw vector $m_{t+1}^{raw}(\rho) = Tp_{t+1}^{env}/\rho_t$ is nonnegative and selected-normalized, but it is not generally an SDF for the test assets because $T^{-1}\mathbf{R}^\top m^{raw}(\rho)$ need not equal zero. The set $\mathcal{A}_+(\rho)$ is nonempty if and only if $0_k \in \text{conv}\{R_1/\rho_1, \ldots, R_T/\rho_T\}$. Whenever $\mathcal{A}_+(\rho)$ is nonempty, $m^{proj} = \Pi_{\mathcal{A}_+(\rho)}(m^{raw}(\rho))$ exists, is unique, nonnegative, and satisfies $T^{-1}\mathbf{R}^\top m^{proj} = 0_k$ and $T^{-1}\rho^\top m^{proj} = 1$. It is therefore the unique closest nonnegative pricing-feasible SDF to the raw tail envelope in Euclidean distance.*

Proof. The ES dual feasible region is the probability simplex with upper bounds $p_{t+1} \le [(1-\tau)T]^{-1}$, so every optimizer is nonnegative and sums to one; optimal mass is assigned to the largest losses, equivalently the lowest residuals, with the remaining boundary mass split across residuals tied at the cutoff. Setting $m_{t+1}^{\text{raw}}(\rho) = Tp_{t+1}^{\text{env}}/\rho_t$ therefore gives $T^{-1}\rho^\top m^{\text{raw}}(\rho) = 1$ and $m^{\text{raw}}(\rho) \ge 0$. For the pricing set, $m \in \mathcal{A}_+(\rho)$ if and only if $p_t = m_{t+1}\rho_t/T$ satisfies $p_t \ge 0$, $\sum_t p_t = 1$, and $\sum_t p_t R_t/\rho_t = 0_k$. This is exactly the statement that 0_k belongs to the convex hull of the normalized observed excess-return vectors; the converse follows by setting $m_{t+1} = Tp_t/\rho_t$ for any such convex

weights. When this condition holds, $\mathcal{A}_+(\rho)$ is a nonempty closed convex subset of the compact selected-normalization set $\{m \geq 0 : \rho^\top m = T\}$. The objective $\|m - m^{\text{raw}}(\rho)\|_2^2$ is continuous and strictly convex in m, so the projection exists and is unique. Because the minimizer belongs to $\mathcal{A}_+(\rho)$, it is nonnegative and satisfies the pricing and normalization equalities by construction. □

Proposition 3 (Soft projection and exact pricing)**.** *Let $\rho \in \mathbb{R}^T$ be a strictly positive normalization vector, with $\rho = e_T$ for mean-one normalization and $\rho = R_f$ for risk-free normalization. Define*

$$\mathcal{C}_+(\rho) = \{m \in \mathbb{R}_+^T : T^{-1}\rho^\top m = 1\}, \qquad \mathcal{A}_+(\rho) = \{m \in \mathcal{C}_+(\rho) : T^{-1}\mathbf{R}^\top m = 0_k\}.$$

For any candidate kernel m^0 and penalty $\lambda > 0$, the soft nonnegative projection

$$m_\lambda^{soft} = \underset{m \in \mathcal{C}_+(\rho)}{\arg\min} \left\{ \|m - m^0\|_2^2 + \lambda \|T^{-1}\mathbf{R}^\top m\|_2^2 \right\}$$

exists and is unique. If $\mathcal{A}_+(\rho)$ is nonempty, then $m_\lambda^{soft} \to \Pi_{\mathcal{A}_+(\rho)}(m^0)$ as $\lambda \to \infty$. If $\mathcal{A}_+(\rho)$ is empty, the soft projection remains nonnegative and normalized, but it is not an exact pricing SDF because the risky-asset pricing residual cannot be zero.

Proof. The set $\mathcal{C}_+(\rho)$ is nonempty, closed, convex, and bounded because $\rho_t > 0$ for all t. The objective is continuous and strictly convex in m because it contains $\|m - m^0\|_2^2$, so a unique minimizer exists. When $\mathcal{A}_+(\rho)$ is nonempty, let $m^* = \Pi_{\mathcal{A}_+(\rho)}(m^0)$. Optimality of m_λ^{soft} gives

$$\|m_\lambda^{\text{soft}} - m^0\|_2^2 + \lambda \|T^{-1}\mathbf{R}^\top m_\lambda^{\text{soft}}\|_2^2 \leq \|m^* - m^0\|_2^2,$$

because the pricing residual of m^* is zero. Hence $\|T^{-1}\mathbf{R}^\top m_\lambda^{\text{soft}}\|_2 \to 0$. Every convergent subsequence has a limit in $\mathcal{A}_+(\rho)$ with distance to m^0 no larger than $\|m^* - m^0\|_2$, and uniqueness of the Euclidean projection implies that this limit is m^*. Compactness of $\mathcal{C}_+(\rho)$ then gives convergence of the full sequence. If $\mathcal{A}_+(\rho)$ is empty, no vector can satisfy nonnegativity, normalization, and zero risky-asset pricing errors simultaneously; the soft solution is therefore a penalized approximation rather than an exact SDF. □

Corollary 1 (Infeasibility certificate)**.** *If $\mathcal{A}_+(\rho)$ is empty, then a separating portfolio exists: there are portfolio weights $a \in \mathbb{R}^k$ and a margin $\delta > 0$ such that*

$$a^\top R_t/\rho_t \geq \delta \qquad \text{for all } t = 1, \ldots, T.$$

Since $\rho_t > 0$, the same portfolio has strictly positive excess returns in every observed state of the sample window. When the observed months are treated as the finite state space and the test assets are frictionlessly tradable, this is a sample arbitrage certificate. In the empirical analysis we use

it more cautiously as a finite-window feasibility diagnostic: it explains why no nonnegative pricing kernel can exactly price the holdout test assets in that window, without claiming a population arbitrage outside the observed scenarios.

Proof. The generalized convex-hull condition follows from $p_t = m_{t+1}\rho_t/T$, which is a probability vector under the normalization $T^{-1}\sum_t m_{t+1}\rho_t = 1$. If the origin is outside the closed convex hull of $\{R_t/\rho_t\}_{t=1}^T$, the finite-dimensional separating-hyperplane theorem gives a and $\delta > 0$ such that $a^\top R_t/\rho_t \geq \delta$ for every observed state after rescaling. Multiplication by $\rho_t > 0$ preserves strict positivity of the excess payoff in the original return units. □

Empirical implementation.

In our implementation it is built by ranking residuals r_t, assigning the dual cap $1/[(1-\tau)T]$ to residuals strictly below the cutoff, splitting the remaining mass across residuals tied at the cutoff, normalizing the resulting tail probability measure with the selected payoff ρ (the risk-free gross return in the empirical application), and projecting when exact pricing is required. We refer to the raw probability-weighted object as the ES-envelope and to the projected pricing-feasible object as the *ES-envelope SDF*. This distinction keeps the tail-state measurement step separate from the no-arbitrage pricing step. The object is interpretable, computationally trivial, and places mass on low-residual states that are candidates for high empirical discount-factor weight.

Part (A) is deliberately a benchmark: it extends the HJ second-moment selection to a state-dependent, regularized setting. The paper's distinct step is part (B), which constructs a simple, tail-focused measure aligned with expected-shortfall duality and the residual structure of our quantile program, then asks whether projection can make it an admissible SDF. Empirically we report both m^{HJ} and the projected ES-envelope SDF m^{ES}, compare their pricing residuals $\mathbf{R}^\top m$, and study their macro-state dependence.

This distinction gives the empirical objects an explicit status, summarized in Table 1. The raw ES-envelope is an identified tail probability measure from residual ranks; it is diagnostic unless it also satisfies the pricing moments. An unconstrained or affine pricing fit is likewise diagnostic if it violates nonnegativity. A soft nonnegative projection is a useful holdout approximation when exact pricing is infeasible, but it is not an exact SDF. A reported SDF is admissible only when it is nonnegative and satisfies the selected normalization and excess-return pricing moments, either by direct constrained estimation or by projection onto $\mathcal{A}_+$.

Table 1: Status of the Empirical Kernel Objects

Object	Selection rule	Pricing status	Role in the paper
Raw ES-envelope	ES dual tail probability measure from quantile residual ranks	Nonnegative and selected-normalized, but risky-asset pricing errors need not be zero	Tail-state diagnostic
Unconstrained or affine fit	HJ or projection criterion without nonnegativity	May price assets but can assign negative state prices	Diagnostic only
Soft nonnegative projection	Nonnegative normalized vector with penalized pricing errors	Preserves nonnegativity and normalization; exact pricing is approximate	Holdout stress test
Reported HJ or ES-envelope SDF	Direct constrained solve or L^2 projection onto $\mathcal{A}_+$	Nonnegative, normalized, and exactly prices the test assets in the relevant sample window	Admissible empirical SDF

Population interpretation and sampling uncertainty.

