Agnostic Stochastic Discount Factors in Small Samples:

Theory, Momentum, and Cross-Market Evidence*

Kuntara Pukthuanthong† , Richard Roll‡ , Junbo Wang§

March 2026

Abstract

We study stochastic discount factor estimation when the cross section is large and the relevant time window is short. We develop a return-only estimator that recovers the pricing kernel directly from traded returns, without macroeconomic proxies, preference restrictions, or an imposed factor structure. The estimator admits a finite-state interpretation and a conditional lag-state representation that links discount-factor estimation to return dynamics. In one application, the lag-state SDF produces a momentum signal with time-varying lag weights and positive average payoffs. In another, the agnostic SDF outperforms Kim–Korajczyk in both domestic and international pricing, while also clarifying how finite-sample noise shapes pricing-error comparisons.

Keywords: stochastic discount factor, asset pricing, momentum, finite-sample bias, market integration

JEL classification: G12, G14, C58

1 Introduction

Empirical asset pricing has made substantial progress in describing the cross section of expected returns, but practical estimation of the stochastic discount factor remains challenging. The difficulty is primarily empirical rather than conceptual. Many existing approaches identify the SDF only after imposing considerable structure, whether through macroeconomic proxies, preference specifications, or an explicit factor model for returns and discount factors (Hansen and Singleton, 1983; Campbell, 1993; Lettau and Ludvigson, 2001; Kelly et al., 2019; Lettau and Pelger, 2020; Gu et al., 2021). Such restrictions can be valuable, but they also make empirical performance depend on the accuracy of the maintained specification (Kozak et al., 2020; Patton and Weller, 2022). This concern is

*We thank [Name 1], [Name 2], [Name 3], [Name 4], and [Name 5] for very helpful comments and suggestions. We also thank seminar participants at [Institution 1], [Institution 2], and [Institution 3], and conference participants at [Conference 1] and [Conference 2]. All remaining errors are our own.

†Department of Finance, University of Missouri. Email: pukthuanthongk@missouri.edu.

‡California Institute of Technology. Email: rroll@caltech.edu.

§Department of Finance, Louisiana State University. Email: junbowang@lsu.edu.

especially important in settings with a rich cross section of traded assets and a relatively short time series over which the relevant pricing object must be estimated. That environment arises naturally in rolling asset-pricing exercises and conditional strategy design.

This paper studies SDF estimation in that large-cross-section, short-window environment. Its starting point is a simple idea: if the purpose of the SDF is to price traded payoffs, then one should ask how much of that discounting information can be recovered directly from traded returns, without first passing through a maintained macroeconomic model or a prespecified factor structure (Hansen and Jagannathan, 1991; A"it-Sahalia and Lo, 1998; Rosenberg and Engle, 2002). That perspective is closely related to Cochrane's payoff-based view of intertemporal asset pricing (Cochrane, 2014). In his analysis of long-horizon portfolio problems, Cochrane emphasizes that the basic pricing logic can be applied to streams of payoffs by treating date and state symmetrically and by focusing on final payoffs rather than on the detailed composition of dynamic trading strategies (Cochrane, 2014). That observation is especially useful here. The objective of the present paper is not to solve a fully specified dynamic portfolio problem. It is to recover, from observed returns, a practically useful discounting rule that prices a rich cross section of payoffs over a short horizon.

The paper develops a return-only, agnostic estimator for that purpose. The estimator is agnostic in a precise sense. It does not require macroeconomic proxies or preference assumptions, and it does not impose a factor structure on either returns or the SDF. Instead, it treats the pricing restriction itself as the identifying object. Intuitively, at any given date the econometrician observes a panel of returns across many assets. The estimator searches for the single discounting rule that prices the cross section, using only information contained in the observed return panel. This approach is attractive because it does not rely on distributional restrictions or on the correctness of a particular factor representation.

The first contribution of the paper is to recast the agnostic estimator in a more transparent and economically useful way. The key step is to view the estimator through a finite-state pricing problem. In a finite-state environment, the cross section of payoffs identifies a vector of state prices, and the SDF follows once those state prices are scaled by probabilities. This representation makes clear what a return-only SDF is doing: it extracts, from the span of traded payoffs, the discounting weights required to price a unit payoff across states (Cochrane, 2014; Ross, 2015). That same logic becomes operational once recent return cross sections are interpreted as the relevant set of states. At a formation date, the econometrician observes the last several return cross sections across many assets. Treating those lagged cross sections as states yields a conditional SDF that takes the form of a vector of lag weights. Those weights summarize which parts of recent return history are most relevant for pricing the current cross section. In that sense, past returns are not used merely as predictors. They are used as the conditioning information set that supports a pricing kernel estimated from the cross section itself (Hansen and Richard, 1987; Ferson and Siegel, 2001).

This lag-state interpretation leads naturally to the paper's first application. Standard momentum strategies aggregate past returns using a fixed lag structure chosen in advance (Jegadeesh and Titman, 1993; Moskowitz and Grinblatt, 1999; Asness et al., 2013). The approach in this

paper instead lets the lag profile be disciplined by asset-pricing restrictions. The resulting signal remains momentum-like in the sense that it aggregates past returns into a cross-sectional score, but it does so using weights that are estimated from the pricing problem at the formation date. The construction is implemented in two steps. First, the recent return window is used to estimate the lag-state SDF. Second, those lag weights are mapped into an asset-level score by forming a weighted average of each stock's demeaned lagged returns. The demeaning matters economically because it removes the common component of the lagged cross section and ties the strategy to a self-financing long-short trade rather than to the level of the market. This construction yields a momentum signal whose effective horizon is allowed to change over time as the pricing relevance of different lags changes (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016; Huang, 2022; Jegadeesh et al., 2025).

The empirical evidence indicates that this conditional SDF construction delivers an economically meaningful momentum signal. In the baseline implementation, the SDF-weighted long–short strategy generates a positive and statistically significant average payoff, and its performance weakens when stale lag weights are used, which is consistent with time variation in the pricing relevance of different lags. The shape of the estimated lag profile also contains information about downside outcomes: a simple state variable based on the recency-weight ratio helps identify periods in which crash risk is lower, and a managed version of the strategy improves left-tail behavior at the cost of some average return. Finally, the characteristic-based evidence shows that the signal is strongest in economically intuitive regions of the cross section and weakens, sometimes sharply, in the smallest, lowest-price, and most illiquid stocks, where microstructure frictions, transient price pressure, and reversal-like dynamics are more likely to matter. Taken together, these results support the interpretation of the lag-state SDF as a feasible, theory-based proxy for the continuation direction in lag space rather than as a fixed-window momentum rule.

The second contribution is empirical and comparative. The paper evaluates the agnostic estimator against the Kim–Korajczyk approach, which imposes additional factor structure on both returns and the SDF (Kim and Korajczyk, 2024). That comparison is important because the gain from structure is potentially real when the imposed model is correct, but the cost of structure can also be substantial when the factor representation is incomplete or misspecified (Kelly et al., 2019; Kozak et al., 2020; Lettau and Pelger, 2020; Patton and Weller, 2022). The empirical design therefore asks a sharper question than the usual in-sample horse race. Using individual U.S. stocks to construct the agnostic SDF over short rolling windows, the paper studies how the resulting pricing kernel performs on the broad test-asset universes of Chen and Zimmermann (2022) and Jensen et al. (2023) and how that performance compares with KK in both domestic and international settings.

The results point to a coherent set of findings. In domestic data, the agnostic SDF estimated from individual CRSP stocks outperforms KK on the Chen–Zimmermann portfolio universe (Chen and Zimmermann, 2022). In international data, the agnostic estimator again dominates KK, suggesting that the return-only SDF retains meaningful pricing information even when transported

across markets (Jensen et al., 2023). The combined evidence indicates that useful discount-factor variation can be extracted directly from returns without imposing a restrictive factor structure. At the same time, the domestic–international contrast highlights an important practical issue: the evaluation of SDFs is highly sensitive to finite-sample noise, especially in cross-market pricing exercises where pricing errors are measured from time-varying discount factors on heterogeneous test assets (Chen and Knez, 1995; Bekaert and Harvey, 1995; Pukthuanthong and Roll, 2009; Freyberger et al., 2025). This tension between transferability and finite-sample noise is central to the paper's interpretation of the empirical results.

The paper's contribution can be summarized in four parts. First, it provides a return-only and economically transparent interpretation of the agnostic SDF by recasting it as a finite-state pricing object and extending that logic to a conditional lag-state setting. Second, it delivers a momentum application in which lag weights are not imposed ex ante but are learned from the pricing kernel itself, thereby linking a familiar return-predictability strategy to an explicitly asset-pricing-based object. Third, it shows in domestic pricing exercises that a short-window agnostic SDF can outperform both a constant benchmark and a structured alternative based on factor restrictions. Fourth, it shows in international pricing exercises that the same agnostic estimator continues to dominate structured nontrivial alternatives, while the apparent success of the constant benchmark can be traced to the mechanics of estimation noise in short samples.

The remainder of the paper proceeds as follows. Section 2 situates the paper in the literature on SDF estimation, conditional return predictability, and cross-market pricing. Section 3 develops the return-only agnostic SDF, beginning with the finite-state interpretation and then extending it to the conditional lag-state construction used in the momentum application. Section 4 outlines the simulation framework for the large-N, small-T regime. Section 5 studies momentum through the lens of the conditional SDF. Section 6 compares the agnostic estimator with KK in pricing domestic and international returns and discusses the implications for market integration. Section 7 concludes.

2 Related Literature

This paper is related to five strands of the literature. The first develops the stochastic discount factor as the central object in asset pricing and studies how conditional information enters pricing restrictions. The second studies empirical pricing kernels, state prices, and return-based restrictions without fully specifying a macroeconomic model. The third estimates SDFs and risk factors using structured, latent-factor, or high-dimensional methods. The fourth studies momentum and related forms of conditional return continuation. The fifth examines cross-market pricing, international integration, and segmentation. Our paper lies at the intersection of these strands: it develops a return-only SDF estimator for a large-cross-section, short-window environment, interprets that object through a finite-state and lag-state representation, and uses it in two applications, a momentum strategy and a domestic-versus-international pricing comparison.

2.1 Foundations of the SDF and conditional pricing

Modern asset pricing is built around the stochastic discount factor representation of security prices. The foundational contributions of Lucas (1978), Breeden (1979), Hansen and Singleton (1982), Hansen and Singleton (1983), and Hansen and Richard (1987) establish the link between intertemporal marginal rates of substitution, conditional information, and asset returns. Hansen and Jagannathan (1991) provide the canonical nonparametric restrictions on admissible discount factors, while Ferson and Harvey (1991), Campbell (1993), Jagannathan and Wang (1996), Ferson and Siegel (2001), and Lettau and Ludvigson (2001) show how conditioning information materially changes the empirical content of linear factor models and SDF restrictions. Related work by Alvarez and Jermann (2005), Cochrane (2011), Cochrane (2014), Ross (2015), and Martin (2017) further emphasizes that discount-factor variation is economically central even when the underlying object is inferred from traded assets rather than from direct consumption data.

The present paper is close to this literature in spirit but different in emphasis. We do not seek to recover a fully structural discount factor from macroeconomic quantities. Instead, we ask how much discounting information can be recovered directly from traded returns when the econometrician observes a rich cross section but only a short effective time-series window. This payoff-based orientation is especially close to Cochrane (2014), who emphasizes that intertemporal asset-pricing problems can be formulated directly in terms of payoff streams rather than one-period returns. That perspective motivates our finite-state representation and our focus on a return-only pricing kernel estimated from the span of traded payoffs.

2.2 Empirical pricing kernels, state prices, and return-only restrictions

A second branch of the literature studies pricing kernels and state prices directly from asset-market data. Nonparametric and semiparametric approaches include A"it-Sahalia and Lo (1998), A"it-Sahalia and Lo (2000), Rosenberg and Engle (2002), and Adrian et al. (2013). Related contributions such as Long (1990), Ross (1976), and Ross (2015) study alternative return-based ways of characterizing the pricing kernel or the relevant stochastic discount factor. At a broader level, this literature shows that much can be learned about discounting from asset returns even when the econometrician does not begin with a fully specified macroeconomic model.

Our paper belongs most naturally in this return-based tradition. The agnostic estimator does not start from preferences, consumption, or a maintained factor structure. Instead, it starts from the pricing restriction itself and asks which time-varying discounting rule prices the observed cross section. In this sense, the paper is more closely related to return-based pricing-kernel and state-price estimation than to macro-finance models in which the SDF is parameterized from the outset. A recent paper that reinforces this distinction is Delikouras and Dittmar (2022), who show that return-spanned pricing kernels and structural Euler-equation restrictions need not line up cleanly once one imposes investment-side optimality conditions. That is precisely why the present paper works directly with the pricing restriction and traded payoffs.

2.3 Structured, latent-factor, and high-dimensional SDF estimation

A third literature estimates discount factors and pricing factors by imposing additional structure on the return-generating process. The classic starting point is the factor-model tradition, including Fama and French (1993), Hou et al. (2015), Fama and French (2015), Barillas and Shanken (2017), and Fama and French (2018). This literature is closely connected to the broader cross-sectional predictor literature summarized in Green et al. (2017), Bessembinder et al. (2019), and the data and replication efforts of Chen and Zimmermann (2022) and Jensen et al. (2023).

More recent work estimates SDFs directly in high-dimensional settings. Kelly et al. (2019) show that characteristics can be understood as covariances in a unified model of risk and return. Kozak et al. (2018) reinterpret factor models from an SDF perspective, and Kozak et al. (2020) estimate a high-dimensional SDF using shrinkage across a large set of candidate factor returns. Lettau and Pelger (2020) estimate latent asset-pricing factors designed to fit both the time series and the cross section, while Feng et al. (2020) study model selection in the factor zoo. Freyberger et al. (2020) show how nonparametric methods can uncover which characteristics matter for expected returns, and Gu et al. (2020) and Gu et al. (2021) develop machine-learning and autoencoder approaches to empirical asset pricing subject to no-arbitrage restrictions.

Post-2020 papers continue this structured turn. Patton and Weller (2022) show that heterogeneity in risk prices across segments is empirically widespread. Bybee et al. (2023) combine text, latent-factor structure, and variable selection to estimate interpretable systematic risk factors. Freyberger et al. (2025) address one of the main practical issues in high-dimensional panels—missing data—within a generalized method-of-moments framework. These papers are closely related to our paper, but they share an important feature that distinguishes them from the agnostic approach: each begins with a maintained structure, whether through observable characteristics, latent factors, shrinkage, segmentation, or nonlinear architecture. Our paper is complementary but distinct. The agnostic estimator does not posit a factor representation for either returns or the discount factor. It recovers a return-only SDF directly from the span of traded payoffs and studies how that object behaves when estimated over short rolling windows.

2.4 Momentum, reversal, and conditional continuation

The first empirical application of the paper is related to the large literature on momentum and conditional return continuation. The foundational evidence includes Lehmann (1990), Lo and MacKinlay (1990), Jegadeesh (1990), and especially Jegadeesh and Titman (1993), who document intermediate-horizon continuation in stock returns. Subsequent work studies the sources and robustness of momentum, including Chan et al. (1996), Conrad and Kaul (1998), Moskowitz and Grinblatt (1999), Hong et al. (2000), Grundy and Martin (2001), and Hvidkjaer (2006). The evidence that momentum extends well beyond U.S. equities is documented by Moskowitz et al. (2012) and Asness et al. (2013).