The projection results above are finite-sample statements. Their population counterpart is a pricing-moment set defined by the maintained test-asset span and information set. Let Z_{t+1} denote the payoff-relevant state vector at date $t+1$, including any information on which the population pricing kernel is allowed to depend. With excess returns R^e_{t+1}, a population admissible kernel belongs to

$$\mathcal{A}^0_+ = \left\{ m(Z_{t+1}) \geq 0 : E[m(Z_{t+1})] = 1,\ E[m(Z_{t+1})R^e_{t+1}] = 0_k \right\}.$$

The sample set $\mathcal{A}_+$ replaces these moments by sample analogues over the observed window, and the ES-envelope estimates a tail probability measure generated by the fitted residual ranking. The paper does not claim point identification of marginal utility. It studies whether economically interpretable selections from the admissible set, especially the projected ES-envelope and HJ benchmark, are stable across penalty choices, state partitions, bootstrap resamples, and rolling holdout windows. Sampling uncertainty enters through three channels: the return moments that define the pricing set, the residual ranks that define the tail envelope, and the regularization and asset-penalty choices that select among admissible kernels. The moving-block bootstrap and holdout diagnostics are therefore interpreted as stability checks for state-dependence summaries and pricing feasibility, not as full asymptotic inference for a structural SDF.

2.6. Implementation Details

The quantile grid is user-specified. The in-sample baseline reported below uses the compact grid $\boldsymbol{\tau} = \{0.50, 0.75, 0.95\}$ for computational reproducibility, while the rolling holdout diagnostics use

the dense grid $\{0.05, 0.10, \ldots, 0.95\}$ as a robustness run. The reported tail setting $\tau = 0.95$ denotes the ES confidence level for the residual-loss envelope; the corresponding tail probability entering the envelope weights is $1 - \tau = 0.05$. Penalty parameters $(\lambda^{(1)}, \lambda^{(2)})$ can be chosen by cross-validation or fixed at stable elastic-net values; ℓ_1 penalties may be activated to induce sparsity in the portfolio weights or the conditioning coefficients. One possible tuning rule is to choose penalties by holdout performance. Let λ collect the candidate penalty values for the portfolio, quantile-surface, and SDF steps, and let m_λ denote the kernel obtained after estimating the corresponding specification on a training block. The fitted object is then evaluated, without re-estimation, on the next holdout block. The holdout score is

$$\mathcal{S}(\lambda) = \widehat{ES}_{\text{hold}}(\lambda) + \left\| \widehat{E}_{\text{hold}}[m_\lambda R^e] \right\|_2 + \left| \widehat{E}_{\text{hold}}[m_\lambda R_f] - 1 \right| + \widehat{E}_{\text{hold}}[(-m_\lambda)_+] \,.$$

Lower values of $\mathcal{S}(\lambda)$ are preferred. The four terms penalize holdout tail loss, risky-asset pricing errors, risk-free normalization error, and negative holdout kernel mass. Thus the selector favors penalties that produce a stable tail fit and a kernel that remains close to admissible on data not used for estimation. In the empirical application below we report a fixed baseline penalty vector and then vary the penalties in the robustness grid, rather than choosing the main specification by minimizing $\mathcal{S}(\lambda)$ on the same state-dependence patterns used for interpretation. When multiple τ levels are integrated to impose the mean constraint, the trapezoidal rule over the chosen grid provides a transparent finite-grid approximation. Because endpoints outside the grid are omitted in the reported implementation, the mean-link restriction should be read as a grid-level calibration at x^*, not as a full conditional-mean identity over unobserved quantile support.

First, pure ℓ_1 specifications and optional cardinality constraints are handled as linear or mixed-integer linear programs. Second, elastic-net formulations and SDF estimation are handled as continuous convex quadratic programs. The optimization path is chosen by the penalty structure and constraint set. For SDF estimation, the Hansen–Jagannathan minimum-second-moment SDF, the raw ES-envelope tail measure, and the projected ES-envelope SDF are produced as separate empirical objects.

Budget $e_k^\top w = 1$ and target return $\mu_R^\top w \geq \mu_0$ are imposed in-sample. Box constraints $\underline{w} \leq w \leq \overline{w}$ and optional no-short-selling restrictions $w \geq 0$ are supported. Group exposure constraints of the form $a^\top w \in [\underline{b}, \overline{b}]$ (e.g., sector or region limits) can be included. In the reported SDF subproblem, normalization follows either $\frac{1}{T}\sum_t m_{t+1} = 1$ or $\frac{1}{T}\sum_t m_{t+1} R_{f,t} = 1$, and nonnegativity $m_{t+1} \geq 0$ is imposed. Unconstrained kernels are useful diagnostics but are not admissible pricing kernels when they take negative values. The projection routines post-check finite values, nonnegativity, the selected normalization, and the exact excess-return pricing moments; a vector that fails these checks is not passed on to the reported diagnostics.

State variables X_t are standardized and lagged by one period so month-t returns are paired with the previous-month macro state. Because the empirical design uses currently available public histories rather than real-time data vintages, this timing convention should be read as historical-state alignment, not as a fully real-time trading signal. When daily or weekly macroeconomic series are used, they are first aggregated to monthly frequency (using means or end-of-month values) before merging with portfolio returns. A constant term is automatically prepended to X_t to ensure an intercept in the conditional quantile system.

2.7. Outputs and Inference

The implemented procedure returns four objects:

(i) optimal portfolio weights w,

(ii) estimated $\mathrm{VaR}_\tau(x^*)$ and return-tail $\mathrm{ES}_\tau^{\mathrm{return}}(x^*)$ at a chosen quantile and state,

(iii) the conditional quantile surface $\{\beta(\tau_j)\}_{j=1}^J$ on the chosen grid,

(iv) two complementary pricing-kernel objects:

- the Hansen–Jagannathan minimum-second-moment SDF $(m_{t+1}^{\mathrm{HJ}}, \gamma, c)$ with diagnostics, and
- the ES-envelope tail measure, derived from residual ranks and converted into m_{t+1}^{ES} after the normalization and projection steps.

Pricing accuracy is evaluated via mean absolute pricing errors $T^{-1}\|\mathbf{R}^\top m\|_1$ (expected to be near zero) and the Hansen–Jagannathan variance proxy $\mathrm{Var}(m)$. Macro-state dependence is analyzed by conditioning m_{t+1} on quantiles or bins of selected X_t, and by inspecting γ for priced risk directions and c for state sensitivities. Comparing m_{t+1}^{HJ} and m_{t+1}^{ES} shows how tail-based and mean-variance pricing diagnostics differ across macroeconomic regimes.

2.8. Rolling Holdout Diagnostic Design

The in-sample pricing errors reported below are feasibility diagnostics: they show that the estimated kernels satisfy the pricing restrictions imposed by the optimization or projection step on the estimation sample. They do not, by themselves, establish performance away from the estimation window. We therefore use a rolling holdout diagnostic. For each origin, the quantile portfolio and SDF coefficients are estimated on a trailing window, and the fitted HJ representation $m_{t+1} = X_t^\top c - R_t^\top \gamma$ is evaluated on the subsequent holdout window. The holdout diagnostics

record risky-asset pricing errors, risk-free normalization errors, kernel variance, and any violations of nonnegativity. In parallel, the trained portfolio and quantile surface are used to rank holdout residuals; the ES-envelope is then rebuilt from those holdout residual ranks and evaluated as a raw nonnegative tail kernel. The ES rows are therefore sample-separated holdout pricing diagnostics, not ex-ante forecasts of the tail state-price vector. For both HJ and ES objects, we report a nonnegative soft projection that preserves risk-free normalization while penalizing risky-asset pricing errors. When the holdout nonnegative pricing set is nonempty, we also report the exact projected pricing kernel. If exact nonnegative pricing is infeasible on a holdout block, the failure is reported rather than silently dropping the origin. This distinction is important because exact pricing on a holdout block is itself a feasibility restriction and may fail when the test assets admit no nonnegative state-price vector on that block.

This design separates four objects. Training-only estimates determine the shared quantile portfolio, the conditional quantile surface, and the HJ coefficients. Holdout evaluation then prices the subsequent return block with those fitted HJ coefficients and records pricing and nonnegativity failures. The ES-envelope row uses the training portfolio and quantile surface but constructs tail probability mass from the ordering of holdout residuals; it is therefore a sample-separated tail-state diagnostic, not a real-time forecast of the state-price vector. The soft and exact projections are post-holdout admissibility audits: they ask how much pricing error remains after imposing nonnegativity and the selected risk-free normalization, and whether the exact nonnegative pricing set is empty.

3. Empirical Design

We use monthly test assets from the Fama–French library (the 25 size $\times$ value portfolios) and the risk-free rate to construct excess returns $R_t \in \mathbb{R}^k$, $k = 25$. Macroeconomic state variables are public FRED series measuring the consumer price index, industrial production, the unemployment rate, the 10-year Treasury yield, the 3-month T-bill rate, and the VIX. Daily yield and volatility series are aggregated to monthly means, and monthly macroeconomic series are used at month-end levels. After transforming and lagging the macro variables, we keep the intersection with available Fama–French returns. The aligned baseline sample runs from February 1990 through January 2026 and contains 431 monthly observations; all in-sample tables and figures use this effective aligned panel. Relative to the complete calendar span, the effective panel omits November 2025 because the lagged state for that return month uses the October 2025 macro row, where the inflation and unemployment inputs are incomplete. The one-month lag removes contemporaneous return-month state information from the reported in-sample alignment, but it does not by itself address publication lags or later data revisions in the public macro histories. For this reason, we interpret the macro-state and holdout evidence as historical diagnostics of the selected pricing kernels, not

as a real-time-vintage trading experiment.

Table 2 reports the sample-construction summary used by the empirical analysis. The table makes the effective date window, state timing, test-asset span, and reporting state explicit before the results are interpreted.

Table 2: Sample Construction Summary

Notes: The table uses the same aligned return and state matrices as the empirical analysis. The omitted-month row compares the effective aligned months with the complete monthly calendar between the first and last included observations.

Design item	Baseline value
Effective aligned sample	February 1990–January 2026
Monthly observations	431
Calendar months in span	432
Omitted month inside span	November 2025
Reason for omitted month	November 2025: incomplete lagged macro state from October 2025 (year-over-year inflation, unemployment rate)
Test assets	25 Fama–French 5×5 size–value excess-return portfolios
Macro state variables	Inflation, Industrial production growth, Unemployment rate, Term spread, VIX
State timing	standardized and lagged one month
Risk-free series	Fama–French risk-free gross return
Reporting state x^*	January 2026 final aligned lagged state

We form transformations

$$\begin{aligned} \text{12-month inflation}_t &= 100\left(\frac{P_t}{P_{t-12}} - 1\right), \\ \text{12-month industrial-production growth}_t &= 100\left(\frac{I_t}{I_{t-12}} - 1\right), \\ \text{Treasury term spread}_t &= Y_t^{10y} - Y_t^{3m}, \end{aligned}$$

where P_t is the price level, I_t is industrial production, and Y_t^{10y} and Y_t^{3m} are the long and short Treasury yields. We collect the state vector

$$\begin{aligned} Z_t = (&\text{12-month inflation}_t,\ \text{12-month industrial-production growth}_t, \\ &\text{unemployment rate}_t,\ \text{Treasury term spread}_t,\ \text{VIX}_t)^\top. \end{aligned}$$

Each column is standardized in-sample and lagged one month so returns are conditioned on the

previous-month macro state:

$$X_t = \left(1,\ \tilde{Z}_{t-1}^{\top}\right)^{\top}, \qquad \tilde{Z}_t = \frac{Z_t - \bar{Z}}{s(Z)}.$$

$\{R_t\}_{t=1}^{T}$ and $\{X_t\}_{t=1}^{T}$ are aligned on a common monthly index and use $\bar{r} = \frac{1}{T}\sum_t R_t$ as the in-sample expected return vector.

We estimate the in-sample joint quantile program (Section 2) on $\tau \in \{0.50, 0.75, 0.95\}$, sharing one portfolio w across τ and enforcing the grid-level mean-link constraint via the trapezoidal rule at the current state x^*. For the reported baseline, $x^* = X_T$, the final available observed state vector in the aligned lagged state matrix $\mathbf{X}$ (equivalently, the last row of $\mathbf{X}$ written in column-vector convention), namely the standardized lagged macro state associated with January 2026. This choice fixes the VaR/ES reporting point and the portfolio target-return constraint; the SDF pricing restrictions, state-dependence tables, and figures use all 431 aligned monthly observations. We then focus on $\tau = 0.95$ for reporting $\mathrm{VaR}_\tau(x^*)$ and return-tail $\mathrm{ES}_\tau^{\mathrm{return}}(x^*)$. These two reported quantities summarize the fitted portfolio-return quantile at x^* and the associated positive-residual return-tail average. They are separate from the ES-envelope probability weights, which are built later from the lower residual-loss tail $\ell_t = -r_t$ and then normalized and projected for pricing. Portfolio constraints include budget $e_k^{\top} w = 1$, a target mean $\bar{r}^{\top} w \geq \mu_0 = 0.003$ (per month), box constraints $0 \leq w_i \leq 0.20$, and long-only ($w \geq 0$). The reported baseline fixes the elastic-net penalties ex ante: for portfolio weights, $\lambda_w^{(1)} = 0$ and $\lambda_w^{(2)} = 10^{-3}$; for conditioning coefficients, $\lambda_\beta^{(1)} = 0.4$ and $\lambda_\beta^{(2)} = 10^{-3}$. Thus the baseline uses quadratic shrinkage for the portfolio weights and combines sparsity with mild quadratic shrinkage for the conditioning effects. These values are a transparent design choice, not estimates of an optimal tuning rule; Table 11 reports lower and higher penalty variants to show which conclusions survive this choice.

Table 3 reports the fitted quantile-program quantities that anchor the tail-ranking step. The baseline solve is optimal, the selected portfolio exceeds the monthly target mean, and the reported $\mathrm{VaR}_{0.95}(x^*)$ and return-tail $\mathrm{ES}_{0.95}^{\mathrm{return}}(x^*)$ are the return-convention summaries at the final aligned macro state. These quantities are not pricing-kernel diagnostics; they document the portfolio and conditional-quantile fit from which the residual ranks are constructed.

Table 4 reports the fitted conditional-quantile coefficients. The coefficients are regularized first-stage objects for the portfolio return $R_t^{\top}\hat{w}$, not structural prices of macro risk. They show which standardized lagged macro variables move the fitted portfolio-return quantiles before the residual ranking is converted into an ES-envelope.

Table 5 reports the active positions of the shared quantile portfolio. The fitted allocation is concentrated but not degenerate: the long-only box constraint binds for several large-cap value

Table 3: Baseline Quantile-Program Fit

Notes: VaR and return-tail ES are reported in the return convention at the baseline reporting state x^* and $\tau = 0.95$. This return-tail ES entry audits the fitted portfolio-return tail and is not the ES-envelope probability vector used for SDF construction. The target and sample means are monthly excess-return quantities. Active positions are weights above one basis point.

Quantity	Value
Solver status	optimal
Quantile grid	0.50, 0.75, 0.95
$\mathrm{VaR}_{0.95}(x^*)$	5.08%
$\mathrm{ES}_{0.95}^{\mathrm{return}}(x^*)$	6.25%
Target mean μ_0	0.30%
Sample mean $\bar{r}^\top \hat{w}$	0.80%
Maximum weight	20.00%
Active positions	7
Non-crossing constraints	enforced in sample
Mean-link constraint	trapezoidal at x^*

Table 4: Conditional Quantile Coefficients

Notes: Entries are coefficients from the shared-w conditional quantile fit. State variables are standardized and lagged one month. The intercept is the first element of $x^\top \beta(\tau)$. Coefficients are descriptive first-stage quantile coefficients; pricing-kernel interpretation enters only after the ES-envelope normalization and projection steps.

State variable	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.95$
Intercept	0.0127	0.0353	0.0601
Inflation	-0.0071	-0.0039	-0.0017
Industrial production growth	0.0038	0.0016	0.0018
Unemployment rate	0.0058	0.0031	0.0015
Term spread	-0.0037	-0.0051	-0.0040
VIX	0.0044	0.0120	0.0212

and middle book-to-market portfolios, while a small number of additional positions complete the target-return and tail-fit restrictions. Reporting these weights makes the residual-ranking step auditable, because the ES-envelope is built from the fitted portfolio return $R_t^\top \hat{w}$ rather than from an abstract market proxy.

Table 5: Baseline Shared Quantile-Portfolio Weights

Notes: The table reports portfolio weights above one basis point in the baseline shared-w quantile program. Weights are sorted in descending order and expressed in percent.

Test asset	Weight (%)
Large firms, low book-to-market	20.00
Large firms, book-to-market quintile 2	20.00
Large firms, book-to-market quintile 3	20.00
Large firms, book-to-market quintile 4	20.00
Small firms, book-to-market quintile 4	13.41
Small firms, high book-to-market	4.07
Size quintile 4, book-to-market quintile 2	2.52
Active positions	7
Total active weight	100.00

The two empirical SDFs are produced as described in Section 2.5, (i) the Hansen–Jagannathan (HJ) minimum-second-moment SDF m_t^{HJ} and (ii) the projected ES-envelope SDF m_t^{ES}.

For any SDF m, we calculate the normalization $\bar{m} = \frac{1}{T}\sum_t m_{t+1}$, the pricing residuals $\overline{mR} = \frac{1}{T}R^\top m$, the mean absolute pricing error

$$\mathrm{MAE}(m) = \frac{1}{k}\sum_{j=1}^{k}\left|\frac{1}{T}\sum_{t=1}^{T} m_{t+1}R_{t,j}\right|,$$

the ℓ_2 norm $\|\overline{mR}\|_2$, and the variance proxy $\mathrm{Var}(m)$.

4. Empirical Results

4.1. Pricing Geometry and Macro-State Evidence

In our application, both m^{HJ} and the projected ES-envelope SDF m^{ES} achieve machine-precision pricing errors (MAE and ℓ_2), whereas m^{ES} retains a substantially larger variance than m^{HJ}, consistent with its tail emphasis. Figure 1 puts this contrast in geometric and distributional perspective. Both reported kernels satisfy the admissibility audit, but the ES-envelope projection sits farther

Table 6: Pricing Error Diagnostics and SDF Variance

Notes: MAE is the mean absolute pricing error $\frac{1}{k}\sum_{j=1}^{k}\left|T^{-1}\sum_{t=1}^{T} m_{t+1}R_{t,j}\right|$. The ℓ_2 norm is $\left\|T^{-1}\mathbf{R}^{\top}m\right\|_2$. Variance proxy is $\mathrm{Var}(m)$. With the gross risk-free convention, the normalization is $T^{-1}\sum_t m_{t+1}R_{f,t} = 1$. Errors are at (near) machine precision because the reported kernels impose or project onto the in-sample pricing restrictions. The last two rows audit the nonnegativity condition for admissible kernels.