A second group of papers emphasizes that momentum is state dependent and that its downside risk is forecastable. Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016) show that

momentum crashes are predictable and that dynamic risk management substantially alters performance. More recent work sharpens this point. Hoberg et al. (2022) show that momentum profits vary with buy-side competition and information diffusion. Huang (2022) show that an ex ante state variable predicts future momentum profitability in the United States and internationally. Lou and Polk (2022) connect momentum profitability to arbitrage activity and crowdedness, while Arnott et al. (2023) show that factors themselves display strong cross-sectional momentum. Most recently, Jegadeesh et al. (2025) study the transition from short-term reversal to longer-horizon momentum around the world.

Our contribution to this literature is different from standard anomaly construction. We do not impose a fixed lag window and then ask whether that signal earns a return. Instead, we estimate a lag-state SDF from the pricing kernel itself and use it to determine which parts of recent return history matter most at the formation date. The resulting signal is still momentum-like, but its lag structure is endogenous and time varying. This is why the momentum application in Section 5 should be read as an application of the pricing kernel rather than as a stand-alone return-predictability exercise.

2.5 Cross-market pricing, international integration, and segmentation

The second empirical application of the paper is related to the literature on market integration and cross-market pricing. The theoretical foundations include Solnik (1974), Stulz (1981), Adler and Dumas (1983), and Errunza and Losq (1985), all of which study how asset-pricing restrictions change when markets are internationally integrated or partially segmented. Chen and Knez (1995) develop a pricing-based theory of market integration built on the law of one price and arbitrage restrictions across markets. Bekaert and Harvey (1995) and De Santis and Gérard (1997) show that international integration is empirically incomplete and time varying, while Bekaert et al. (2007) emphasize that local liquidity and segmentation continue to matter for expected returns. Pukthuanthong and Roll (2009) show that simple return correlations are not reliable measures of integration and propose an alternative measure based on common-factor explanatory power.

Our paper complements this literature from the perspective of estimated discount factors. We ask whether an SDF extracted from a broad U.S. stock universe transfers to domestic and international test assets, and whether the answer depends on the finite-sample properties of the pricing-error metric itself. This distinction is central to Section 6. The domestic-versus-international comparison is informative about segmentation, but only after accounting for the fact that cross-sectional pricing-error metrics can mechanically favor a constant benchmark when the candidate SDF is time varying and evaluation histories are short. That concern is closely related to the missing-data and uneven-history issues studied by Freyberger et al. (2025).

Taken together, these strands clarify the paper's position. Relative to the traditional SDF literature, the paper is more return-based and more implementation-oriented. Relative to the structured and high-dimensional factor literature, it is less model dependent and more focused on short-window finite-sample behavior. Relative to the momentum literature, it provides a pricing-kernel-based in-

terpretation of lag weighting. Relative to the integration literature, it studies whether a return-only SDF transports across markets and how such evidence should be interpreted when pricing errors are noisy in finite samples.

3 Theory: Return-Only Identification, Lag States, and Finite-Sample Implementation

This section develops the theoretical framework that underlies the paper's empirical analysis. The objective is to recover a stochastic discount factor from traded returns alone, while allowing the relevant pricing object to vary over time and to be estimated over short rolling windows. The section proceeds in four steps. First, it states the return-only conditional pricing problem and shows how the agnostic estimator arises from a purely return-based set of moment restrictions. Second, it presents a finite-state representation that makes the economic content of the agnostic estimator transparent. Third, it extends the same logic to a conditional lag-state setting in which recent return cross-sections are treated as the relevant state space at the formation date. Finally, it discusses the finite-sample bias, signal-to-noise, and regularization issues that are central to implementation in the large-N, small-T environment emphasized in the introduction.

3.1 A return-only conditional pricing problem

Let R_t denote the $N \times 1$ vector of gross returns at date t, and let Z_t denote the information available at that date. The conditional asset-pricing restriction can be written as

$$E[m(Z_t)R_t \mid Z_t] = \mathbf{1}_N, \tag{1}$$

where $m(Z_t)$ is the scalar stochastic discount factor and $\mathbf{1}_N$ is an $N \times 1$ vector of ones (Hansen and Richard, 1987; Campbell, 1993; Ferson and Siegel, 2001; Hansen and Jagannathan, 1991; Jagannathan and Wang, 1996; Lettau and Ludvigson, 2001). Equation (1) is standard, but its use here is different. Rather than positing a specific macroeconomic state vector or a parametric factor model for $m(Z_t)$, we take (1) itself as the primitive object and ask how much of the pricing kernel can be recovered directly from returns.

A convenient way to convert the conditional restriction into unconditional moment conditions is to interact returns with measurable functions of the conditioning information (Hansen and Richard, 1987; Ferson and Siegel, 2001). Let $f(Z_t)$ be an $M \times 1$ vector of measurable functions and define the normalized transform $\bar{f}(Z_t)$ element-by-element so that each component has unit mean. Then the conditional restriction implies

$$E\big[m(Z_t)\big(R_t \otimes \bar{f}(Z_t)\big)\big] = \mathbf{1}_{NM}, \tag{2}$$

where $\otimes$ denotes the Kronecker product. Thus the original N pricing conditions are replaced by NM

unconditional pricing conditions on augmented returns. This observation is useful for two reasons. First, it shows that conditioning information can be incorporated without specifying a parametric model for the SDF. Second, it clarifies that the agnostic approach is fundamentally a return-only procedure even in a conditional setting: the extra information enters only by augmenting the payoff space. This return-based perspective is also consistent with the broader distinction between return-spanned pricing kernels and structural Euler-equation models emphasized by Delikouras and Dittmar (2022).

Define the augmented return vector

$$\widetilde{R}_t := R_t \otimes \bar{f}(Z_t),$$

and stack observations over $t = 1, \ldots, T$ into the matrix

$$\widetilde{R} := \big[\widetilde{R}_1, \ldots, \widetilde{R}_T\big]',$$

so that $\widetilde{R}$ is a $T \times NM$ matrix. The agnostic estimator of the time series of SDF realizations is then

$$\hat{m}(Z) = T\big(\widetilde{R}\widetilde{R}'\big)^{-1}\widetilde{R}\mathbf{1}_{NM}. \tag{3}$$

When conditioning information is absent, $\bar{f}(Z_t) \equiv 1$ and (3) reduces to

$$\hat{m} = T\big(RR'\big)^{-1}R\mathbf{1}_N, \tag{4}$$

which is the basic return-only agnostic estimator.

Two features of (4) deserve emphasis. First, the matrix RR' is not a covariance matrix. It is the cross-sectional time-product of returns, so the estimator is driven by the pricing restriction itself rather than by a mean–variance tangency argument (Hansen and Jagannathan, 1991). Second, the vector $R\mathbf{1}_N$ appears because the right-hand side of the pricing equation is the same unit payoff for every asset in every period. The estimator therefore solves for the time series of discounting weights that prices a flat payoff across the cross section, rather than for portfolio weights on the asset space. In that sense, the agnostic estimator is a projection in the time dimension, not a projection in the space of assets.

This interpretation becomes especially transparent when one writes the sample analogue of the pricing condition as

$$\frac{R'm}{T} = \mathbf{1}_N + \bar{\varepsilon}, \tag{5}$$

where $\bar{\varepsilon}$ is the vector of average pricing surprises across assets. Premultiplying by R yields

$$\frac{RR'm}{T} = R\mathbf{1}_N + R\bar{\varepsilon},$$

so that the estimator in (4) is the solution obtained by setting the average surprise term to zero.

Equivalently, it is the minimizer of the sum of squared average surprises,

$$\min_m \left(\mathbf{1}_N - \frac{R'm}{T}\right)' \left(\mathbf{1}_N - \frac{R'm}{T}\right). \tag{6}$$

The first-order condition of (6) yields (4). This minimum-surprise interpretation is useful throughout the paper. It clarifies that the agnostic estimator is not attempting to identify a structural preference parameter. It is choosing the realized time path of the pricing kernel that makes the observed cross section look most nearly priced by a common discount factor.

3.2 A finite-state representation

The finite-state formulation makes the economic content of the return-only estimator transparent. It also provides the clearest bridge between the general agnostic estimator and the lag-state construction used later in the momentum application (Cochrane, 2014; Ross, 2015).

Consider a one-period economy with S states. Let $R \in \mathbb{R}^{S\times N}$ collect gross payoffs across N traded assets, with row s equal to the cross section of asset payoffs in state s. Let $\pi \in (0,1)^S$ be the vector of state probabilities, with $\sum_{s=1}^S \pi_s = 1$, and let $\Pi := \mathrm{diag}(\pi)$. A stochastic discount factor is a vector $m \in \mathbb{R}^S$ satisfying

$$R'\Pi m = \mathbf{1}_N. \tag{7}$$

Equation (7) says that the same discounting rule must simultaneously price the entire menu of traded payoffs.

It is useful to define the vector of state prices

$$q := \Pi m.$$

Then (7) becomes

$$R'q = \mathbf{1}_N. \tag{8}$$

This representation isolates what the cross section identifies. The payoff matrix R determines which state-contingent payoffs are spanned by traded assets; the state-price vector q is the unique vector that values those payoffs at one unit if the span is sufficiently rich.

Assumption 1. There exists at least one $q \in \mathbb{R}^S$ such that $R'q = \mathbf{1}_N$.

Assumption 2. $\mathrm{rank}(R) = S$, so that RR' is invertible.

Assumption 3. $\pi_s > 0$ for all $s = 1, \ldots, S$.

Assumption 1 is the finite-state pricing requirement. Assumption 2 is the spanning condition that makes the state-price vector unique. Assumption 3 ensures that state prices can be mapped into an SDF state by state.

Proposition 1 (Exact return-only SDF in the finite-state model)**.** *Under Assumptions 1–3, the unique solution to (7) is*

$$m = \Pi^{-1}(RR')^{-1}R\mathbf{1}_N. \tag{9}$$

If states are equiprobable, so that $\Pi = \frac{1}{S}I_S$*, then*

$$\frac{m}{S} = (RR')^{-1}R\mathbf{1}_N. \tag{10}$$

Proof. Under Assumption 2, the unique solution to (8) is

$$q = (RR')^{-1}R\mathbf{1}_N.$$

Since $m = \Pi^{-1}q$, (9) follows immediately. Equation (10) is the equiprobable special case. □

Proposition 1 is the core theoretical statement in this section. It shows that the agnostic SDF is not an opaque numerical object. It is the state-price vector implied by the traded payoff span, rescaled by probabilities. The term $(RR')^{-1}R\mathbf{1}_N$ is the unique linear combination of state payoff vectors that prices a unit payoff across all traded assets. Multiplying by Π^{-1} converts those state prices into marginal-utility weights by dividing out probabilities state by state. States with high prices relative to probabilities receive high SDF weight.

This representation also clarifies why the agnostic estimator is naturally described as return-only. Equation (10) depends only on the payoff matrix and a normalization. No factor structure, preference specification, or external state variable is required. The estimator simply asks what vector of discounting weights prices the observed payoff span.

3.3 From finite states to lag states

The finite-state representation becomes operational once recent return cross-sections are treated as the relevant state space for short-horizon estimation. This is the key step that links the agnostic estimator to the momentum application (Cochrane, 2014; Jegadeesh and Titman, 1993; Asness et al., 2013).

Fix a formation date $\tau - 1$ and a window length $K \geq 1$. Let

$$\mathcal{F}_{\tau-1}^{(K)} := \sigma(R_{\cdot,\tau-1}, R_{\cdot,\tau-2}, \ldots, R_{\cdot,\tau-K})$$

denote the information generated by the last K return cross-sections. Interpret each lagged cross-section as a state and stack them as rows of the matrix

$$R_{\tau-1}^{(K)} \in \mathbb{R}^{K\times N},$$

with row s equal to the cross section $R'_{\cdot,\tau-s}$ for $s = 1, \ldots, K$. Let $\pi_{\tau-1}^{(K)} \in (0,1)^K$ be a vector of

lag-state probabilities measurable with respect to $\mathcal{F}_{\tau-1}^{(K)}$ and define

$$\Pi_{\tau-1}^{(K)} := \operatorname{diag}\left(\pi_{\tau-1}^{(K)}\right).$$

A conditional lag-state SDF is a vector $m_{\tau-1}^{(K)} \in \mathbb{R}^K$ such that

$$\left(R_{\tau-1}^{(K)}\right)' \Pi_{\tau-1}^{(K)} m_{\tau-1}^{(K)} = \mathbf{1}_N. \tag{11}$$

Equation (11) is the direct analogue of the stacked pricing system in the finite-state model. It requires a single set of lag weights to price the entire cross section inside the K-month window. In this sense, recent returns are not merely predictive covariates. They are treated as the conditioning set that must support a pricing kernel at the formation date (Hansen and Richard, 1987; Ferson and Siegel, 2001).

Proposition 2 (Exact conditional SDF on a K-month window)**.** *Suppose* $\mathbf{1}_N \in \operatorname{col}\left(\left(R_{\tau-1}^{(K)}\right)'\right)$, $\operatorname{rank}\left(R_{\tau-1}^{(K)}\right) = K$*, and all lag-state probabilities are strictly positive. Then the unique solution to* (11) *is*

$$m_{\tau-1}^{(K)} = \left(\Pi_{\tau-1}^{(K)}\right)^{-1} \left(R_{\tau-1}^{(K)} \left(R_{\tau-1}^{(K)}\right)'\right)^{-1} R_{\tau-1}^{(K)} \mathbf{1}_N. \tag{12}$$

If lag states are equiprobable, so that $\Pi_{\tau-1}^{(K)} = \frac{1}{K} I_K$*, then*

$$\frac{m_{\tau-1}^{(K)}}{K} = \left(R_{\tau-1}^{(K)} \left(R_{\tau-1}^{(K)}\right)'\right)^{-1} R_{\tau-1}^{(K)} \mathbf{1}_N. \tag{13}$$

Proof. The proof is identical to that of Proposition 1, with lagged return cross-sections taking the place of primitive states of nature. □

Proposition 2 is the formation-date object used later in the paper. It shows that the agnostic SDF over a short rolling window can be interpreted as a vector of lag weights. Large entries of $m_{\tau-1}^{(K)}$ correspond to lagged cross-sections that are more important for pricing a flat payoff at date $\tau - 1$. Hence the conditional SDF provides an economically interpretable ranking of recent return histories according to their pricing relevance. Specifically, over a short rolling window, the conditional lag-state SDF can be read as a set of relative state prices across recent lags. Interpreting lags as states, a larger component of $m_{\tau-1}^{(K)}$ means that the corresponding lagged return realization is more pricing-relevant for valuing the cross-section at the formation date. The vector $m_{\tau-1}^{(K)}$ therefore does more than assign statistical weights to past returns: it summarizes which parts of recent return history the market is implicitly pricing more heavily at $\tau - 1$.