Metric	HJ	ES-envelope
Mean absolute pricing error	9.89×10^{-19}	2.28×10^{-18}
ℓ_2 norm of pricing errors	6.02×10^{-18}	1.37×10^{-17}
Risk-free pricing error	1.11×10^{-15}	3.33×10^{-16}
Variance proxy $\mathrm{Var}(m)$	0.2402	5.0946
Minimum SDF value	0.0000	0.0000
Negative observations	0	0

up the HJ variance dimension and places more mass in high-SDF states; its larger variance is an economically meaningful tail-emphasis statistic rather than a pricing error.

Figure 2 first summarizes how the reported SDFs vary with each standardized macro state $X_t^{(j)}$ by plotting nonparametric smooths of standardized kernels,

$$m_{t,z}^{\mathrm{HJ}} \;=\; \frac{m_t^{\mathrm{HJ}} - \bar{m}^{\mathrm{HJ}}}{s(m^{\mathrm{HJ}})}, \qquad m_{t,z}^{\mathrm{ES}} \;=\; \frac{m_t^{\mathrm{ES}} - \bar{m}^{\mathrm{ES}}}{s(m^{\mathrm{ES}})},$$

against unemployment, VIX, inflation, industrial-production growth, and the term spread. The tabulated tercile and rank-spread diagnostics below provide the complementary finite-sample summaries.

We discretize each macro variable into terciles (Low/Mid/High), and Table 7 tabulates the mean and dispersion of m_t^{HJ} and m_t^{ES} by tercile. High-VIX terciles exhibit much higher average SDF levels for the ES-envelope kernel and only mild differences for the HJ kernel. Unemployment is less stable: ES-envelope terciles show only a mild high-unemployment elevation, while the quintile-spread diagnostics below are not positive. The inflation terciles are non-monotone in this specification. Macro variables are standardized in-sample and lagged one month prior to returns.

As a robustness check on the state-dependence interpretation, Table 8 reports two summaries that do not rely on the plotted smooths: Spearman rank correlations between each macro state and each SDF, and the high-minus-low quintile spread in SDF levels. The ES-envelope–VIX relation is the most stable positive pattern. The HJ kernel is higher in low term-spread states and, in this sample, lower in high-unemployment states. The ES-envelope unemployment relation is mixed: the rank correlation is slightly positive, but the high-minus-low quintile spread is negative and imprecise.

Table 7: SDF Levels by Macro Terciles

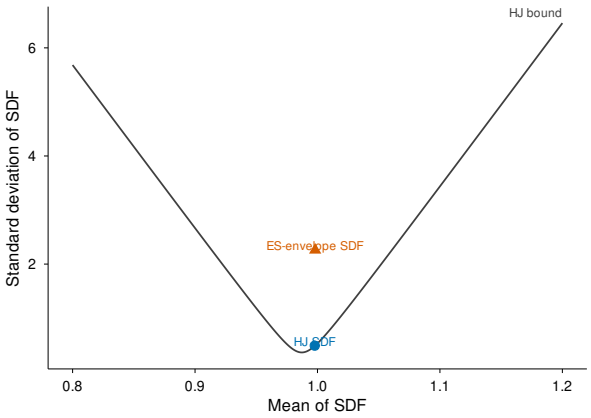
The table reports the mean and standard deviation of the reported SDFs within terciles of each standardized macro variable. Terciles are formed from in-sample empirical quantiles (equal-mass bins; lowest bin includes endpoints); macro variables are standardized and lagged one month. The HJ SDF is the nonnegative projected HJ benchmark used in the reported figures; the ES-envelope SDF is the projected tail-envelope constructed from residual ranks. Means reflect the level of the pricing kernel by state, while SD summarizes within-state dispersion. Larger SDs for the ES-envelope kernel are expected given its tail emphasis.

State variable	Tercile	HJ mean	HJ SD	ES-envelope mean	ES-envelope SD
Inflation	Low	0.97	0.50	1.15	2.23
Inflation	Mid	1.01	0.49	0.69	1.75
Inflation	High	1.01	0.48	1.16	2.69
Industrial production	Low	1.01	0.56	1.07	1.91
Industrial production	Mid	1.02	0.46	1.25	2.93
Industrial production	High	0.96	0.45	0.68	1.72
Unemployment	Low	1.06	0.53	1.08	2.35
Unemployment	Mid	0.98	0.45	0.78	1.84
Unemployment	High	0.95	0.49	1.15	2.55
Term spread	Low	1.08	0.49	0.99	2.18
Term spread	Mid	0.95	0.50	0.82	1.91
Term spread	High	0.96	0.48	1.18	2.63
VIX	Low	1.05	0.41	0.31	0.73
VIX	Mid	0.96	0.53	0.78	1.90
VIX	High	0.99	0.53	1.92	3.15

Figure 1: Pricing Geometry and Distribution of Reported Kernels

Panel (a) plots the unconditional HJ variance bound and locates the reported admissible kernels. The HJ SDF is the minimum-second-moment pricing benchmark; the projected ES-envelope SDF remains pricing-feasible but lies higher in the standard-deviation dimension because it concentrates more mass in tail states. Panel (b) shows kernel densities of the same reported nonnegative HJ and projected ES-envelope SDFs. The raw ES-envelope and unconstrained HJ diagnostics are excluded because they are not the reported admissible pricing kernels.

(a) Unconditional HJ bound



(b) Reported kernel densities

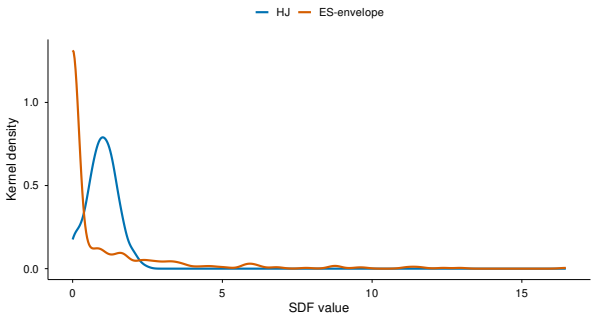


Table 9 adds moving-block bootstrap intervals for the quintile spreads, using 12-month blocks to preserve local serial dependence. The intervals reinforce the same calibration: the ES-envelope–VIX spread is clearly positive, the HJ term-spread and unemployment intervals are negative, and the HJ–VIX, ES-envelope unemployment, and inflation patterns are not precise enough to support stronger claims.

Table 8: Robustness of Macro-State Dependence

ρ_s denotes the Spearman rank correlation between the standardized macro variable and the reported SDF. Q5-Q1 is the difference between mean SDF values in the top and bottom quintiles of the same macro variable.

State variable	ρ_s HJ	Q5-Q1 HJ	ρ_s ES-envelope	Q5-Q1 ES-envelope
Inflation	0.03	0.06	-0.07	0.14
Industrial production	-0.03	-0.10	-0.06	-0.44
Unemployment	-0.11	-0.14	0.02	-0.18
Term spread	-0.11	-0.14	0.03	0.34
VIX	-0.06	-0.06	0.30	2.27

Because the baseline in-sample fit uses a compact quantile grid for computational reproducibility, we also check whether the main admissibility and state-dependence conclusions depend on that grid. Table 10 re-estimates the baseline test-asset span with compact, medium, and dense quantile grids

Figure 2: SDFs Across Macro States

Nonparametric smooths (loess) of standardized SDFs against standardized macro variables. HJ (blue) and the ES-envelope SDF (orange) are scaled to mean 0, variance 1 for comparability. The horizontal lower line in the ES-envelope panels reflects binding nonnegativity constraints in the projection: observations with projected SDF value zero all map to the same standardized value. These zero values are boundary solutions of the finite-sample projection problem, not structural claims that marginal utility is zero in those states. The clearest positive high-volatility pattern belongs to the ES-envelope kernel; HJ variation is milder, and its cleaner tabulated state spread is low term spread. The unemployment pattern is descriptive and mixed across the tabulated summaries.

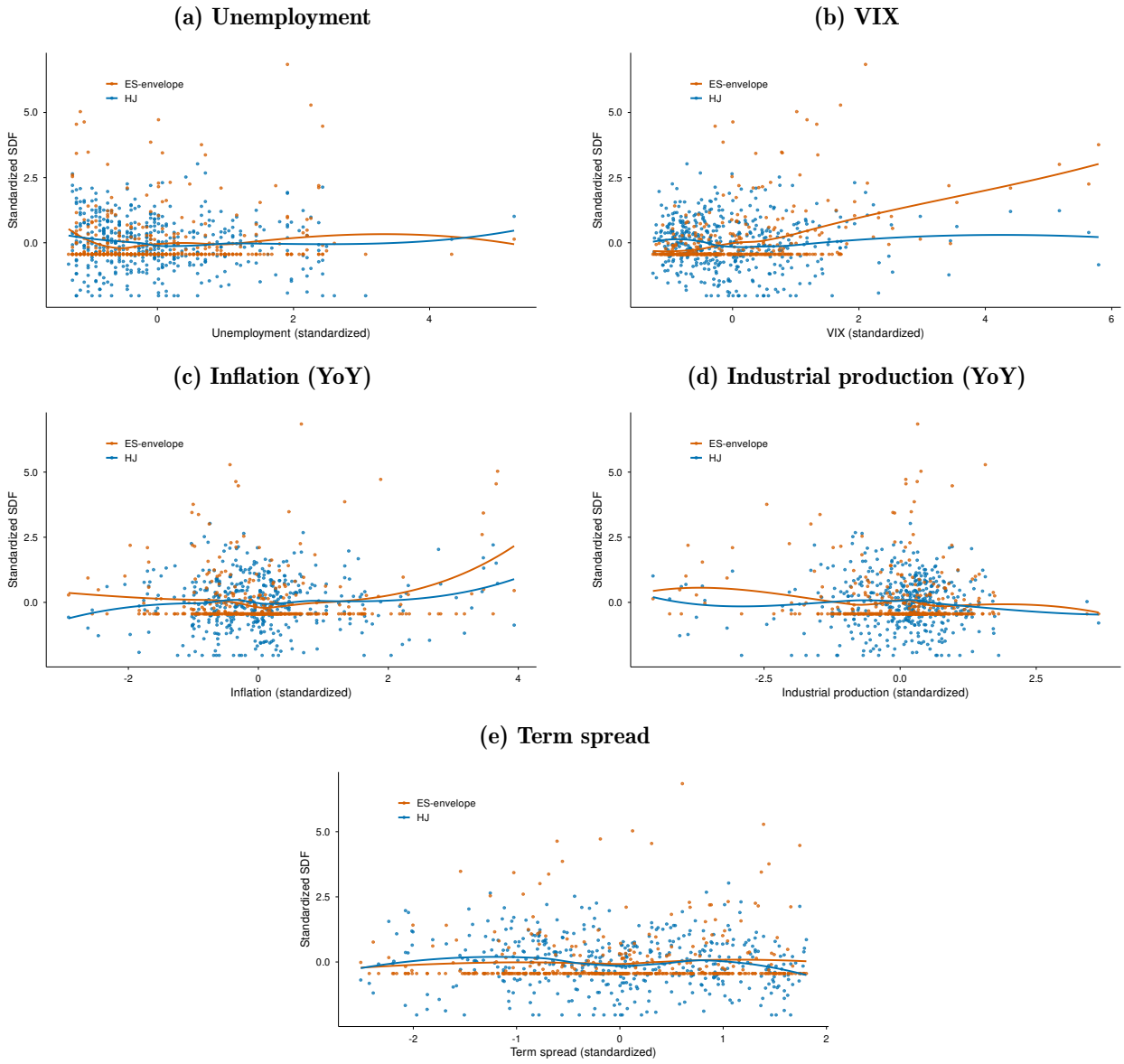


Table 9: Moving-Block Bootstrap Intervals for State Spreads

The table reports 90% percentile intervals for the Q5-Q1 spreads in Table 8, computed from 499 circular moving-block bootstrap replications. Blocks have length 12 months; the bootstrap design uses fixed draws, block length, and seed across the reported state variables. The intervals are stability diagnostics for the reported state spreads, not structural confidence intervals for a preference parameter.

State variable	HJ Q5-Q1 90% CI	ES-envelope Q5-Q1 90% CI
Inflation	[-0.03,0.17]	[-0.66,1.05]
Industrial production	[-0.22,0.05]	[-1.32,0.49]
Unemployment	[-0.25,-0.01]	[-1.03,1.02]
Term spread	[-0.29,-0.01]	[-0.42,1.10]
VIX	[-0.19,0.05]	[1.38,3.11]

while keeping the penalties and constraints fixed. All three grids recover nonnegative projected HJ and ES-envelope kernels with pricing errors at numerical precision. The ES-envelope–VIX rank relation remains positive across grids, while the unemployment spread becomes more negative under the denser grids. This reinforces the main interpretation: the positive volatility relation is not an artifact of the compact grid, whereas the labor-market pattern should not be read as a positive tail-state regularity.

Table 10: Robustness Across Quantile Grids

Notes: The table re-estimates the baseline 25-portfolio specification with alternative quantile grids and fixed baseline penalties. MAE is the mean absolute excess-return pricing error of the reported nonnegative projected SDF. The dense grid is $\{0.05, 0.10, \ldots, 0.95\}$.

Quantile grid	Levels	HJ MAE	ES-envelope MAE	Max weight	$\rho_s^{ES}(VIX)$	ES unemployment Q5-Q1	Status
Compact grid	3	0	0	0.20	0.30	-0.18	Successful
Medium grid	4	0	0	0.20	0.27	-0.28	Successful
Dense grid	19	0	0	0.20	0.27	-0.28	Successful

We also re-estimate the workflow across alternative return spans and penalty strengths. Table 11 reports three test-asset sets–the baseline 25 size-value portfolios, the same portfolios without dividends, and the Fama–French five factors augmented with momentum–under lower, baseline, and higher elastic-net penalties. All nine specifications recover nonnegative projected HJ and ES-envelope kernels with pricing errors at machine precision. The positive VIX rank relation for the ES-envelope kernel is preserved across the two 25-portfolio spans and remains positive, though smaller, in the factor span. In contrast, the ES-envelope unemployment spread is negative across this robustness grid, so the labor-market interpretation is treated as mixed rather than as a positive tail-state regularity.

The HJ representation reports payoff-price directions γ on the asset span and macro loadings c on X_t for the selected admissible kernel. Figure 3 plots the estimated HJ loadings alongside the

Table 11: Robustness Across Test Assets and Penalties

Notes: The table re-estimates the in-sample workflow for alternative traded test-asset spans and three penalty strengths. MAE is the mean absolute excess-return pricing error of the reported nonnegative projected SDF. The status column records whether the fit and the nonnegative pricing projection succeeded.

Test assets	Penalty	Assets	HJ MAE	ES-envelope MAE	$\rho_s^{HJ}(VIX)$	$\rho_s^{ES}(VIX)$	Status
25 size-value portfolios	Lower penalty	25	9.9×10^{-19}	2.4×10^{-18}	-0.06	0.27	Successful
25 size-value portfolios	Baseline penalty	25	9.9×10^{-19}	2.3×10^{-18}	-0.06	0.30	Successful
25 size-value portfolios	Higher penalty	25	9.9×10^{-19}	2.3×10^{-18}	-0.06	0.30	Successful
25 size-value portfolios without dividends	Lower penalty	25	1.1×10^{-18}	2.0×10^{-18}	-0.06	0.27	Successful
25 size-value portfolios without dividends	Baseline penalty	25	1.1×10^{-18}	2.0×10^{-18}	-0.06	0.27	Successful
25 size-value portfolios without dividends	Higher penalty	25	1.1×10^{-18}	2.0×10^{-18}	-0.06	0.27	Successful
Five factors plus momentum	Lower penalty	6	9.8×10^{-19}	2.1×10^{-18}	0.04	0.21	Successful
Five factors plus momentum	Baseline penalty	6	9.8×10^{-19}	2.1×10^{-18}	0.04	0.21	Successful
Five factors plus momentum	Higher penalty	6	9.8×10^{-19}	1.0×10^{-18}	0.04	0.10	Successful

implied ES-envelope loadings in one comparative display. For the ES-envelope SDF, we recover an implied linear representation by regressing

$$m_t^{\mathrm{ES}} \;=\; a + X_t^\top c^{\mathrm{ES}} - R_t^\top \gamma^{\mathrm{ES}} + \varepsilon_t,$$

and report the implied ES-envelope macro and payoff coefficients (with the sign convention aligning $-\hat{\gamma}$ to prices of risk). In practice, the implied ES-envelope representation loads most heavily on VIX, while the slack-related evidence is weaker and not monotone across rank and spread diagnostics.

To summarize slow-moving variation, we smooth both SDFs with an exponential moving average (EWMA) (half-life 60 months) and a 10-year rolling mean,

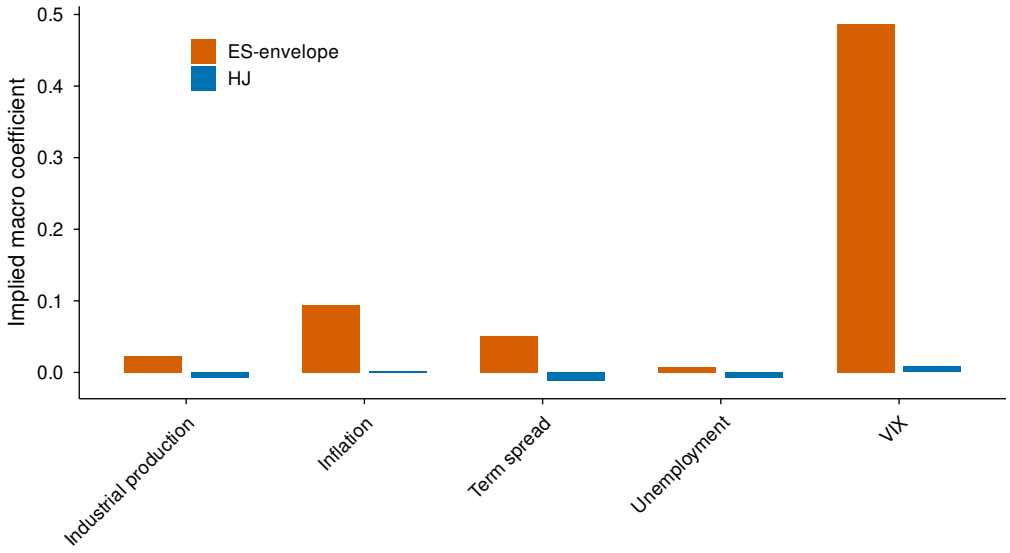
$$\overline{m}_t^{(120)} = \frac{1}{120} \sum_{s=t-119}^{t} m_s,$$

and overlay 12-month smooths. Figure 4 shows that both kernels vary with macro-financial stress; the ES-envelope SDF exhibits sharper amplitudes around stress episodes, partly because moving averages transmit large tail observations forward, whereas the HJ SDF moves more smoothly, reflecting average market pricing rather than downside concentration.