Three remarks are useful. First, the lag-state SDF is measurable with respect to information available at the formation date. That is essential for implementation. Second, the construction is still return-only. The conditioning information is the realized return history itself. Third, the

equiprobable specification in (13) is not merely a convenience. It is the natural analogue of the original agnostic estimator once recent lags are interpreted as states rather than as calendar dates.

3.4 The history direction and the SDF-weighted signal

The lag-state SDF lives in a K-dimensional lag space. To connect that object to momentum, it is useful to introduce the corresponding cross-sectional history space.

For each lag $s = 1, \dots, K$, define the cross-sectional mean return

$$\bar{R}_{\tau-s} := \frac{1}{N}\mathbf{1}_N' R_{\cdot,\tau-s},$$

and the demeaned lagged cross-section

$$R^e_{\cdot,\tau-s} := R_{\cdot,\tau-s} - \bar{R}_{\tau-s}\mathbf{1}_N.$$

Stack these demeaned lagged cross-sections as columns of the $N \times K$ history matrix

$$X_\tau := \left[R^e_{\cdot,\tau-1}\ R^e_{\cdot,\tau-2}\ \cdots\ R^e_{\cdot,\tau-K}\right].$$

By construction,

$$X_\tau'\mathbf{1}_N = 0.$$

Let $y_\tau := R_{\cdot,\tau}$ denote the current return cross section.

The history matrix summarizes how recent return histories vary across assets. A convenient reduced-form representation of the current cross section is

$$y_\tau = \mu_\tau \mathbf{1}_N + X_\tau \gamma_\tau + e_\tau, \tag{14}$$

where $\gamma_\tau \in \mathbb{R}^K$ is the direction in lag space most closely associated with the cross-sectional variation in y_τ and e_τ is the component of the current cross section not explained by recent histories. In a large cross section it is natural to require that the residual has little systematic covariance with the lag space,

$$\|X_\tau' e_\tau\| = o(1) \quad \text{as } N \to \infty. \tag{15}$$

Under (15), the span of recent demeaned lagged cross-sections captures the pricing-relevant history directions at date τ.

Proposition 3 (History direction at date τ)**.** *Define*

$$m_\tau^{\text{hist}} := (X_\tau' X_\tau)^{-1} X_\tau' y_\tau. \tag{16}$$

If $X_\tau' X_\tau$ is invertible and its smallest eigenvalue is bounded away from zero as $N \to \infty$, then

$$\|m_\tau^{\text{hist}} - \gamma_\tau\| = \|(X_\tau' X_\tau)^{-1} X_\tau' e_\tau\| = o(1). \tag{17}$$

Proposition 3 describes an infeasible benchmark. It uses the realized date-τ return cross section to recover the direction of continuation in lag space. That object is useful conceptually but not implementable, because y_τ is not known at $\tau - 1$. The lag-state SDF supplies a feasible proxy that lives in the same K-dimensional lag space and is measurable at the formation date.

Using the conditional SDF to aggregate each stock's own demeaned lagged returns yields the SDF-weighted history statistic

$$P_{i,\tau-1}^{(K)} := \sum_{s=1}^{K} m_{\tau-1,s}^{(K)} R_{i,\tau-s}^{e} = \left(x_{i,\tau}^{e}\right)' m_{\tau-1}^{(K)}, \tag{18}$$

where

$$x_{i,\tau}^{e} := \left(R_{i,\tau-1}^{e}, \ldots, R_{i,\tau-K}^{e}\right)'.$$

Because lagged returns are demeaned lag by lag,

$$\sum_{i=1}^{N} P_{i,\tau-1}^{(K)} = 0.$$

Hence the signal is inherently cross-sectional and naturally suited for long-short portfolio formation.

This demeaning step is economically important. Momentum is a relative-performance strategy (Jegadeesh and Titman, 1993; Moskowitz and Grinblatt, 1999). Removing the common component of each lagged cross section ensures that the signal targets zero-investment variation across stocks rather than the level of the market. The resulting ranking variable remains momentum-like—it is still a weighted sum of lagged returns—but the lag profile is no longer fixed ex ante. It is instead disciplined by the pricing kernel (Huang, 2022; Jegadeesh et al., 2025).

Consider now the cross-sectional regression with an intercept,

$$R_{i,\tau} = a_\tau + b_\tau P_{i,\tau-1}^{(K)} + \varepsilon_{i,\tau}. \tag{19}$$

The associated OLS slope satisfies

$$b_\tau = \frac{\left(m_{\tau-1}^{(K)}\right)' X_\tau' y_\tau}{\left(m_{\tau-1}^{(K)}\right)' X_\tau' X_\tau m_{\tau-1}^{(K)}} = \frac{\left\langle m_{\tau-1}^{(K)}, m_\tau^{\text{hist}} \right\rangle_{X_\tau' X_\tau}}{\| m_{\tau-1}^{(K)} \|_{X_\tau' X_\tau}^2} + o(1), \tag{20}$$

where $\langle u, v \rangle_A := u' A v$ and $\|u\|_A^2 := u' A u$. Thus the slope is large when the feasible lag-state SDF is closely aligned with the realized history direction. Equivalently, the conditional SDF is informative for momentum to the extent that the lag profile required to price the recent window is also the lag profile that best summarizes subsequent cross-sectional continuation. This alignment condition is the economically relevant one for the momentum application. A positive payoff does not require exact recovery of the SDF in level. What matters is whether the feasible lag-state SDF recovers the continuation-relevant direction in lag space well enough to rank stocks along the part of recent

history that subsequently persists. In that sense, the empirical content of the strategy comes from directional recovery rather than from exact in-sample pricing of the formation-window returns.

Equation (20) also yields an implementable mimicking-portfolio interpretation. Let

$$p_\tau := X_\tau m_{\tau-1}^{(K)}.$$

Because $X_\tau' \mathbf{1}_N = 0$, the portfolio with weights

$$w_\tau := \frac{p_\tau}{p_\tau' p_\tau}$$

satisfies $w_\tau' \mathbf{1}_N = 0$. Its realized return is

$$r_\tau^{\text{mom}} = w_\tau' y_\tau = \frac{p_\tau' y_\tau}{p_\tau' p_\tau} = b_\tau.$$

Hence the cross-sectional regression slope is exactly the payoff on the factor-mimicking portfolio for the SDF-implied signal. This is the theoretical bridge between the conditional lag-state SDF and the momentum strategy studied later in the paper.

3.5 Why the agnostic estimator is biased in short windows

The preceding results are exact conditional or finite-state characterizations. Empirical implementation, however, takes place in short samples, and finite-sample distortions are first-order. This subsection develops the two key issues: bias and signal-to-noise. Formal consistency arguments for the original agnostic estimator are developed in Kim (2017).

The starting point is the minimum-surprise representation in (6). Because the dependent object in that least-squares problem is the vector of ones, the resulting estimator is not a standard regression coefficient with orthogonal regressors and disturbances (Britten-Jones, 1999). Indeed, writing the exact sample relation as

$$R' m / T = \mathbf{1}_N + \bar{\varepsilon},$$

one obtains

$$\hat{m} - m = -T(RR')^{-1} R\bar{\varepsilon}.$$

The bias arises because the regressors and the pricing surprises are mechanically linked in finite samples. At a given date, the pricing error for asset i is constructed from the same return $R_{i,t}$ that appears in the regressor matrix. When T is large, the effect of this dependence on average pricing surprises is attenuated. When T is short, it need not be (Britten-Jones, 1999; Kim and Korajczyk, 2024; Kim, 2017).

This feature has an important empirical implication. A direct finite-sample test of the Euler restriction can be difficult to interpret because the estimated SDF itself is noisy and biased (Kim

and Korajczyk, 2024). That is why the empirical sections of the paper emphasize objects whose economic meaning depends primarily on the direction and stability of the estimated SDF rather than on unbiased pointwise pricing-error measurement. The momentum application is one such object. The domestic-versus-international pricing comparison is another, provided that pricing error metrics are interpreted in light of the finite-sample noise they contain.

The signal-to-noise problem can be described more explicitly under a simple decomposition of augmented returns. Suppose

$$\widetilde{R}_t = \alpha + BF_t + v_t, \tag{21}$$

where F_t is a low-dimensional vector of systematic pricing directions and v_t is the idiosyncratic component (Ross, 1976; Kelly et al., 2019). Then

$$\widetilde{R}\widetilde{R}' = F'B'BF + V'V + \text{cross terms}.$$

When the cross section is large, the cross terms are small under standard averaging arguments, and the agnostic estimator is driven by the balance between the systematic term $F'B'BF$ and the idiosyncratic term $V'V$. This provides a useful signal-to-noise interpretation. The systematic component is the signal: it contains the pricing-relevant directions in returns. The idiosyncratic component is the noise: it inflates the matrix that must be inverted to recover the SDF.

Under the additional approximation that aggregate idiosyncratic volatility is not time-varying and that the idiosyncratic component has little serial dependence, one has

$$V'V \approx \psi^2 I_T,$$

where ψ^2 is aggregate idiosyncratic variance. In that case, the agnostic estimator can be written heuristically as

$$\hat{m}(Z) \approx T\big(F'B'BF + \psi^2 I_T\big)^{-1} F'B'\mathbf{1}_{NM}, \tag{22}$$

up to terms that vanish with cross-sectional aggregation. Equation (22) formalizes the basic implementation problem. When the signal component $F'B'BF$ is strong relative to $\psi^2 I_T$, the pricing kernel is identified sharply. When idiosyncratic noise is large, the estimator is heavily shrunk toward smoother time profiles and becomes less informative.

3.6 Ridge regularization and the bias–estimation-error tradeoff

The same decomposition yields a natural ridge interpretation of the agnostic estimator. Under the approximation above, (3) is equivalent to the coefficient vector in a ridge-type problem in which the pricing kernel is penalized in proportion to aggregate idiosyncratic variance. This interpretation is useful because it clarifies both the appeal and the cost of regularization (Kozak et al., 2020; Lettau and Pelger, 2020).

A regularized version of the conditional agnostic estimator is

$$\hat{m}_\kappa(Z) = T\big(\widetilde{R}\widetilde{R}' + \kappa_T I_T\big)^{-1}\widetilde{R}\mathbf{1}_{NM}, \tag{23}$$

where $\kappa_T \geq 0$ is a shrinkage parameter. In the lag-state application, the corresponding regularized estimator is

$$\frac{\hat{m}^{(K)}_{\tau-1}}{K} = \Big(R^{(K)}_{\tau-1}\big(R^{(K)}_{\tau-1}\big)' + \kappa_K I_K\Big)^{-1} R^{(K)}_{\tau-1}\mathbf{1}_N. \tag{24}$$

Regularization addresses two related problems. First, when the effective rank of the return panel is low, RR' or $R^{(K)}_{\tau-1}(R^{(K)}_{\tau-1})'$ can be ill-conditioned. Second, when the number of assets is not much larger than the length of the estimation window, the estimator can be overly sensitive to idiosyncratic noise. Adding κI stabilizes the inversion and lowers estimation variance. The cost is additional bias.

This tradeoff is particularly sharp in the large-N, small-T region emphasized in this paper. Holding the cross section fixed, a longer estimation window reduces the finite-sample bias in pricing because average surprises are taken over more periods. At the same time, a longer window can worsen the signal-to-noise ratio of the inversion problem if the systematic component of returns is low rank relative to the time dimension. In the decomposition above, increasing T reduces the relative importance of the idiosyncratic term for bias but can make the inverse problem less stable because the systematic matrix need not fill the expanded time dimension. Conversely, shortening the window can improve stability and reduce estimation error, but at the cost of greater bias in pricing restrictions. This is the central econometric tension in the paper (Kozak et al., 2020; Kim and Korajczyk, 2024).

The importance of this point cannot be overstated. The simulation framework in Section 4 is designed precisely to map out this bias–estimation-error frontier in the empirical region in which the estimator is actually used. The domestic pricing exercise, the international pricing comparison, and the momentum strategy all operate on short rolling windows. Their interpretation therefore depends on understanding how the agnostic estimator behaves when the cross section is large but the effective time dimension is limited.

3.7 Summary and implications for the empirical sections

The main message of this section is that the agnostic SDF admits a transparent return-only representation. In a finite-state economy, the cross section identifies state prices directly, and the SDF follows by scaling state prices with probabilities. When recent return cross-sections are interpreted as states, the same pricing logic yields a conditional lag-state SDF that is measurable at the formation date and that naturally takes the form of a vector of lag weights. Those lag weights, in turn, produce an SDF-weighted history statistic that is centered in the cross section and leads directly to an implementable momentum signal.

At the same time, empirical implementation over short windows raises a nontrivial finite-sample problem. The agnostic estimator is biased in short samples because the pricing surprises are

mechanically linked to the return panel used to estimate the SDF. Its precision also depends on the signal-to-noise ratio of the return panel: systematic pricing directions sharpen identification, while idiosyncratic variation inflates the inverse problem. Regularization improves stability but adds bias. These considerations explain why the rest of the paper emphasizes economically interpretable applications of the estimated pricing kernel and why the simulation framework in Section 4 is organized around the large-N, small-T regime rather than around generic asymptotics.

4 Simulation Design

This section describes the Monte Carlo framework used to evaluate the finite-sample properties of the return-only agnostic estimator in the environments that matter for the paper's two empirical applications. The simulations are not intended as a generic large-sample validation exercise. Their purpose is narrower and more practical: to quantify how the estimator behaves when the cross section is rich, the relevant estimation window is short, and the econometrician wishes to compare a return-only SDF with structured alternatives. The design therefore targets the same large-N, small-T region that appears in the domestic pricing exercise, the international pricing comparison, and the conditional momentum application.

The simulation framework has four objectives. First, it quantifies the bias–estimation-error tradeoff emphasized in Section 3. Second, it compares the agnostic estimator with structured alternatives under correct specification and under factor misspecification. Third, it studies whether the lag-state SDF recovers the continuation-relevant direction in settings designed to mimic the momentum application. Fourth, it evaluates the extent to which cross-sectional pricing metrics can favor mechanically simple benchmarks, such as a constant SDF, when the evaluation window is short and the candidate SDF is time varying.

4.1 Data-generating framework

All simulation environments begin from a common monthly panel of returns,

$$R_{it} = \alpha_i + \beta_i' F_t + u_{it}, \tag{25}$$

where F_t is a low-dimensional vector of systematic pricing directions and u_{it} is the idiosyncratic component. The baseline calibration uses the same empirical ingredients that motivate the paper's applications: the volatility of observed factor returns, the scale of cross-sectional loadings, and the residual dispersion of broad portfolio or stock-level returns. To preserve continuity with the return-only perspective developed in Section 3, the true pricing kernel is defined as $m_t^\star = \delta_0 + \delta_F' F_t$, with (δ_0, δ_F) chosen so that the Euler restrictions hold in population. This construction allows the simulations to separate two questions that are distinct in the paper: whether an estimator recovers the time path of the true SDF, and whether it produces economically useful pricing or ranking objects even when it does not recover that path exactly.