The macro–state plots in Figure 2 show state dependence across several macro dimensions. The clearest positive state pattern is the ES-envelope–VIX relation. The admissible HJ kernel is higher in low term-spread states and lower in high-unemployment states in the quintile-spread diagnostics, while its VIX spread is modest. The inflation and unemployment patterns for the ES-envelope kernel are weaker and not monotone, so we interpret them cautiously. The ES-envelope SDF exhibits larger estimated slopes–particularly along the VIX dimension–reflecting its emphasis on extreme downside states, whereas the HJ SDF is smoother and less volatile. Tercile summaries

Figure 3: State-Price Coefficients on Macroeconomic Variables

Bars report macro loadings on standardized variables (intercept omitted) for the state-dependent HJ SDF and the implied linear representation of the ES-envelope SDF. Positive (negative) values indicate that the selected empirical kernel is higher (lower) when the corresponding macro variable rises in the fitted linear representation. For the ES-envelope panel these are descriptive implied loadings, not structural preference parameters.



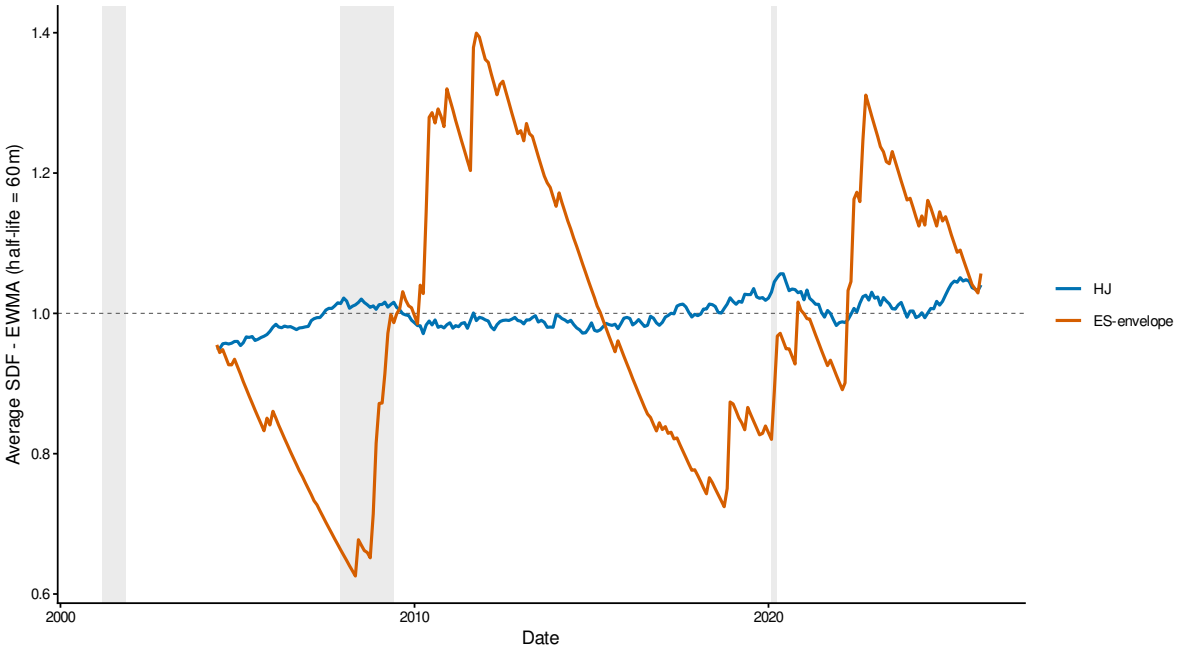
(Table 7) and rank/spread diagnostics (Table 8) mirror these patterns. Quantile-grid robustness (Table 10) and the asset/penalty robustness grid (Table 11) show that admissibility and the positive ES-envelope–VIX relation are not artifacts of the compact baseline grid, one return span, or one penalty value, while the unemployment relation is not robust as a positive tail-state claim. The comparative coefficient display (Figure 3) is consistent with the state and payoff channels: the HJ loadings are moderate, while the ES-envelope-implied loadings tilt more strongly toward states associated with drawdowns and volatility spikes. Finally, the 10-year rolling means (Figure 4) show that both kernels are time-varying, but the ES-envelope SDF displays sharper cyclical amplitudes–again consistent with a pricing kernel that concentrates mass in tail states.

Taken together, the macro analysis suggests two complementary empirical kernels. The HJ SDF behaves like a diversified mean–variance benchmark whose selected kernel level moves smoothly with term-spread and real-activity states. The ES-envelope SDF behaves like a tail-sensitive kernel that places more weight on downside states; its robust positive state pattern is volatility, not labor-market slack. Their joint use separates average pricing restrictions from tail-state emphasis.

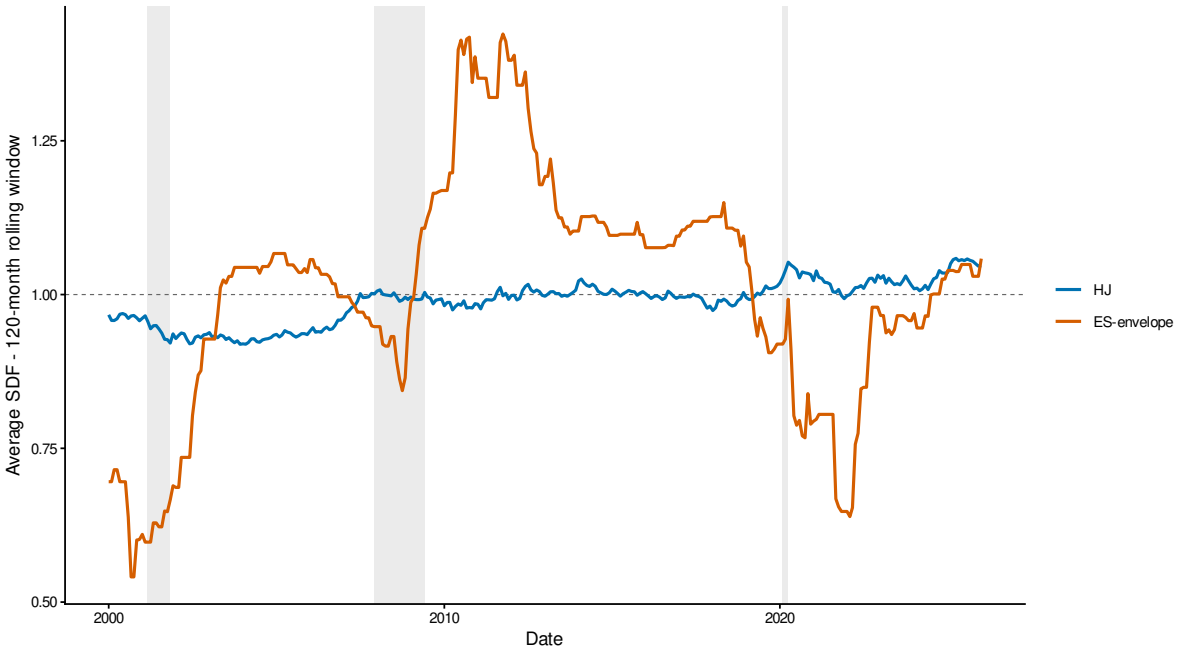
Figure 4: Time Varying Dynamics of the SDF

Low-frequency dynamics of the SDFs. Figure 4a shows an EWMA with a half-life of 60 months. Figure 4b shows a moving average over 120 lagged months. The EWMA offers less lag; the 120-month moving average is shown for robustness.

(a) EWMA (half-life = 60 months)



(b) 120-month rolling average



4.2. Rolling Holdout Diagnostics

The in-sample evidence above is a feasibility result: it shows that the reported kernels satisfy the pricing restrictions imposed by direct constrained estimation or projection on the estimation sample. The rolling holdout diagnostics in Tables 12 and 13 separate this feasibility result from pricing behavior away from the estimation window. The design uses all annual origins available from the sample with a 20-year training window, a five-year holdout window, and the dense quantile grid $\{0.05, 0.10, \ldots, 0.95\}$.

The same rolling run first audits the quantile portfolio itself. Table 12 shows that the fitted 0.95 portfolio-return quantile has a mean holdout hit rate of 0.8682, with origin-level rates ranging from 0.7667 to 0.9667. Ten of the 11 origin-level hit rates are below the 0.95 target. Pooled across the overlapping holdout windows, the fitted quantile records 573 hits in 660 origin-month evaluations. The corresponding exact binomial p-value, 4.23×10^{-16}, is a descriptive calibration audit rather than an independent-observation test because the five-year holdout windows overlap. We therefore do not treat the conditional quantile surface as a calibrated real-time forecast. Its role in the holdout exercise is narrower: it supplies the training-window portfolio and residual ordering used to construct the ES-envelope diagnostic. The average holdout portfolio return is 0.0102 with standard deviation 0.0417; the mean empirical lower-tail ES in the holdout windows is 0.0886.