The baseline grid for the time dimension is $T \in \{12, 24, 60, 120, 240\}$, which spans the short rolling windows used in the empirical analysis and the longer windows that help reveal the bias frontier. The cross-sectional grid is $N \in \{500, 1{,}500, 2{,}500, 10{,}000\}$. The smaller values mimic portfolio-based construction and evaluation environments; the larger values mimic stock-level SDF construction.

4.2 Simulation environments

The framework is organized around five environments, each designed to support a specific claim in the empirical sections.

The first environment studies the bias–estimation-error frontier directly, mapping out how bias, estimation error, and regularization interact in the region in which the estimator is actually used. The second is tailored to the momentum application: it asks whether the lag-state SDF recovers the continuation-relevant direction in lag space even when the SDF itself is biased or noisy in level. The third gives structured methods such as KK their best case by imposing a correctly specified common factor structure. The fourth studies misspecification, segmentation, and instability—the environments in which the agnostic estimator should have its greatest comparative advantage. The fifth mirrors the domestic and international comparisons in Section 6, studying whether pricing-error metrics favor the same method when the pricing kernel is transported to a shifted asset universe and when evaluation histories are short or uneven.

4.3 Evaluation criteria

The simulation evaluates estimators using three objects corresponding to the three roles played by the SDF in the paper. SDF recovery is summarized by the regression $m_t^\star = a_m + b_m \hat{m}_t + \nu_t$ and its R^2. Direct pricing performance is summarized by the cross-sectional mean squared pricing error and the bias–variance decomposition of pricing errors. Directional and signal recovery, which is central for the momentum application, is summarized by the correlation between the feasible lag-state SDF and the true continuation direction, the rank correlation between feasible and infeasible cross-sectional scores, and the average return on long–short portfolios formed on the feasible score.

Full simulation results are in preparation and will appear in the next version of the paper. The empirical applications that follow should therefore be interpreted in light of the econometric considerations developed in Section 3 rather than as claims already supported by completed Monte Carlo evidence.

5 Empirical Application I: Momentum from the Conditional SDF

This section studies the first empirical application of the return-only agnostic estimator developed in Section 3. The central object is the lag-state conditional SDF. At each formation date, the estimator uses a short window of recent return cross-sections to recover a vector of lag weights, and those weights are then mapped into a one-dimensional cross-sectional signal. The resulting

strategy is momentum-like in a familiar sense—it ranks stocks on the basis of their recent return histories—but the lag profile is not fixed ex ante. Instead, it is disciplined by the pricing restriction that identifies the conditional SDF. This is the main economic distinction between the present construction and conventional momentum strategies based on a predetermined formation window (Jegadeesh and Titman, 1993; Moskowitz and Grinblatt, 1999; Asness et al., 2013).

The section proceeds in four parts. First, it describes the empirical implementation and relates the feasible signal to the infeasible continuation benchmark discussed in Section 3. Second, it shows how the signal can be interpreted through lagged, cross-sectionally demeaned pricing errors and explains why this application is substantially less sensitive to the finite-sample pricing-bias problem that arises in direct Euler-equation tests. Third, it reports baseline results and simple implementation variants. Fourth, it examines where in the cross-section the SDF-weighted momentum signal is strongest by studying characteristic-specific heterogeneity.

5.1 Empirical design, timing, and signal construction

The empirical implementation follows the two-step logic developed in Section 3. The stock universe is a screened monthly CRSP-style panel of common stocks. The reported sample runs from 1930:07 to 2024:11, and the SDF step is implemented only in months with at least 500 stocks. The baseline lag window is $K = 12$ months. This design keeps the empirical implementation aligned with the large-cross-section, short-window environment emphasized in the theory and simulation sections (Jegadeesh and Titman, 1993; Asness et al., 2013).

At each month t, the lag-state conditional SDF is estimated using only the most recent K return cross-sections available at formation. Let $\hat{m}_t = (\hat{m}_{t,1}, \ldots, \hat{m}_{t,K})'$ denote the resulting vector of lag weights. Consistent with Section 3, those lag weights summarize which parts of recent return history are most relevant for pricing the cross-section at formation month t. Given $\hat{m}_t$, the stock-level SDF-weighted history score is

$$S_{i,t} = \sum_{k=1}^{K} \hat{m}_{t,k}\, \widetilde{R}_{i,t+1-k}, \tag{26}$$

where $\widetilde{R}_{i,t+1-k}$ is stock i's lagged return demeaned cross-sectionally month by month. The demeaning matters because it removes the level component of the market and maps the lag-state SDF into a self-financing long–short ranking variable. In the empirical implementation, the raw score is standardized within month,

$$Z_{i,t} = \frac{S_{i,t} - \bar{S}_t}{\mathrm{sd}_t(S_{i,t})}, \tag{27}$$

and $Z_{i,t}$ is used both for portfolio sorts and for cross-sectional regressions.

This construction is most easily understood relative to the infeasible benchmark discussed earlier. The ideal reduced-form object is the continuation direction in lag space: the linear combination of recent lagged returns that best summarizes the next cross-section of returns. That object is not

observable at formation because it depends on realized date-$t+1$ returns. The role of the conditional lag-state SDF is to provide a feasible proxy in the same lag space. Rather than being identified from a predictive regression, the vector $\hat{m}_t$ is identified from the pricing restriction using only information available at time t. The empirical question is whether that feasible lag direction contains enough information about subsequent continuation to generate a positive average payoff (Hansen and Richard, 1987; Ferson and Siegel, 2001).

The signal is evaluated in two complementary ways. First, each month stocks are sorted into deciles on $Z_{i,t}$ and the next-month high-minus-low spread is computed as the return on the top decile minus the return on the bottom decile. Second, we run the monthly cross-sectional regression

$$R_{i,t+1} = a_{t+1} + b_{t+1} Z_{i,t} + \varepsilon_{i,t+1}, \tag{28}$$

and summarize the signal's marginal predictive content through the Fama–MacBeth average of the slope coefficient (Fama and MacBeth, 1973). The portfolio-spread and regression-slope approaches carry the same economic content. The spread emphasizes implementability; the slope isolates the per-unit mapping from the SDF-weighted history direction into subsequent expected returns.

5.2 Pricing errors, residual momentum, and the ex-ante nature of the signal

A useful way to interpret the SDF-weighted momentum signal is through pricing errors. For each stock i and date t, define the realized Euler pricing error

$$\mathrm{PE}_{i,t} := \hat{m}_t R_{i,t} - 1,$$

and the corresponding cross-sectionally demeaned component

$$\widetilde{\mathrm{PE}}_{i,t} := \hat{m}_t \widetilde{R}_{i,t},$$

where $\widetilde{R}_{i,t}$ denotes the demeaned return used throughout the section. Since the cross-sectional mean is removed, $\widetilde{\mathrm{PE}}_{i,t}$ retains only the self-financing component of the pricing residual. Under this notation, the SDF-weighted signal in (26) can be written as

$$S_{i,t} = \sum_{k=1}^{K} \hat{m}_{t,k} \widetilde{R}_{i,t+1-k} = \sum_{k=1}^{K} \widetilde{\mathrm{PE}}_{i,t+1-k}. \tag{29}$$

Thus the signal is a cumulated history of lagged, demeaned, SDF-weighted residual performance.

Equation (29) provides a natural link to residual momentum. Standard residual-momentum strategies first specify a factor model, remove the factor-spanned component of returns, and then apply a momentum signal to the resulting residuals (Blitz et al., 2011; Asness et al., 2013). The present approach reaches a closely related object without specifying a factor model. Instead, the relevant history direction is identified directly from the return-only pricing kernel. In that sense, residual momentum is nested inside a more general SDF-based ranking rule.

The close connection to residual momentum is economically natural and should not be viewed as a weakness of the present construction. Both objects aim to isolate continuation in the self-financing component of returns rather than in raw returns. The difference is that conventional residual momentum first specifies a factor model and then applies a fixed lag-window ranking rule to the resulting residuals, whereas the present signal uses the lag-state conditional SDF to determine ex ante which parts of recent return history are most pricing-relevant at the formation date. Thus the signal is best interpreted as a theory-disciplined, state-dependent residual-momentum object. Under that interpretation, one should not expect the SDF-weighted signal to be uniformly strong in every residual-momentum environment. Instead, its payoff can be concentrated in the parts of the cross-section where residual continuation is already pronounced, while weakening—or even turning negative—in regions where recent residual-return dynamics are dominated by correction, illiquidity, or reversal. That pattern would indicate conditional alignment with residual momentum, not that the signal is mechanically identical to a standard residual-momentum sort.

This interpretation also clarifies why the present application is substantially less sensitive to the finite-sample pricing-bias problem emphasized in direct tests of the Euler equation (Kim and Korajczyk, 2024). The most bias-sensitive exercise is an in-sample test of whether the estimated SDF prices the same return panel used to construct it. In that setting, the estimated pricing kernel and the pricing-error statistic are mechanically linked, and small-sample bias can strongly distort inference. The momentum application is different. The lag-state SDF is estimated at month t using only information available at formation, and the resulting score is then used to forecast the return cross-section at month $t+1$. The signal is therefore predetermined with respect to the next return innovation. The empirical question is not whether the formation window is priced exactly in sample, but whether the feasible lag profile extracted from the pricing restriction contains information about subsequent cross-sectional continuation.

Three distinctions are especially important. First, the object of interest is the *direction* selected by the lag-state SDF, not the level of the in-sample pricing error. Second, the signal is formed ex ante and evaluated on future returns, rather than on the same realization used to estimate the SDF. Third, the empirical signal is constructed from cross-sectionally demeaned lagged returns, so common level distortions in the estimated SDF are much less consequential than they would be in a direct pricing-error calculation. This does not mean that estimation error is irrelevant. If the estimated lag-state SDF is noisy, the ranking variable will be attenuated and the resulting long–short spread will be weaker. But that is an accuracy problem rather than the direct finite-T pricing-bias problem that arises in in-sample Euler-equation tests (Kim and Korajczyk, 2024).

5.3 Baseline results and implementation variants

We begin with three implementations of the SDF-weighted momentum strategy.

1. **Current-window strategy.** The baseline specification uses the contemporaneous lag-state SDF $\hat{m}_t$ estimated at the formation month.

2. **Past-estimated strategy.** This variant uses stale weights $\hat{m}_{t-1}$ from the previous month. It is a natural diagnostic because it weakens the alignment between the current pricing-relevant lag profile and the next return cross-section.

3. **Managed strategy.** This variant scales exposure using a state variable derived from the shape of the lag-weight vector, defined below.

To summarize whether the positive weight mass is concentrated in the most recent lags, define the recency-weight ratio

$$\text{RecencyRatio}_t = \frac{\sum_{k=1}^{2} \max(\hat{m}_{t,k}, 0)}{\sum_{k=3}^{12} \max(\hat{m}_{t,k}, 0) + \varepsilon}, \qquad \varepsilon = 10^{-8}. \tag{30}$$

A higher value indicates that, relative to the medium-horizon region (months 3–12), the lag-state SDF places more nonnegative emphasis on the most recent lags. Economically, RecencyRatio_t is a compact summary of how front-loaded the SDF-implied history relevance is at time t.

Using this state variable, the managed long–short payoff is

$$r_{t+1}^{\text{managed}} = g_t \, r_{t+1}^{\text{LS}}, \qquad g_t = \min\{1, c/\text{RecencyRatio}_t\}, \tag{31}$$

for a fixed constant $c > 0$. This rule reduces exposure when the estimated lag profile becomes unusually front-loaded. The economic interpretation is not that short-horizon dynamics are necessarily undesirable. Rather, when they dominate the lag structure, the strategy may be more exposed to transient price pressure, liquidity shocks, or reversal-like behavior that are less consistent with the canonical medium-horizon continuation mechanism (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016; Huang, 2022; Lou and Polk, 2022).

Table 1 reports mean return, volatility, Sharpe ratio, expected shortfall at the 5% level, and the Newey–West ($L = 6$) t-statistic for the mean. The current-window strategy produces a positive and statistically significant payoff of 43.5 basis points per month. Using stale weights reduces both the mean and the Sharpe ratio, which is consistent with economically meaningful time variation in the lag profile. The managed variant gives up some average return but lowers volatility and improves left-tail behavior, suggesting that the shape of the lag-state SDF contains information about the downside properties of the long–short spread (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016).

To connect the lag-weight shape to downside risk more directly, we estimate a univariate predictive regression of the crash indicator

$$\mathbf{1}\{r_{t+1} < -2\hat{\sigma}_t\}$$

on RecencyRatio_t, where $\hat{\sigma}_t$ is a rolling volatility estimate of the baseline long–short payoff. Table 2 reports a negative and statistically significant slope. In the present sample, higher values of the recency-weight ratio are associated with a lower incidence of large negative payoffs under this

crash definition. A useful economic summary is that the shape of the lag-state SDF contains implementation-relevant information about the strategy's tail behavior (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016; Huang, 2022).

Two implementation lessons follow. First, the strategy is meaningfully sensitive to whether the lag-state weights are updated contemporaneously or used with a lag, which is exactly what one would expect if the pricing relevance of different lags changes over time. Second, the recency-weight ratio—a simple summary of the shape of the estimated lag profile—contains systematic information about crash realizations, and an exposure rule based on that state variable improves left-tail outcomes in the reported sample.

5.4 Momentum effects across firm characteristics

The baseline evidence establishes that the conditional lag-state SDF generates an economically meaningful momentum signal on average. A natural next question is where in the cross-section that signal is strongest. The goal of this subsection is explicitly economic: to learn in which parts of the cross-section the continuation mechanism is actually operative. The point is not merely to document heterogeneity in average spread magnitudes. A positive within-bin spread or slope indicates that, in that characteristic environment, the lag-state SDF is isolating a direction of recent history that is rewarded through continuation. By contrast, a weak or negative within-bin payoff indicates that the same signal is more naturally read as capturing correction or reversal dynamics rather than persistent continuation. This continuation-versus-correction distinction is central to the economic interpretation of the double sorts reported below.

A large literature studies whether momentum is stronger among large versus small firms, liquid versus illiquid stocks, and high versus low past-return firms (Jegadeesh and Titman, 1993; Moskowitz and Grinblatt, 1999; Hong et al., 2000; Hvidkjaer, 2006; Hoberg et al., 2022). The standard summary is a double sort: first sort firms into characteristic bins and then sort within each bin on a momentum signal. We adopt that design here, but the comparison is sharper because we evaluate both the SDF-weighted signal and the conventional momentum benchmark in the same characteristic environments.

Two considerations motivate the exercise. First, the SDF-weighted signal is formed from the lag-state pricing kernel, so it can alter the economic geography of momentum profits relative to a conventional fixed-window construction. Second, within-bin portfolio spreads can blur interpretation. A large spread may reflect genuine continuation in that environment, or it may arise simply because the top and bottom portfolios are far apart in the signal. For that reason, we report both within-bin high–minus–low portfolio spreads and within-bin cross-sectional regression slopes.

5.5 Design and reporting conventions

Each month, stocks are first sorted into quintiles by a characteristic $X_{i,t}$. Within each quintile, stocks are then sorted into deciles by the signal of interest, and the next-month high–minus–low return spread is computed. The baseline signal is the SDF-weighted score $Z_{i,t}$. For comparison, we

repeat the same exercise using traditional momentum,

$$\text{Mom}(3\text{–}12),$$

defined as the average return from months $t-11$ to $t-2$, excluding the two most recent months (Jegadeesh and Titman, 1993).

We report the evidence in three complementary ways.

1. **Within-quintile portfolio spreads using the SDF-weighted signal.** These show where the conditional SDF generates continuation profits or reversals across the characteristic space.
2. **The same within-quintile portfolio spreads using traditional momentum.** This provides a direct benchmark in exactly the same bins.
3. **Within-quintile regression slopes using the SDF-weighted signal.** These measure the marginal payoff to moving along the SDF-weighted momentum dimension in a given characteristic environment.

Panel A uses CRSP-only characteristics. Panel B uses Chen–Zimmerman characteristics (Chen and Zimmermann, 2022). All reported entries are time-series means in basis points per month, with Newey–West ($L = 6$) t-statistics in parentheses.

5.6 Results with the SDF-weighted signal

Table 3 reports the within-quintile high–minus–low spreads for the SDF-weighted signal. The most important feature of the table is that continuation is not uniform across the cross-section. Instead, the payoff to the SDF-weighted signal varies systematically across characteristic environments, and in some regions the signal weakens sharply or flips into reversal.

Several cases are especially informative. For turnover, the within-bin spread rises across quintiles, which is consistent with continuation being stronger in more actively traded stocks (Hvidkjaer, 2006; Hoberg et al., 2022). For size and price proxies, the spread is negative or weak in the smallest and lowest-price quintiles and turns strongly positive in larger and higher-price segments, a pattern consistent with microstructure noise and episodic price pressure being more severe among small and low-priced stocks (Hong et al., 2000). For illiquidity (`Amihud` and `zerotrade1M`), the spread deteriorates markedly in the most illiquid quintiles and becomes negative in the most illiquid region. This is exactly the type of pattern one would expect if continuation is weaker where frictions, underreaction-to-correction cycles, or reversal-like dynamics are more important (Hvidkjaer, 2006; Lou and Polk, 2022).

The Chen–Zimmerman characteristics convey the same message in sharper form. For `Mom12m`, the SDF-weighted spread is strongly negative in low-momentum quintiles and close to zero in the highest quintile. For `ResidualMomentum`, the spread increases monotonically across quintiles, moving from negative values in the low-residual-momentum bins toward positive values only in the

highest quintile. This pattern confirms that the SDF-weighted signal is closely related to residual-return dynamics, but it does not imply that the signal is merely residual momentum in disguise. A more natural interpretation is conditional alignment: the SDF-weighted ranking variable is most effective when the cross-section already exhibits a strong residual-continuation component, whereas in low-residual-momentum environments the same lag-history direction is more likely to reflect correction or reversal (Blitz et al., 2011). For `zerotrade1M`, the spread is strongest in the most liquid quintile and turns negative in the most illiquid quintile. Taken together, these results suggest that the conditional SDF should not be viewed as delivering a uniformly positive continuation signal across all characteristic environments. Rather, it sorts stocks along a lag-history direction whose payoff depends strongly on the economic environment. The close relation to residual momentum is economically natural, but the characteristic dependence indicates that the SDF-based signal is better understood as a state-dependent continuation signal than as a mechanical relabeling of a standard momentum or residual-momentum sort.

5.7 Traditional momentum in the same characteristic bins

Table 4 repeats exactly the same design using conventional momentum, Mom(3–12), as the within-quintile sort variable. This benchmark is useful because it shows which parts of the cross-section deliver continuation under the standard fixed-lag construction (Jegadeesh and Titman, 1993; Asness et al., 2013).

Relative to Table 3, the traditional benchmark often generates broadly positive and more uniform within-bin spreads. That is informative, but it is also economically less discriminating. Because the lag window is fixed for all stocks and all dates, it can blur the possibility that some parts of the cross-section are governed more by correction and reversal than by continuation (Jegadeesh et al., 2025).

Two comparisons stand out. First, for `Mom12m`, the conventional signal shows only modest heterogeneity across quintiles, whereas the SDF-weighted signal produces a large and statistically strong contrast. Second, for `zerotrade1M`, the conventional signal deteriorates only moderately in illiquid stocks, while the SDF-weighted signal weakens sharply and turns negative in the most illiquid quintile. That sharper pattern is consistent with the idea that the lag-state SDF is better able to reveal where the continuation mechanism is genuinely operative and where short-horizon distortions dominate.

5.8 Per-unit slope heterogeneity within characteristic bins

Portfolio spreads are useful for investors, but they do not fully isolate the economic strength of the continuation mechanism within a characteristic environment. A spread can be large because the top and bottom portfolios are far apart in the signal, even if the local return–signal relation is weak. To address that issue, Table 5 reports within-quintile regression slopes using the SDF-weighted signal $Z_{i,t}$.

These slopes are a mechanism check. They measure whether, within a given characteristic environment, moving along the SDF-weighted momentum dimension is systematically rewarded or penalized. In this sense, they distinguish continuation from simple portfolio separation.

The CRSP-only characteristics line up well with standard economic intuition. For turnover, the slope rises across quintiles, indicating that the marginal payoff to the SDF-weighted signal is larger in more actively traded stocks (Hvidkjaer, 2006; Hoberg et al., 2022). For size and price, the slope is negative in the smallest and lowest-price quintiles and positive in larger and higher-price segments, which is consistent with microstructure noise and transient price pressure being most severe among small and low-price stocks (Hong et al., 2000). For `Amihud`, the slope turns negative in the most illiquid quintile, again suggesting that severe illiquidity weakens continuation and can induce correction rather than trend persistence.

The Chen–Zimmerman characteristics sharpen the interpretation further (Chen and Zimmermann, 2022). For `MaxRet`, the slopes are essentially flat, suggesting that the payoff heterogeneity seen in spread-based tables is not coming from large differences in the marginal return–signal relation. For `RealizedVol`, the slope declines across quintiles and the $Q5 - Q1$ difference is negative, which is consistent with the idea that high-volatility environments are less conducive to clean continuation patterns. For `zerotrade1M`, the slope falls steadily and becomes negative in the most illiquid quintile, matching the spread-based evidence.

5.9 Discussion

Several conclusions emerge from this section.

First, the lag-state conditional SDF generates a positive and statistically meaningful momentum payoff when evaluated out of sample. This is the central empirical result of the section.

Second, the implementation variants are informative about the economics of the lag profile. Using stale weights weakens the strategy, which suggests that the pricing relevance of different lags varies over time. The recency-weight ratio captures a distinct dimension of that variation and is systematically related to downside outcomes (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016; Huang, 2022).

Third, the characteristic results show that the SDF-weighted signal does not produce uniform continuation everywhere in the cross-section. Instead, it is strongest in economically intuitive regions and weakens—sometimes turning negative—in environments where illiquidity, microstructure frictions, or reversal-like dynamics are likely to be more important (Hong et al., 2000; Hvidkjaer, 2006; Hoberg et al., 2022). That is exactly what one would expect if the lag-state SDF is capturing which parts of recent return history the market is currently pricing.

Finally, the pricing-error interpretation clarifies how this application relates to the broader message of the paper. The signal can be viewed as a ranking variable based on lagged, cross-sectionally demeaned Euler residuals. Its empirical content is therefore not a direct in-sample pricing test of the SDF. It is an ex-ante application of the estimated pricing kernel to the next return cross-section. This is why the momentum exercise is substantially less sensitive to the finite-

sample pricing-bias problem that complicates direct Euler-equation tests in short samples (Kim and Korajczyk, 2024).

The next section turns to the second empirical application and asks a different question: how does the agnostic return-only SDF compare with structured alternatives such as Kim–Korajczyk in pricing domestic and international returns?

6 Empirical Application II: Agnostic versus KK in Pricing Returns

This section studies the second empirical application of the paper. The question is whether the return-only agnostic SDF developed in Section 3 delivers economically useful pricing information when compared with a structured alternative. The natural benchmark is the Kim–Korajczyk (KK) estimator, which imposes a common factor structure on both returns and the pricing kernel (Kim and Korajczyk, 2024). The comparison is informative for two reasons. First, it provides a direct empirical counterpart to the structured-versus-agnostic comparison emphasized in Sections 2 and 4. Second, it clarifies how the agnostic SDF behaves when it is estimated from a rich U.S. cross section and then applied to domestic and international test assets.

Specifically, we estimate the agnostic SDF from individual U.S. stocks using short rolling windows, and we compare its pricing performance with that of the KK estimator. The domestic evidence evaluates both objects on the Chen–Zimmermann (CZ) portfolio universe (Chen and Zimmermann, 2022). The international evidence applies the same U.S.-estimated pricing kernels out of sample to Jensen–Kelly–Pedersen (JKP) international portfolios (Jensen et al., 2023). This structure aligns closely with the interpretation developed in Section 3. If the agnostic estimator is extracting time-varying discounting information from the U.S. cross section, it should perform well on domestic assets constructed from the same market. Whether that information transfers to international portfolios is an empirical question about both market integration and the finite-sample properties of the pricing-error metric itself (Chen and Knez, 1995; Bekaert and Harvey, 1995; Pukthuanthong and Roll, 2009).

6.1 Construction of the SDFs and evaluation metrics

Let $\hat{m}_t^A(W)$ denote the agnostic SDF estimated at month t using a rolling window of length W, where the baseline implementation uses $W = 12$. Consistent with Sections 3 and 4, the agnostic SDF is estimated from individual CRSP stocks and normalized so that its sample mean is one. Let $\hat{m}_t^{KK}$ denote the KK SDF, estimated from the CZ portfolio universe using a factor-based PCA construction with regularization.

Table 6 summarizes the construction and basic time-series properties of the competing SDFs. The agnostic SDF is estimated over a much shorter rolling window than KK and from a broader cross section of underlying assets. Its time variation is somewhat larger than KK's, as reflected in

the standard deviation of the realized SDF series. That difference will matter later, because the pricing-error variance of any non-trivial SDF depends directly on the variance of $m_t R_{it}$.

For any candidate SDF $\hat{m}_t$, the pricing error on test asset i is

$$\bar{e}_i(\hat{m}) = \frac{1}{T_i^{\text{eval}}} \sum_{t \in \mathcal{T}_i} \hat{m}_t R_{i,t} - 1, \tag{32}$$

where $\mathcal{T}_i$ is the set of months in which both the SDF and the test asset are observed and $T_i^{\text{eval}} := |\mathcal{T}_i|$. The baseline evaluation criteria are the cross-sectional mean squared pricing error,

$$\text{XS-MSE}(\hat{m}) = \frac{1}{N} \sum_{i=1}^{N} \bar{e}_i(\hat{m})^2, \tag{33}$$

and the mean absolute pricing error,

$$\text{Mean}|PE|(\hat{m}) = \frac{1}{N} \sum_{i=1}^{N} \left|\bar{e}_i(\hat{m})\right|. \tag{34}$$

These metrics are simple and transparent, but their interpretation depends on the finite-sample properties of $\bar{e}_i(\hat{m})$, especially when $\hat{m}_t$ is itself time varying and the evaluation history is short or uneven. That issue will matter in interpreting the international comparison below (Kim and Korajczyk, 2024; Freyberger et al., 2025).

6.2 Domestic evidence: pricing the CZ portfolio universe

We begin with domestic test assets. Table 7 reports pricing errors for the 1,624 Chen–Zimmermann portfolios (Chen and Zimmermann, 2022). The main result is unambiguous. The agnostic SDF estimated with a 12-month rolling window outperforms the KK estimator on both metrics. At $W = 12$, the agnostic SDF achieves an XS-MSE of 0.76 basis points and a mean absolute pricing error of 78.8 basis points, compared with 6.54 and 221.5 basis points for KK.

This result is striking for two reasons. First, the agnostic estimator beats KK on KK's own natural construction universe. Second, the result is obtained with a much shorter rolling window and a much broader stock-level construction sample. Thus the domestic evidence suggests that the time variation extracted by the agnostic method is economically informative rather than merely a noisy reweighting of returns (Kozak et al., 2020; Lettau and Pelger, 2020).

Table 7 also reports the agnostic estimator for longer rolling windows. Performance is best at the short end of the window grid, particularly at $W = 12$ and $W = 60$. As the window length increases further, the pricing-error metrics deteriorate. This pattern is exactly consistent with the bias–variance logic developed in Sections 3 and 4. Short windows help the agnostic inversion focus on the most relevant pricing directions in the current cross section, whereas longer windows reduce bias but can introduce more estimation noise and less locally relevant time variation (Kim and Korajczyk, 2024; Kozak et al., 2020).

The domestic evidence therefore supports the main empirical premise of the paper. A return-only SDF estimated from a large cross section of individual stocks over a short rolling window can contain useful time-varying pricing information for domestic test assets. The next question is whether that information transfers out of sample to international portfolios.

6.3 International evidence: pricing JKP portfolios

Table 8 reports the same metrics for the 7,697 Jensen–Kelly–Pedersen international portfolios (Jensen et al., 2023). Here the agnostic and KK SDFs are both constructed from U.S. data and applied out of sample to international assets. The results differ sharply from the domestic case, but the main comparison remains clear.

The agnostic estimator dominates KK by a wide margin. At the baseline window, the agnostic SDF produces an XS-MSE of 2.52 basis points, compared with 17.59 basis points for KK. The same pattern appears for mean absolute pricing error. Thus the return-only method continues to extract more useful pricing information than the structured alternative even when both are transported from the U.S. stock market to international portfolios.

Portfolio-level comparisons tell the same story. On the 1,115 JKP portfolios with at least 60 months of overlapping data for both the agnostic and KK estimators, the agnostic SDF has smaller absolute pricing error than KK for 71.2% of portfolios. This result is important because it shows that the international evidence is not simply a fragile average effect driven by a small set of portfolios. Rather, the agnostic method remains the stronger of the two competing pricing kernels on a broad share of the international cross section.

At the same time, the international evidence is weaker than the domestic evidence in absolute terms. The domestic–international contrast therefore points to two issues at once. One is transferability: pricing information extracted from U.S. stocks does not necessarily map one-for-one into international portfolios. The other is finite-sample noise: pricing errors for time-varying SDFs can be measured imprecisely when evaluation histories are short or heterogeneous. The next subsection develops that point formally.

6.4 Finite-sample interpretation of international pricing errors

The domestic–international contrast should not be interpreted solely as a pass–fail test of market integration. It also reflects the sampling properties of the pricing-error metric itself. To see this, note that for any candidate SDF $\hat{m}_t$ the pricing error

$$\bar{e}_i(\hat{m}) = \frac{1}{T_i^{\text{eval}}} \sum_{t \in \mathcal{T}_i} \hat{m}_t R_{i,t} - 1$$

has sampling variance approximately equal to

$$\operatorname{Var}\big(\bar{e}_i(\hat{m})\big) \approx \frac{\operatorname{Var}(\hat{m}_t R_{i,t})}{T_i^{\text{eval}}}. \tag{35}$$

Thus the cross-sectional MSE can be written heuristically as

$$E\big[\bar{e}_i(\hat{m})^2\big] = \text{bias}_i(\hat{m})^2 + \text{Var}\big(\bar{e}_i(\hat{m})\big). \tag{36}$$

Even if a time-varying SDF contains genuine pricing information, it will be mechanically penalized by the variance term whenever $\text{Var}(\hat{m}_t R_{i,t})$ is large and the evaluation history T_i^{eval} is short (Kim and Korajczyk, 2024).