The raw ES-envelope remains nonnegative and risk-free normalized by construction, but it does not price the risky test assets: across 11 annual holdout origins, raw ES-envelope MAE is 1.07×10^{-1}. The unconstrained HJ diagnostic has lower MAE, 1.08×10^{-2}, but produces 50 negative holdout kernel observations and is therefore interpreted only as diagnostic evidence, not as an admissible SDF estimate. For both HJ and ES-envelope objects, soft nonnegative projections reduce MAE to 7.80×10^{-5} for HJ and 1.64×10^{-4} for the ES-envelope while preserving nonnegativity and risk-free normalization. Exact nonnegative projections succeed in 9 of 11 origins; when feasible, pricing errors are at numerical precision. The two exact-projection failures occur in the Feb 2010–Jan 2015 and Feb 2012–Jan 2017 holdout windows. They fail together because exact nonnegative pricing is a property of the holdout return span: the common admissible set is empty in those windows, so no starting kernel can be projected onto it. The separating-portfolio certificates in the last panel confirm this geometry: normalized to unit gross leverage, the certificate portfolio has a strictly positive monthly excess return in all 60 months of each infeasible holdout block. The holdout evidence therefore supports the projection discipline rather than the raw tail measure as a stand-alone holdout SDF, and the infeasible windows are reported rather than silently dropped.

The rolling evidence sharpens the interpretation of the in-sample results. Exact nonnegative pricing is an economic feasibility restriction, not a guaranteed numerical post-processing step, and it fails in two of the eleven holdout windows. Because the HJ and ES-envelope exact projections face

Table 12: Rolling Holdout Portfolio-Quantile Diagnostics

Notes: The table summarizes the portfolio-return diagnostics from the same rolling origins used in Table 13. The hit rate is the share of holdout portfolio returns below the fitted 0.95 conditional quantile. The reported binomial p-value is a descriptive pooled coverage audit; it treats origin-month evaluations as if independent and is reported only to document the size of the calibration gap because the rolling holdout windows overlap. VaR and return-tail ES are training-window fitted summaries in the return convention; holdout lower-tail ES is the empirical mean loss in the lowest five percent of holdout portfolio returns.

Diagnostic	Value
Rolling origins	11
Target quantile hit rate	0.9500
Mean fitted-quantile hit rate	0.8682
Holdout hit-rate range	0.7667–0.9667
Origins below target	10/11
Aggregate hit months	573/660
Aggregate hit-rate 95 percent CI	0.8400–0.8931
Descriptive binomial p-value	4.23e-16
Mean training VaR	0.0562
Mean training return-tail ES	0.0641
Mean holdout lower-tail ES	0.0886
Mean holdout portfolio return	0.0102
Mean holdout portfolio SD	0.0417

Table 13: Rolling Holdout SDF Diagnostics

Notes: The table reports out-of-sample diagnostics across rolling annual origins. The training window is 240 months and the holdout window is 60 months. The quantile program uses the dense grid $\tau \in \{0.05, 0.10, \ldots, 0.95\}$. MAE, ℓ_2, risk-free error, negative observations, and variance are defined as in Table 6. The unconstrained HJ diagnostic is included only because it reveals the nonnegativity failure of an affine fit. Soft HJ and soft ES-envelope projections are nonnegative mean-R_f projections with a quadratic pricing-error penalty. The status and failure-window panels record when exact nonnegative pricing is infeasible. The certificate panel reports a unit-gross-leverage separating portfolio for each infeasible holdout window. The scaled margin is the minimum payoff after the risk-free normalization used in the convex-hull certificate; excess returns are monthly basis points.

Model	Origins	MAE	ℓ_2	Risk-free error	Negative observations	Var(m)
Unconstrained HJ diagnostic	11	1.08×10^{-2}	5.94×10^{-2}	1.15×10^{-1}	50	0.7081
Soft HJ projection	11	7.80×10^{-5}	4.94×10^{-4}	2.12×10^{-16}	0	0.9891
Exact HJ projection	9	2.37×10^{-18}	1.54×10^{-17}	1.48×10^{-16}	0	1.0906
Raw ES-envelope diagnostic	11	1.07×10^{-1}	5.48×10^{-1}	4.04×10^{-17}	0	18.9653
Soft ES-envelope projection	11	1.64×10^{-4}	1.08×10^{-3}	2.62×10^{-16}	0	2.1819
Exact ES-envelope projection	9	4.43×10^{-18}	2.78×10^{-17}	1.97×10^{-16}	0	2.1035

Model	Status	Rows
Unconstrained HJ diagnostic	Successful	11
Soft HJ projection	Successful	11
Exact HJ projection	Exact projection infeasible	2
Exact HJ projection	Successful	9
Raw ES-envelope diagnostic	Successful	11
Soft ES-envelope projection	Successful	11
Exact ES-envelope projection	Exact projection infeasible	2
Exact ES-envelope projection	Successful	9

Origin	Holdout window	Affected exact projections	Interpretation
240	Feb 2010–Jan 2015	HJ and ES-envelope	Common nonnegative pricing set infeasible
264	Feb 2012–Jan 2017	HJ and ES-envelope	Common nonnegative pricing set infeasible

Origin	Holdout window	Scaled margin (bp)	Gross lev.	Min. excess return (bp)	Mean excess return (bp)	Positive months
240	Feb 2010–Jan 2015	2.3720	1.0000	2.3720	20.2032	60/60
264	Feb 2012–Jan 2017	1.9652	1.0000	1.9652	13.8616	60/60

the same holdout admissible set, they fail in the same windows. The certificates show that these failures have an economic, not merely numerical, content: in each failed window the holdout return span contains a separating excess-return direction that is positive in every observed month. The soft projections show that admissibility can be preserved with small holdout pricing errors when exact pricing is infeasible, but they are penalized approximations rather than exact SDF recoveries. We therefore interpret the holdout exercise as support for the projection discipline and as a stress test of the selected kernels, not as a claim that the raw ES-envelope or unconstrained HJ fit is a stand-alone forecasting model.

5. Conclusion

This paper develops a unified, optimization-based workflow that links portfolio choice, conditional quantile estimation, and stochastic discount factor (SDF) recovery. By solving a joint quantile program with elastic-net regularization, we obtain tail-aware portfolios and a conditional quantile surface on a chosen grid. The fitted residual ranking then yields an ES-envelope tail measure implied by the dual structure of expected-shortfall; after normalization and, when feasible, projection onto the nonnegative pricing set, this tail measure becomes the reported ES-envelope SDF. A state-dependent Hansen–Jagannathan (HJ) minimum-second-moment SDF provides the mean–variance benchmark. The approach is transparent, computationally tractable, and can impose the no-arbitrage restrictions $E[m_{t+1}R^e_{t+1}] = 0$ with either $E[m_{t+1}] = 1$ or $E[m_{t+1}R_{f,t+1}] = 1$ whenever the relevant nonnegative pricing set is feasible; otherwise it reports soft projections and feasibility failures as diagnostics.

Applied to Fama–French test assets and standard U.S. macro instruments, both recovered SDFs satisfy pricing restrictions in-sample after projection or direct imposition, while revealing macro-state dependence. The HJ SDF behaves like a diversified mean–variance kernel–smoothly varying and relatively low variance, with its most reliable state spread in low term-spread states. The ES-envelope SDF concentrates mass on downside states and therefore exhibits the clearest positive volatility relation, while the labor-market relation is mixed and not robust as a positive tail-state claim. Taken together, these objects are two empirical selections inside the same finite-sample pricing discipline: one selects the minimum-second-moment admissible kernel, while the other selects the admissible kernel closest to the expected-shortfall tail envelope.

Methodologically, three features are central. First, the shared-w multi-quantile fit stabilizes tail estimation and connects VaR/ES to investable portfolios. Second, explicit ℓ_1/ℓ_2 penalties control complexity in both the portfolio and conditioning blocks, yielding interpretable loadings and mitigating multicollinearity. Third, the ES-envelope construction provides a direct summary of

downside states without requiring a separate dual solve, and can be projected back onto the HJ pricing space when the nonnegative pricing set is feasible.

These findings place the paper between four ongoing debates. For the conditioning-information literature, the message is that macro instruments are useful, but the empirical object remains a selection from a finite test-asset pricing set. For the HJ-distance, misspecified-SDF, and SDF-applicability literatures, the rolling projection failures show why admissibility has to be audited separately from average pricing fit. For the flexible-SDF and machine-learning literature, the results provide a transparent benchmark in which no-arbitrage restrictions, nonnegativity, macro interpretation, and failure modes are visible rather than absorbed into predictive performance. For the expected-shortfall and conditional-quantile state-price literatures, the distinction between the raw ES-envelope and its projected pricing-kernel version is essential: ES dual weights identify tail states, while admissible pricing additionally requires normalization, nonnegative state prices, and risky-asset pricing moments.