This logic is especially relevant in the international comparison. The JKP portfolios have much more heterogeneous histories than the domestic CZ portfolios, and the transfer exercise places a U.S.-estimated time-varying pricing kernel on a shifted asset universe. As a result, measured pricing errors can be materially noisier internationally even when the agnostic SDF still dominates the structured benchmark. This is why the paper interprets the domestic–international contrast as evidence not only about transferability across markets, but also about the practical sensitivity of SDF evaluation to finite-sample noise (Chen and Knez, 1995; Bekaert and Harvey, 1995; Pukthuanthong and Roll, 2009; Freyberger et al., 2025).

6.5 Tail behavior of international pricing errors

A natural concern is that the agnostic estimator's international performance relative to KK might be driven by a small number of extreme outliers. Table 9 addresses that concern directly. Panel A reports the share of total cross-sectional MSE accounted for by the worst portfolios in the squared-pricing-error distribution for each method. Panel B then trims the portfolios with the largest agnostic squared pricing errors and recomputes the agnostic and KK XS-MSE statistics.

The purpose of this exercise is to distinguish two possibilities. One is that the international comparison is driven mainly by a handful of unusually poor agnostic realizations. The other is that the agnostic-versus-KK difference reflects a broader contrast in pricing performance across the international cross section. The trimmed-MSE evidence is designed to separate those interpretations.

Taken together, the evidence in Table 9 helps discipline the interpretation of the international results. If the agnostic estimator continues to compare favorably with KK after the most extreme portfolios are removed, then the case for economically meaningful pricing information becomes stronger. More broadly, the exercise reinforces the paper's main point that practical SDF evaluation can be sensitive to the interaction between time-varying discount factors and noisy pricing-error estimates.

6.6 Interpretation: market integration and segmentation

The domestic and international comparisons jointly provide a natural test of market integration. If the U.S. stock market and the international test-asset universe were priced by exactly the same discount factor and if the finite-sample metric were not strongly distorted by estimation noise, then the same U.S.-estimated SDF should perform comparably well in both environments. That is not what the data show.

Domestically, the agnostic SDF clearly beats KK. Internationally, it continues to beat KK, but the overall pricing exercise is weaker in absolute terms and more exposed to sampling variation. The most natural interpretation is that the agnostic SDF estimated from U.S. stocks contains genuine time-varying pricing information for U.S. assets, but that this information does not fully transfer to the international portfolio universe. This pattern is consistent with partial market segmentation rather than full integration (Chen and Knez, 1995; Bekaert and Harvey, 1995; Bekaert et al., 2007; Pukthuanthong and Roll, 2009).

At the same time, the structured alternative does not rescue the international exercise. The agnostic estimator remains the stronger performer in both the domestic and international comparisons. This is important because it shows that the international result is not simply an indictment of all U.S.-estimated discount factors. Instead, it reinforces the comparative advantage of the return-only approach emphasized throughout the paper. When the maintained factor structure is not the right one, KK performs poorly even before broader questions of transferability and sampling noise enter the comparison (Kelly et al., 2019; Kozak et al., 2020; Lettau and Pelger, 2020; Patton and Weller, 2022).

The section therefore delivers two empirical conclusions. First, the return-only agnostic estimator extracts useful time-varying pricing information from the U.S. cross section, as shown by its domestic performance and by its continued dominance over KK internationally. Second, the domestic–international contrast highlights that practical SDF evaluation is highly sensitive to finite-sample noise, particularly when pricing errors are measured over short or heterogeneous return histories. These conclusions are closely aligned with the broader argument of the paper. The agnostic SDF is most useful in settings where the cross section is rich, the relevant time window is short, and the researcher wishes to avoid potentially fragile factor restrictions. The domestic evidence shows that such an estimator can perform very well in practice. The international evidence clarifies both the limits of transferring that pricing kernel across markets and the care required in interpreting pricing-error comparisons when the estimated SDF is itself time varying.

7 Conclusion

This paper studies how to estimate and use stochastic discount factors when the empirically relevant environment is a large cross section of traded assets and a short effective time-series window. In that setting, the main challenge is not the conceptual existence of an SDF. It is the practical recovery of a pricing kernel without relying too heavily on potentially fragile macroeconomic proxies, preference restrictions, or an imposed factor structure. The paper's response is a return-only agnostic estimator that treats the pricing restriction itself as the identifying object.

The first contribution of the paper is theoretical. The paper shows that the agnostic estimator admits a transparent finite-state interpretation: traded payoffs identify state prices, and the SDF follows once those state prices are scaled by probabilities. It then extends the same logic to a conditional lag-state setting in which recent return cross-sections are treated as the relevant state

space at the formation date. In that representation, the conditional SDF becomes a vector of lag weights. This makes the economic content of the estimator much clearer. The agnostic SDF is not an opaque numerical object. It is a return-based pricing rule that ranks recent histories according to their current pricing relevance.

The second contribution is methodological. The paper emphasizes that short-window implementation necessarily creates a tradeoff between bias and estimation error. In finite samples, the agnostic estimator is biased because the pricing surprises are mechanically linked to the same return panel used to estimate the SDF. At the same time, precision depends on the signal-to-noise ratio of the return panel, and regularization improves stability at the cost of additional bias. The simulation framework outlined in Section 4 is designed around the large-N, small-T region in which the estimator is actually used, rather than around generic asymptotics. Its role is to clarify what should and should not be inferred from short-window empirical performance.

The first empirical application uses the conditional lag-state SDF to construct a momentum signal. Instead of imposing a fixed lag structure, the paper lets the lag profile be learned from the pricing kernel at the formation date. The resulting signal is still momentum-like, but its effective horizon varies over time with the estimated pricing relevance of different lags. The evidence indicates that this SDF-weighted momentum strategy generates a positive and statistically meaningful long–short payoff. Its performance weakens when stale lag weights are used, and the shape of the estimated lag profile contains information about downside outcomes. The characteristic-based evidence further suggests that the signal is strongest in economically intuitive regions of the cross section and weakens, sometimes substantially, in environments where illiquidity, microstructure frictions, or reversal-like dynamics are likely to matter more. These findings support the interpretation of the lag-state SDF as a feasible, theory-based proxy for the continuation direction in lag space rather than as a fixed-window momentum rule.

The second empirical application compares the agnostic estimator with the Kim–Korajczyk approach in pricing returns. In domestic data, the agnostic SDF outperforms KK on the Chen–Zimmermann portfolio universe, which suggests that the return-only approach extracts economically meaningful time variation from the U.S. stock cross section. In international data, the agnostic estimator again dominates KK, indicating that useful discount-factor information can still be recovered from returns even when the pricing kernel is transported across markets. At the same time, the domestic–international contrast shows that cross-market pricing comparisons are sensitive to finite-sample noise, particularly when pricing errors are evaluated for time-varying discount factors on heterogeneous test assets. The international evidence is therefore informative both about the limits of transferability and about the practical interpretation of pricing-error metrics in finite samples.

Taken together, the paper's results support a simple conclusion. A return-only SDF can be both economically interpretable and empirically useful in the short-window environments that arise naturally in rolling asset-pricing exercises, conditional strategy design, and out-of-sample pricing comparisons. The main contribution of the paper is therefore not only to propose an estimator,

but also to show how that estimator should be interpreted, where it is most useful, and how its empirical performance should be judged in finite samples. In that sense, the paper argues for a practical middle ground in empirical asset pricing: one can extract meaningful discount-factor variation directly from traded returns, provided that the resulting object is interpreted through the pricing restriction, evaluated with care, and applied in settings where short-window information is economically relevant.

References

Adler, M. and Dumas, B. (1983). International portfolio choice and corporation finance: A synthesis. *Journal of Finance*, 38(3):925–984.

Adrian, T., Crump, R. K., and Moench, E. (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics*, 110(1):110–138.

A"it-Sahalia, Y. and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, 53(2):499–547.

A"it-Sahalia, Y. and Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics*, 94(1–2):9–51.

Alvarez, F. and Jermann, U. J. (2005). Using asset prices to measure the persistence of the marginal utility of wealth. *Econometrica*, 73(6):1977–2016.

Arnott, R. D., Kalesnik, V., and Linnainmaa, J. T. (2023). Factor momentum. *Review of Financial Studies*, 36(8):3034–3070.

Asness, C. S., Moskowitz, T. J., and Pedersen, L. H. (2013). Value and momentum everywhere. *Journal of Finance*, 68(3):929–985.

Barillas, F. and Shanken, J. (2017). Which alpha? *Review of Financial Studies*, 30(4):1316–1338.

Barroso, P. and Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, 116(1):111–120.

Bekaert, G. and Harvey, C. R. (1995). Time-varying world market integration. *Journal of Finance*, 50(2):403–444.

Bekaert, G., Harvey, C. R., and Lundblad, C. (2007). Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies*, 20(6):1783–1831.

Bessembinder, H., Cooper, M. J., and Zhang, F. (2019). Characteristic-based benchmark returns and corporate events. *Review of Financial Studies*, 32(1):75–125.

Blitz, D., Huij, J., and Martens, M. (2011). Residual momentum. *Journal of Empirical Finance*, 18(3):506–521.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296.

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance*, 54(2):655–671.

Bybee, L., Kelly, B., and Su, Y. (2023). Narrative asset pricing: Interpretable systematic risk factors from news text. *Review of Financial Studies*, 36(12):4759–4787.

Campbell, J. Y. (1993). Intertemporal asset pricing without consumption data. *American Economic Review*, 83(3):487–512.

Chan, L. K. C., Jegadeesh, N., and Lakonishok, J. (1996). Momentum strategies. *Journal of Finance*, 51(5):1681–1713.

Chen, A. Y. and Zimmermann, T. (2022). Open source cross-sectional asset pricing. *Critical Finance Review*, 11(2):207–264.

Chen, Z. and Knez, P. J. (1995). Measurement of market integration and arbitrage. *Review of Financial Studies*, 8(2):287–325.

Cochrane, J. H. (2011). Discount rates. *Journal of Finance*, 66(4):1047–1108.

Cochrane, J. H. (2014). A mean-variance benchmark for intertemporal portfolio theory. *Journal of Finance*, 69(1):1–49.

Conrad, J. and Kaul, G. (1998). An anatomy of trading strategies. *Review of Financial Studies*, 11(3):489–519.

Daniel, K. and Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2):221–247.

De Santis, G. and Gérard, B. (1997). International asset pricing and portfolio diversification with time-varying risk. *Journal of Finance*, 52(5):1881–1912.

Delikouras, S. and Dittmar, R. F. (2022). Do investment-based models explain equity returns? evidence from euler equations. *Review of Financial Studies*, 35(8):3823–3866.

Errunza, V. R. and Losq, E. (1985). International asset pricing under mild segmentation: Theory and test. *Journal of Finance*, 40(1):105–124.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fama, E. F. and French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, 128(2):234–252.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Feng, G., Giglio, S., and Xiu, D. (2020). Taming the factor zoo: A test of new factors. *Journal of Finance*, 75(3):1327–1370.

Ferson, W. E. and Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy*, 99(2):385–415.

Ferson, W. E. and Siegel, A. F. (2001). The efficient use of conditioning information in portfolios. *Journal of Finance*, 56(3):967–982.

Freyberger, J., H"oppner, B., Neuhierl, A., and Weber, M. (2025). Missing data in asset pricing panels. *Review of Financial Studies*, 38(3):760–802.

Freyberger, J., Neuhierl, A., and Weber, M. (2020). Dissecting characteristics nonparametrically. *Review of Financial Studies*, 33(5):2326–2377.

Green, J., Hand, J. R. M., and Zhang, X. F. (2017). The characteristics that provide independent information about average U.S. monthly stock returns. *Review of Financial Studies*, 30(12):4389–4436.

Grundy, B. D. and Martin, J. S. (2001). Understanding the nature of the risks and the source of the rewards to momentum investing. *Review of Financial Studies*, 14(1):29–78.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, 33(5):2223–2273.

Gu, S., Kelly, B., and Xiu, D. (2021). Autoencoder asset pricing models. *Journal of Econometrics*, 222(1):429–450.

Hansen, L. P. and Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy*, 99(2):225–262.

Hansen, L. P. and Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, 55(3):587–613.

Hansen, L. P. and Singleton, K. J. (1982). Generalized instrumental variables estimation of non-linear rational expectations models. *Econometrica*, 50(5):1269–1286.

Hansen, L. P. and Singleton, K. J. (1983). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy*, 91(2):249–265.

Hoberg, G., Kumar, N., and Prabhala, N. (2022). Buy-side competition and momentum profits. *Review of Financial Studies*, 35(1):254–298.

Hong, H., Lim, T., and Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, 55(1):265–295.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28(3):650–705.

Huang, S. (2022). The momentum gap and return predictability. *Review of Financial Studies*, 35(7):3303–3336.

Hvidkjaer, S. (2006). A trade-based analysis of momentum. *Review of Financial Studies*, 19(2):457–491.

Jagannathan, R. and Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, 51(1):3–53.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance*, 45(3):881–898.

Jegadeesh, N., Luo, J., Subrahmanyam, A., and Titman, S. (2025). Short-term reversals and longer-term momentum around the world: Theory and evidence. *Review of Financial Studies*, 38(12):3673–3728.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1):65–91.

Jensen, T. I., Kelly, B. T., and Pedersen, L. H. (2023). Is there a replication crisis in finance? *Journal of Finance*, 78(5):2465–2518.

Kelly, B. T., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524.

Kim, S. (2017). Properties of the pukthuanthong-roll stochastic discount factor. Georgia Tech working paper.

Kim, S. and Korajczyk, R. A. (2024). Large sample estimators of the stochastic discount factor. *Journal of Financial Econometrics*, 22(5):1672–1713.

Kozak, S., Nagel, S., and Santosh, S. (2018). Interpreting factor models. *Journal of Finance*, 73(3):1183–1223.

Kozak, S., Nagel, S., and Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics*, 105(1):1–28.

Lettau, M. and Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, 56(3):815–849.

Lettau, M. and Pelger, M. (2020). Factors that fit the time series and cross-section of stock returns. *Review of Financial Studies*, 33(5):2274–2325.

Lo, A. W. and MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3(2):175–205.

Long, J. B. (1990). The numeraire portfolio. *Journal of Financial Economics*, 26(1):29–69.

Lou, D. and Polk, C. (2022). Comomentum: Inferring arbitrage activity from return correlations. *Review of Financial Studies*, 35(7):3272–3302.

Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica*, 46(6):1429–1445.

Martin, I. W. R. (2017). What is the expected return on the market? *Quarterly Journal of Economics*, 132(1):367–433.

Moskowitz, T. J. and Grinblatt, M. (1999). Do industries explain momentum? *Journal of Finance*, 54(4):1249–1290.

Moskowitz, T. J., Ooi, Y. H., and Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, 104(2):228–250.

Patton, A. J. and Weller, B. M. (2022). Risk price variation: The missing half of empirical asset pricing. *Review of Financial Studies*, 35(11):5127–5184.