Three scope conditions govern the interpretation. First, the finite test-asset span identifies an admissible set, not a unique marginal utility process; the HJ and ES-envelope kernels are selections from that set. Second, the ES-envelope is a tail probability measure before projection, so its pricing interpretation depends on the normalization and projection step when risky-asset pricing moments are required. Third, the macro-state evidence is descriptive for the reported U.S. monthly sample rather than a structural preference estimate. These limits are design features of the empirical exercise: they make clear which object is estimated, which restrictions are imposed, and which claims remain outside the present evidence.

The framework is extensible. On the data side, it can incorporate richer test assets (characteristic-sorted portfolios, industry/sector ETFs, cross-asset macro factors) and higher-frequency macro instruments. On the modeling side, the quantile surface can be enriched with nonlinear bases or interaction terms; penalties can be tuned via the holdout criterion above or by stability selection; and the non-crossing and mean-link constraints can be adapted to different target functionals. Inference can proceed via block bootstrap for functionals of m_{t+1} and for the integrated mean constraint. The rolling diagnostics, quantile-grid robustness, and asset/penalty robustness reported here are stability checks, but broader expanding-window comparisons of tuning rules remain useful next steps. Finally, connecting the ES-envelope SDF to equilibrium preferences (e.g., disappointment aversion or recursive utility) is a promising avenue to map the empirical tail emphasis into structural parameters.

For practice, the message is twofold. First, conditioning on macro states is associated with changes in the shape and volatility of the pricing kernel; incorporating such information can inform risk measurement and portfolio design, particularly in stress regimes. Second, tail-aware kernels identify

observed states in which downside payoffs carry high empirical discount-factor weight, offering a disciplined way to study downside protection while preserving transparency about the projection and test-asset span that support the claim. Overall, the proposed regularized conditional quantile framework offers an interpretable bridge between portfolio optimization and state-dependent asset pricing, and it makes explicit where tail risk receives high empirical discount-factor weight in the pricing of returns.

Data and Replication Availability

The empirical analysis uses public Fama–French portfolio and factor data together with public FRED macroeconomic series. The replication package accompanying this manuscript refreshes these inputs, reproduces the reported tables and figures, rebuilds the article and presentation, and runs internal consistency checks. It also records the computational environment, package versions, data-source manifests, solver settings, and checksums for the paper-facing outputs. The accompanying documentation maps each table and figure to the corresponding reproduction workflow for the reported evidence.

Declaration of Generative AI and AI-Assisted Technologies

During the preparation of this manuscript, the authors used AI-assisted tools, including OpenAI's ChatGPT and Codex, to support language editing, literature organization, code review, and consistency checks. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Acerbi, C. and D. Tasche (2002): "On the Coherence of Expected Shortfall," *Journal of Banking and Finance*, 26, 1487–1503.

Adrian, T., N. Boyarchenko, and D. Giannone (2019): "Vulnerable Growth," *American Economic Review*, 109, 1263–1289.

Ait-Sahalia, Y. and A. W. Lo (2000): "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics*, 94, 9–51.

Ang, A., J. Chen, and Y. Xing (2006): "Downside Risk," *The Review of Financial Studies*, 19, 1191–1239.

Bassett, G. W., R. Koenker, and G. Kordas (2004): "Pessimistic Portfolio Allocation and Choquet Expected Utility," *Journal of Financial Econometrics*, 2, 477–492.

Bollerslev, T., V. Todorov, and L. Xu (2015): "Tail Risk Premia and Return Predictability," *Journal of Financial Economics*, 118, 113–134.

Bonaccolto, G., M. Caporin, and S. Paterlini (2018): "Asset allocation strategies based on penalized quantile regression," *Computational Management Science*, 15, 1–32.

Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009): "Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns," *The Review of Financial Studies*, 22, 3411–3447.

Brodie, J., I. Daubechies, C. De Mol, D. Giannone, and I. Loris (2009): "Sparse and Stable Markowitz Portfolios," *Proceedings of the National Academy of Sciences*, 106, 12267–12272.

Chen, L., M. Pelger, and J. Zhu (2024): "Deep Learning in Asset Pricing," *Management Science*, 70, 714–750.

Cochrane, J. H. (2005): *Asset Pricing*, Princeton, NJ: Princeton University Press, revised edition ed.

Farago, A. and R. Tedongap (2018): "Downside Risks and the Cross-Section of Asset Returns," *Journal of Financial Economics*, 129, 69–86.

Ferson, W. E. and A. F. Siegel (2001): "The Efficient Use of Conditioning Information in Portfolios," *The Journal of Finance*, 56, 967–982.

——— (2003): "Stochastic Discount Factor Bounds with Conditioning Information," *The Review of Financial Studies*, 16, 567–595.

Fissler, T. and J. F. Ziegel (2016): "Higher Order Elicitability and Osband's Principle," *The Annals of Statistics*, 44, 1680–1707.

Freyberger, J., A. Neuhierl, and M. Weber (2020): "Dissecting Characteristics Nonparametrically," *The Review of Financial Studies*, 33, 2326–2377.

Gospodinov, N., R. Kan, and C. Robotti (2014): "Misspecification-Robust Inference in Linear Asset Pricing Models with Irrelevant Risk Factors," *Review of Financial Studies*, 27, 2139–2170.

Grammig, J., C. Hanenberg, C. Schlag, and J. Soenksen (2025): "Diverging Roads: Theory-Based vs. Machine-Learning-Implied Stock Risk Premia," *Journal of Financial Econometrics*, 23, 1–55.

Grammig, J. and E.-M. Kuechlin (2018): "A Two-Step Indirect Inference Approach to Estimate the Long-Run Risk Asset Pricing Model," *Journal of Econometrics*, 205, 6–33.

Gu, S., B. Kelly, and D. Xiu (2020): "Empirical Asset Pricing via Machine Learning," *The Review of Financial Studies*, 33, 2223–2273.

——— (2021): "Autoencoder Asset Pricing Models," *Journal of Econometrics*, 222, 429–450.

Hansen, L. P. and R. Jagannathan (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

——— (1997): "Assessing Specification Errors in Stochastic Discount Factor Models," *The Journal of Finance*, 52, 557–590.

Hansen, L. P. and S. F. Richard (1987): "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica*, 55, 587–613.

Kelly, B. T., S. Malamud, and K. Zhou (2024): "The Virtue of Complexity in Return Prediction," *The Journal of Finance*, 79, 459–503.

Kelly, B. T., S. Pruitt, and Y. Su (2019): "Characteristics Are Covariances: A Unified Model of Risk and Return," *Journal of Financial Economics*, 134, 501–524.

Kim, S. and R. A. Korajczyk (2024): "Large Sample Estimators of the Stochastic Discount Factor," *Journal of Financial Econometrics*, 22, 1672–1713.

Koenker, R. (2005): *Quantile Regression*, Cambridge University Press.

Korsaye, S. A., A. Quaini, and F. Trojani (2025): "Smart Stochastic Discount Factors," *Management Science*, articles in Advance.

Kozak, S., S. Nagel, and S. Santosh (2020): "Shrinking the Cross-Section," *Journal of Financial Economics*, 135, 271–292.

Lettau, M. and S. Ludvigson (2001): "Consumption, Aggregate Wealth, and Expected Stock Returns," *The Journal of Finance*, 56, 815–849.

Lettau, M. and M. Pelger (2020): "Factors That Fit the Time Series and Cross-Section of Stock Returns," *The Review of Financial Studies*, 33, 2274–2325.

Merlo, L., L. Petrella, and V. Raponi (2021): "Forecasting VaR and ES using a joint quantile regression and its implications in portfolio allocation," *Journal of Banking & Finance*, 133, 106248.

Metaxoglou, K. and A. Smith (2017a): "Forecasting Stock Returns Using Option-Implied State Prices," *Journal of Financial Econometrics*, 15, 427–473.

——— (2017b): "State Prices of Conditional Quantiles: New Evidence on Time Variation in the Pricing Kernel," *Journal of Applied Econometrics*, 32, 192–217.

Patton, A. J., J. F. Ziegel, and R. Chen (2019): "Dynamic Semiparametric Models for Expected Shortfall (and Value-at-Risk)," *Journal of Econometrics*, 211, 388–413.

Pezzo, L., Y. Zhu, M. K. Hassan, and J. Tian (2024): "Testing the Boundaries of Applicability of Standard Stochastic Discount Factor Models," *Journal of Financial Stability*, 72, 101268.

Rockafellar, R. T. and S. Uryasev (2000): "Optimization of Conditional Value-at-Risk," *Journal of Risk*, 2, 21–42.

——— (2002): "Conditional Value-at-Risk for General Loss Distributions," *Journal of Banking and Finance*, 26, 1443–1471.

Ross, S. A. (2015): "The Recovery Theorem," *The Journal of Finance*, 70, 615–648.

Rytchkov, O. and X. Zhong (2025): "Macroeconomic Content of Characteristics-Based Asset Pricing Models: A Machine Learning Analysis," SSRN working paper, last revised January 30, 2025.

Soenksen, J. and J. Grammig (2021): "Empirical Asset Pricing with Multi-Period Disaster Risk: A Simulation-Based Approach," *Journal of Econometrics*, 222, 805–832.

Taylor, J. W. (2019): "Forecasting Value at Risk and Expected Shortfall Using a Semiparametric Approach Based on the Asymmetric Laplace Distribution," *Journal of Business & Economic Statistics*, 37, 121–133.

Zou, H. and T. Hastie (2005): "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 301–320.