Pukthuanthong, K. and Roll, R. (2009). Global market integration: An alternative measure and its application. *Journal of Financial Economics*, 94(2):214–232.

Rosenberg, J. V. and Engle, R. F. (2002). Empirical pricing kernels. *Journal of Financial Economics*, 64(3):341–372.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360.

Ross, S. A. (2015). The recovery theorem. *Journal of Finance*, 70(2):615–648.

Solnik, B. H. (1974). An equilibrium model of the international capital market. *Journal of Economic Theory*, 8(4):500–524.

Stulz, R. M. (1981). A model of international asset pricing. *Journal of Financial Economics*, 9(4):383–406.

Table 1: **Performance of the SDF-weighted momentum strategy under alternative implementations.** This table reports monthly performance statistics for three implementations of the momentum strategy constructed from the lag-state stochastic discount factor estimated at the formation date. In each month, the baseline strategy estimates the lag-state SDF from the most recent $K = 12$ cross-sections of stock returns and uses the implied lag weights to form the standardized stock-level signal $Z_{i,t}$. Stocks are then sorted on that signal, and the reported payoff is the next-month return on the top-decile portfolio minus the return on the bottom-decile portfolio. The first row uses the contemporaneous lag-state SDF estimated at the formation month. The second row uses one-month stale lag weights. The third row applies a managed-exposure overlay based on the recency-weight ratio defined in equation (30). The columns report the number of monthly observations (T), mean monthly return, monthly standard deviation, monthly Sharpe ratio, expected shortfall at the 5% level, and the Newey–West t-statistic for the mean using six lags. Returns and volatility measures are in percent per month. Expected shortfall is the average payoff in the worst 5% of months.

Strategy	T	Mean	SD	Sh.	$\mathrm{ES}_{5\%}$	t_{NW6}
Current-window SDF weights $\hat{m}_t$	1163	0.43	4.84	0.31	-12.67	3.06
Past-estimated weights $\hat{m}_{t-1}$	1163	0.37	5.77	0.22	-15.21	2.18
Managed (fixed rule on $\mathrm{RecencyRatio}_t$)	1139	0.33	4.46	0.26	-11.92	2.50

Table 2: **Crash-risk regression using the shape of the estimated lag profile.** This table reports the predictive regression of a crash indicator for the baseline SDF-weighted long–short momentum strategy on the recency-weight ratio RecencyRatio_t. The dependent variable is the indicator $\mathbf{1}\{r_{t+1} < -2\hat{\sigma}_t\}$, where r_{t+1} is the next-month return on the baseline SDF-weighted long–short strategy and $\hat{\sigma}_t$ is a rolling estimate of its volatility. The explanatory variable summarizes whether the positive mass of the estimated lag-state SDF is concentrated in the most recent two lags relative to lags 3–12. The rows report the intercept and slope coefficients, and the columns report the point estimate and the Newey–West t-statistic computed with six lags. Note: A negative slope indicates that periods with a more front-loaded lag profile are associated with a lower probability of a large negative next-month payoff under this crash definition.

Term	Coef.	t (NW6)
α	0.084	4.50
β	-0.252	-3.05

Table 3: **Double-sort evidence for the SDF-weighted momentum signal across firm characteristics.** This table reports characteristic-specific double-sort evidence for the SDF-weighted momentum strategy. Each month, stocks are first sorted into quintiles based on the listed characteristic. Within each characteristic quintile, stocks are then sorted into deciles based on the standardized SDF-weighted signal $Z_{i,t}$ constructed from the lag-state SDF estimated at formation. For each characteristic quintile, the reported entry is the time-series mean of the next-month high-minus-low return, in basis points per month, where the high-minus-low spread is the return on the top decile minus the return on the bottom decile within the same characteristic quintile. Rows correspond to the characteristic used in the first sort. Columns Q1 through Q5 report the high-minus-low spread within each characteristic quintile, and the final column reports the difference between the spread in quintile 5 and the spread in quintile 1. Panel A uses CRSP-only characteristics. Panel B uses Chen–Zimmerman characteristics. Newey–West t-statistics with six lags are shown in parentheses.

Panel A: CRSP-only characteristics (within-quintile sort on $Z_{i,t}$)

Characteristic	Q1	Q2	Q3	Q4	Q5	$HLQ5 - HLQ1$
Turnover	18.3 (1.96)	25.1 (2.34)	28.7 (2.62)	36.0 (2.95)	52.4 (3.98)	34.1 (2.77)
LogMcap	-30.9 (-2.94)	11.8 (1.11)	24.6 (2.21)	41.2 (3.11)	47.2 (3.58)	78.1 (5.70)
LogPrice	-38.3 (-3.86)	34.2 (3.03)	39.7 (3.39)	45.0 (3.53)	41.5 (3.22)	79.8 (5.83)
LogVol	23.6 (2.48)	24.8 (2.28)	18.6 (1.69)	35.5 (2.97)	36.7 (2.91)	13.1 (1.08)
LogDolVol	-4.4 (-0.44)	9.0 (0.86)	28.7 (2.53)	44.5 (3.57)	51.0 (3.84)	55.4 (4.06)
Amihud	48.7 (4.22)	45.1 (3.92)	26.9 (2.30)	6.8 (0.60)	-20.3 (-2.03)	-69.0 (-5.66)
Vol12	12.4 (1.26)	21.1 (1.97)	33.7 (2.98)	38.2 (3.06)	17.0 (1.44)	4.6 (0.36)
Ret12	-47.7 (-4.77)	-24.6 (-1.79)	-10.0 (-0.78)	7.6 (0.71)	28.3 (2.52)	76.0 (5.56)
Ret1	-15.2 (-1.68)	18.3 (1.86)	36.9 (3.37)	55.2 (4.81)	74.8 (5.91)	90.0 (6.45)
LogShrout	12.7 (1.31)	15.6 (1.47)	17.8 (1.61)	31.4 (2.57)	34.2 (2.71)	21.5 (1.73)

Panel B: Chen–Zimmerman characteristics (within-quintile sort on $Z_{i,t}$)

Characteristic	Q1	Q2	Q3	Q4	Q5	$HLQ5 - HLQ1$
Mom12m	-53.5 (-6.88)	-67.0 (-7.16)	-44.1 (-5.19)	-33.8 (-4.46)	-4.2 (-0.65)	49.3 (5.48)
ResidualMomentum	-43.4 (-2.45)	-36.8 (-2.58)	-32.8 (-2.23)	-23.6 (-1.75)	20.2 (1.29)	63.6 (3.18)
MaxRet	15.5 (1.93)	9.5 (1.18)	16.2 (1.93)	17.1 (1.98)	15.8 (1.98)	0.3 (0.03)
RealizedVol	18.9 (2.27)	22.4 (2.75)	24.6 (2.78)	17.1 (1.88)	1.1 (0.12)	-17.8 (-1.93)
zerotrade1M	39.5 (4.31)	25.0 (2.54)	16.8 (1.67)	4.2 (0.39)	-8.4 (-0.91)	-47.9 (-4.37)
InvGrowth	12.0 (0.90)	11.7 (0.94)	14.6 (1.08)	22.9 (1.77)	32.5 (2.30)	20.5 (1.45)
OPLeverage	5.1 (0.43)	19.8 (1.86)	9.2 (0.79)	14.6 (1.35)	21.7 (1.98)	16.6 (1.36)
TotalAccruals	19.4 (1.76)	13.5 (1.18)	7.5 (0.62)	15.9 (1.32)	14.8 (1.21)	-4.6 (-0.42)
RoE	0.6 (0.04)	2.8 (0.18)	4.4 (0.26)	8.9 (0.56)	23.8 (1.54)	23.2 (1.52)
RDS	14.1 (0.86)	1.0 (0.06)	8.1 (0.46)	2.4 (0.14)	3.4 (0.19)	-10.7 (-0.60)

Notes: All returns are in basis points per month.

Table 4: **Double-sort evidence for conventional momentum across firm characteristics.** This table repeats the double-sort design of Table 3, but uses conventional momentum, Mom(3–12), as the within-quintile sort variable instead of the SDF-weighted signal. Traditional momentum is defined as the average return from months $t-11$ to $t-2$, excluding the two most recent months. Within each characteristic quintile, stocks are sorted into deciles on the conventional momentum signal, and the reported entry is the time-series mean of the next-month high-minus-low return, in basis points per month. Rows correspond to the characteristic used in the first sort. Columns Q1 through Q5 report the high-minus-low spread within each characteristic quintile, and the final column reports the difference between the spread in quintile 5 and the spread in quintile 1. Panel A uses CRSP-only characteristics. Panel B uses Chen–Zimmerman characteristics. Newey–West t-statistics with six lags are shown in parentheses. The table is intended to provide a benchmark for the cross-sectional geography of conventional momentum profits relative to the SDF-weighted signal in Table 3.

Panel A: CRSP-only characteristics (within-quintile sort on Mom(3–12)**)**

Characteristic	Q1	Q2	Q3	Q4	Q5	$HLQ5-HLQ1$
Turnover	81.0 (7.66)	82.1 (6.02)	85.6 (5.66)	86.7 (5.60)	105.7 (5.44)	24.7 (1.41)
LogMcap	59.9 (4.58)	78.7 (4.74)	92.2 (5.43)	109.1 (6.09)	79.8 (4.55)	19.8 (1.18)
LogPrice	52.2 (3.26)	115.9 (7.66)	110.4 (7.50)	111.4 (6.57)	93.0 (5.19)	40.8 (2.30)
LogVol	88.2 (7.57)	83.2 (5.75)	68.0 (3.98)	90.2 (5.16)	91.1 (4.61)	2.9 (0.16)
LogDolVol	76.9 (6.21)	72.7 (4.60)	77.5 (4.56)	100.1 (5.82)	104.1 (5.20)	27.2 (1.40)
Amihud	78.3 (4.26)	101.4 (5.67)	79.9 (4.78)	71.2 (4.44)	82.5 (6.46)	4.1 (0.24)
Vol12	43.3 (4.69)	70.2 (5.91)	85.2 (5.65)	113.6 (7.05)	91.9 (4.94)	48.5 (2.68)
Ret12	87.5 (7.19)	121.8 (9.58)	96.9 (7.61)	95.1 (6.86)	73.1 (4.32)	-14.4 (-0.76)
Ret1	7.0 (0.41)	58.6 (3.61)	84.6 (5.56)	105.7 (6.64)	175.6 (9.84)	168.7 (9.99)
LogShrout	83.3 (6.71)	79.4 (4.75)	63.3 (3.70)	89.4 (4.84)	72.3 (4.18)	-11.0 (-0.72)

Panel B: Chen–Zimmerman characteristics (within-quintile sort on Mom(3–12)**)**

Characteristic	Q1	Q2	Q3	Q4	Q5	$HLQ5-HLQ1$
Mom12m	20.3 (1.76)	23.8 (2.01)	20.4 (1.71)	1.9 (0.15)	34.9 (2.19)	14.5 (0.78)
ResidualMomentum	38.9 (2.36)	37.5 (2.58)	45.5 (2.92)	43.3 (2.68)	58.7 (3.30)	19.9 (1.00)
MaxRet	68.1 (5.69)	67.4 (5.03)	75.4 (4.88)	77.7 (4.54)	96.8 (4.91)	28.7 (1.62)
RealizedVol	67.7 (7.47)	90.5 (7.56)	90.4 (6.13)	84.4 (5.01)	80.7 (4.02)	13.0 (0.70)
zerotrade1M	111.6 (5.81)	84.4 (5.18)	86.0 (5.93)	72.4 (5.09)	74.3 (6.55)	-37.3 (-2.05)
InvGrowth	72.4 (3.37)	65.2 (3.78)	87.1 (4.71)	71.6 (3.52)	117.9 (5.28)	45.6 (2.49)
OPLeverage	86.6 (4.73)	107.3 (5.38)	75.7 (4.15)	95.9 (5.20)	82.0 (4.63)	-4.5 (-0.25)
TotalAccruals	97.8 (5.10)	83.9 (5.09)	74.5 (4.49)	101.2 (5.40)	69.5 (3.27)	-28.3 (-1.66)
RoE	58.1 (2.57)	59.3 (2.96)	73.0 (3.90)	82.6 (4.03)	100.3 (4.80)	42.2 (2.29)
RDS	88.1 (3.18)	48.0 (1.88)	80.6 (3.62)	59.8 (2.74)	56.4 (2.31)	-31.8 (-1.83)

Notes: All returns are in basis points per month.

Table 5: **Within-characteristic slope heterogeneity for the SDF-weighted signal.** This table reports within-characteristic heterogeneity in the per-unit return to the SDF-weighted signal. Each month, stocks are first sorted into quintiles based on the listed characteristic. Within each characteristic quintile, we run a cross-sectional regression of next-month stock returns on the standardized SDF-weighted signal $Z_{i,t}$ and an intercept. The reported entry is the time-series mean of the monthly slope coefficient, in basis points per one-standard-deviation increase in the signal. Rows correspond to the characteristic used in the first sort. Columns Q1 through Q5 report the average slope within each characteristic quintile, and the final column reports the difference between the slope in quintile 5 and the slope in quintile 1. Panel A uses CRSP-only characteristics. Panel B uses Chen–Zimmerman characteristics. Newey–West t-statistics with six lags are shown in parentheses. The table is intended to isolate the local strength of the continuation mechanism within each characteristic environment, separate from the magnitude of portfolio separation.

Panel A: CRSP-only characteristics

Characteristic	Q1	Q2	Q3	Q4	Q5	$Q5 - Q1$
Turnover	4.4 (1.06)	6.0 (1.40)	8.6 (1.96)	9.0 (2.28)	15.9 (3.75)	11.5 (2.74)
LogMcap	-7.6 (-2.19)	9.8 (2.49)	15.6 (3.71)	19.6 (4.04)	17.1 (3.11)	24.7 (4.73)
LogPrice	-9.9 (-2.27)	20.5 (5.53)	21.4 (5.44)	22.3 (4.70)	21.5 (4.03)	31.3 (5.71)
LogVol	7.0 (2.03)	6.4 (1.63)	7.6 (1.74)	13.4 (3.13)	13.9 (2.74)	7.0 (1.54)
LogDolVol	0.5 (0.13)	3.3 (0.79)	10.2 (2.36)	17.6 (3.93)	22.4 (4.34)	21.9 (4.30)
Amihud	18.6 (3.65)	20.6 (4.38)	13.7 (3.27)	4.8 (1.19)	-8.4 (-2.39)	-27.0 (-5.79)
Vol12	6.2 (1.26)	7.7 (1.75)	12.1 (2.70)	15.1 (3.57)	6.2 (1.60)	-0.0 (-0.00)
Ret12	-15.4 (-2.68)	-9.0 (-1.03)	-6.8 (-0.79)	6.4 (0.80)	12.9 (2.59)	28.3 (3.86)
Ret1	-4.8 (-1.23)	15.0 (3.40)	22.7 (5.11)	30.5 (7.23)	39.6 (8.81)	44.4 (10.40)
LogShrout	6.5 (2.03)	7.3 (1.80)	6.4 (1.54)	11.8 (2.46)	12.7 (2.30)	6.2 (1.31)

Panel B: Chen–Zimmerman characteristics

Characteristic	Q1	Q2	Q3	Q4	Q5	$Q5 - Q1$
Mom12m	-40.8 (-7.04)	-56.1 (-7.61)	-29.6 (-3.92)	-27.9 (-4.17)	0.9 (0.19)	41.7 (5.66)
ResidualMomentum	-17.1 (-2.78)	-14.9 (-2.50)	-11.2 (-1.87)	-5.4 (-1.04)	6.3 (1.31)	23.4 (3.29)
MaxRet	9.3 (2.21)	7.0 (1.70)	11.7 (2.77)	10.7 (2.47)	10.3 (2.55)	1.0 (0.24)
RealizedVol	12.9 (3.33)	14.9 (3.75)	13.9 (3.27)	10.8 (2.56)	3.7 (0.92)	-9.2 (-2.15)
zerotrade1M	19.3 (4.59)	11.5 (2.54)	8.7 (1.94)	2.5 (0.55)	-2.3 (-0.62)	-21.6 (-5.19)
InvGrowth	6.9 (1.31)	8.1 (1.61)	8.3 (1.48)	11.5 (2.15)	18.0 (3.32)	11.1 (2.65)
OPLeverage	2.9 (0.47)	15.0 (2.86)	7.2 (1.42)	9.5 (1.96)	12.0 (2.67)	9.1 (1.71)
TotalAccruals	12.5 (2.57)	8.0 (1.69)	4.1 (0.80)	11.7 (2.15)	9.5 (1.82)	-3.0 (-0.84)
RoE	-1.3 (-0.24)	3.0 (0.53)	3.7 (0.58)	6.5 (1.05)	16.9 (3.13)	18.2 (4.16)
RDS	10.0 (1.33)	0.1 (0.01)	6.0 (0.94)	1.5 (0.26)	2.6 (0.44)	-7.4 (-1.60)

Notes: Slope coefficients are reported in basis points per one-standard-deviation increase in the SDF-weighted signal.

Table 6: **Construction and time-series properties of the competing stochastic discount factors.** This table summarizes how the agnostic and Kim–Korajczyk (KK) stochastic discount factors are constructed and describes the time-series properties of the realized SDF series used later in the pricing exercises. The agnostic SDF is estimated from individual CRSP stocks using a rolling window of length $W = 12$ months. The KK SDF is estimated from the Chen–Zimmermann (CZ) portfolio universe using a factor-based PCA construction with regularization and a much longer expanding estimation window. Rows describe the construction sample, window length, method, time coverage, number of months available for evaluation, median number of assets or portfolios used in each estimation window, and summary statistics of the realized SDF series. The entries in the last four rows report the mean, standard deviation, minimum, and maximum of each realized SDF. The purpose of the table is to make transparent that the agnostic and KK estimators differ not only in structure but also in the size of the underlying construction universe and the amount of time variation in the resulting discount factor.

	Agnostic	**KK**
Construction assets	Individual CRSP stocks	CZ portfolios ($N \approx 1{,}624$)
Window	Rolling $W = 12$ months	Expanding, $W \geq 120$ months
Method	Vanilla (no ridge)	PCA with regularization
Months covered	1928:11–2024:12	1963:07–2022:12
Number of months	1,086	714
Median N per window	2,524	1,624
$E[m]$	1.0000	1.0000
$\mathrm{sd}(m)$	0.4601	0.3677
$\min(m)$	-0.44	-2.02
$\max(m)$	2.64	2.71

Table 7: **Pricing errors on the Chen–Zimmermann domestic portfolio universe.** This table reports the domestic pricing performance of the agnostic SDF and the Kim–Korajczyk (KK) estimator on the 1,624 Chen–Zimmermann portfolios. For each candidate SDF, the pricing error for portfolio i is the sample average of $\hat{m}_t R_{i,t} - 1$ over the months in which both the SDF and the test portfolio are observed. The first numeric column reports the cross-sectional mean squared pricing error (XS-MSE), measured in basis points. The second numeric column reports the cross-sectional mean absolute pricing error, also in basis points. The first two rows compare the baseline agnostic SDF estimated with a 12-month rolling window to KK. The final three rows show how the pricing performance of the agnostic estimator changes as the rolling window length increases from 12 to 240 months. The table is intended to show whether a return-only SDF estimated from a large cross section of individual U.S. stocks contains economically useful pricing information for a broad domestic portfolio universe.

Method	**XS-MSE (bp)**	**Mean$\lvert PE\rvert$ (bp)**
Agnostic ($W = 12$)	0.76	78.8
KK (PCA, CZ)	6.54	221.5
Agnostic ($W = 60$)	0.95	77.1
Agnostic ($W = 120$)	1.53	92.2
Agnostic ($W = 240$)	2.05	116.1

Table 8: **Pricing errors on the Jensen–Kelly–Pedersen international portfolio universe.** This table reports the international pricing performance of the agnostic SDF and the Kim–Korajczyk (KK) estimator on the 7,697 Jensen–Kelly–Pedersen international portfolios. Both SDFs are estimated from U.S. data and then applied out of sample to international test assets. For each candidate SDF, the pricing error for portfolio i is the sample average of $\hat{m}_t R_{i,t} - 1$ over the overlapping months in which both the SDF and the international portfolio are observed. The first numeric column reports the cross-sectional mean squared pricing error (XS-MSE), in basis points. The second numeric column reports the mean absolute pricing error, also in basis points. The table is designed to show whether discounting information extracted from a broad U.S. stock universe transfers to international test assets, and how the return-only agnostic estimator compares with a structured alternative in that out-of-sample setting.

Method	**XS-MSE (bp)**	**Mean$\lvert PE\rvert$ (bp)**
Agnostic ($W = 12$)	2.52	131.9
KK (PCA, CZ)	17.59	375.8

Table 9: **Tail concentration of international pricing errors and trimmed-MSE evidence.** This table examines whether the international pricing results are driven by a small number of extreme outliers. Panel A reports the share of total cross-sectional MSE accounted for by the portfolios in the top 5% and top 10% of the squared-pricing-error distribution for each method. Columns correspond to the agnostic SDF and the Kim–Korajczyk (KK) estimator. Panel B reports the cross-sectional mean squared pricing error, in basis points, after progressively trimming the portfolios with the largest agnostic squared pricing errors. Rows correspond to the full international sample and then to samples obtained by removing the worst 1%, 5%, 10%, and 20% of portfolios according to the agnostic squared pricing error. Columns report the resulting XS-MSE for the agnostic SDF and KK. The table is intended to show whether the international agnostic-versus-KK comparison is driven primarily by a small number of extreme portfolios or reflects a broader difference in pricing performance.

	Agnostic	**KK**
Panel A: % of total MSE in tails		
Top 5% of portfolios	43.3%	33.2%
Top 10% of portfolios	62.8%	49.1%
Panel B: XS-MSE (bp) after trimming Agnostic's worst portfolios		
Full sample ($N = 1{,}115$)	9.94	16.02
Trim 1% ($N = 1{,}103$)	8.70	16.10
Trim 5% ($N = 1{,}059$)	5.93	16.13
Trim 10% ($N = 1{,}003$)	4.11	16.19
Trim 20% ($N = 892$)	2.52	14.46

Appendix

A Proofs and Additional Derivations

This appendix collects the proofs and supporting derivations omitted from the main text. To keep the paper concise, we do not repeat empirical discussions or interpretation already developed in the body of the paper. Instead, the appendix focuses on four objects: (i) proofs of the propositions used in Section 3, (ii) derivations supporting the momentum application in Section 5, (iii) the ex-ante generated-regressor argument underlying the predictive interpretation of the SDF-weighted signal, and (iv) the bias–variance decomposition used in Section 6.

B Proof of Proposition 1

Write the stacked pricing restriction in the finite-state model as

$$R'\Pi m = \mathbf{1}_N. \tag{37}$$

Define the state-price vector

$$q := \Pi m.$$

Because Π is diagonal with strictly positive entries by Assumption 3, it is invertible, so (37) is equivalent to

$$R'q = \mathbf{1}_N. \tag{38}$$

Existence follows from Assumption 1, which states exactly that there exists at least one q satisfying (38). Uniqueness follows from Assumption 2. Since $\text{rank}(R) = S$, the matrix R has full row rank, so RR' is invertible. Therefore, the unique solution to (38) is

$$q = (RR')^{-1}R\mathbf{1}_N.$$

Mapping back from state prices to the SDF gives

$$m = \Pi^{-1}q = \Pi^{-1}(RR')^{-1}R\mathbf{1}_N,$$

which is equation (9).

If states are equiprobable, $\Pi = \frac{1}{S}I_S$, so $\Pi^{-1} = SI_S$, and therefore

$$m = S(RR')^{-1}R\mathbf{1}_N, \qquad \frac{m}{S} = (RR')^{-1}R\mathbf{1}_N,$$

which is equation (10). This proves Proposition 1.

C Proof of Proposition 2

The lag-state system is the direct analogue of the finite-state problem, with lagged return cross-sections replacing primitive states of nature. The pricing restriction is

$$\left(R_t^{(K)}\right)' \Pi_t^{(K)} m_t^{(K)} = \mathbf{1}_N. \tag{39}$$

Define the lag-state price vector

$$q_t^{(K)} := \Pi_t^{(K)} m_t^{(K)}.$$

Since all lag-state probabilities are strictly positive, $\Pi_t^{(K)}$ is invertible, so (39) is equivalent to

$$\left(R_t^{(K)}\right)' q_t^{(K)} = \mathbf{1}_N. \tag{40}$$

Under the stated rank condition, $R_t^{(K)}\left(R_t^{(K)}\right)'$ is invertible. Hence the unique solution to (40) is

$$q_t^{(K)} = \left(R_t^{(K)}\left(R_t^{(K)}\right)'\right)^{-1} R_t^{(K)} \mathbf{1}_N.$$

Therefore

$$m_t^{(K)} = \left(\Pi_t^{(K)}\right)^{-1} \left(R_t^{(K)}\left(R_t^{(K)}\right)'\right)^{-1} R_t^{(K)} \mathbf{1}_N,$$

which is equation (12). If lag states are equiprobable, $\Pi_t^{(K)} = \frac{1}{K} I_K$, and so

$$\frac{m_t^{(K)}}{K} = \left(R_t^{(K)}\left(R_t^{(K)}\right)'\right)^{-1} R_t^{(K)} \mathbf{1}_N,$$

which is equation (13). This proves Proposition 2.

D Proof of Proposition 3

By definition,

$$y_t = \mu_t \mathbf{1}_N + X_t \gamma_t + e_t.$$

Premultiplying by X_t' and using $X_t' \mathbf{1}_N = 0$ gives

$$X_t' y_t = X_t' X_t \gamma_t + X_t' e_t.$$

Since $X_t' X_t$ is invertible,

$$(X_t' X_t)^{-1} X_t' y_t = \gamma_t + (X_t' X_t)^{-1} X_t' e_t.$$

By definition, the left-hand side is m_t^{hist}, so

$$m_t^{\text{hist}} - \gamma_t = (X_t' X_t)^{-1} X_t' e_t.$$

Taking norms,

$$\|m_t^{\text{hist}} - \gamma_t\| \leq \|(X_t'X_t)^{-1}\| \, \|X_t'e_t\|.$$

If the smallest eigenvalue of $X_t'X_t$ is bounded away from zero, then $\|(X_t'X_t)^{-1}\| = O(1)$. Under condition (15), $\|X_t'e_t\| = o(1)$ as $N \to \infty$. Hence

$$\|m_t^{\text{hist}} - \gamma_t\| = o(1).$$

This proves Proposition 3.

E Derivation of equation (20)

Let

$$p_t := X_t m_t^{(K)}.$$

Because $X_t'\mathbf{1}_N = 0$, the vector p_t is cross-sectionally demeaned, so the OLS slope in the cross-sectional regression

$$R_{i,t+1} = a_{t+1} + b_{t+1} P_{i,t}^{\text{SDF}} + \varepsilon_{i,t+1}$$

is

$$b_{t+1} = \frac{p_t' y_{t+1}}{p_t' p_t} = \frac{\left(m_t^{(K)}\right)' X_t' y_{t+1}}{\left(m_t^{(K)}\right)' X_t' X_t m_t^{(K)}}.$$

Now use the reduced-form decomposition

$$y_{t+1} = \mu_{t+1}\mathbf{1}_N + X_t\gamma_{t+1} + e_{t+1}.$$

Premultiplying by X_t' and using $X_t'\mathbf{1}_N = 0$ gives

$$X_t' y_{t+1} = X_t' X_t \gamma_{t+1} + X_t' e_{t+1}.$$

Substituting into the slope expression yields

$$b_{t+1} = \frac{\left(m_t^{(K)}\right)' X_t' X_t \gamma_{t+1}}{\left(m_t^{(K)}\right)' X_t' X_t m_t^{(K)}} + \frac{\left(m_t^{(K)}\right)' X_t' e_{t+1}}{\left(m_t^{(K)}\right)' X_t' X_t m_t^{(K)}}.$$

Using the notation

$$\langle u, v \rangle_A := u'Av, \qquad \|u\|_A^2 := u'Au,$$

this becomes

$$b_{t+1} = \frac{\langle m_t^{(K)}, \gamma_{t+1} \rangle_{X_t'X_t}}{\|m_t^{(K)}\|_{X_t'X_t}^2} + \frac{\left(m_t^{(K)}\right)' X_t' e_{t+1}}{\left(m_t^{(K)}\right)' X_t' X_t m_t^{(K)}}.$$

If $\|X_t' e_{t+1}\| = o(1)$ and the smallest eigenvalue of $X_t' X_t$ is bounded away from zero, then the remainder term is $o(1)$. Using Proposition 3, one can also substitute

$$m_{t+1}^{\text{hist}} = \gamma_{t+1} + o(1)$$

to obtain the representation in equation (20).

Finally, because $p_t = X_t m_t^{(K)}$ and $p_t' \mathbf{1}_N = 0$, the portfolio with weights

$$w_t = \frac{p_t}{p_t' p_t}$$

is self-financing, and its realized return satisfies

$$w_t' y_{t+1} = \frac{p_t' y_{t+1}}{p_t' p_t} = b_{t+1}.$$

Thus the slope is exactly the return on the factor-mimicking portfolio for the SDF-weighted signal.

F Pricing-error representation of the SDF-weighted signal

Define the demeaned pricing error at date t for asset i as

$$\widetilde{\text{PE}}_{i,t} := m_t \widetilde{R}_{i,t}.$$

In the lag-state application, the SDF-weighted signal is

$$P_{i,t}^{\text{SDF}} = \sum_{k=1}^{K} m_{t,k}^{(K)} \widetilde{R}_{i,t+1-k}.$$

By definition of the demeaned pricing error at each lag,

$$m_{t,k}^{(K)} \widetilde{R}_{i,t+1-k} = \widetilde{\text{PE}}_{i,t+1-k}.$$

Summing over lags gives

$$P_{i,t}^{\text{SDF}} = \sum_{k=1}^{K} \widetilde{\text{PE}}_{i,t+1-k},$$

which is equation (29) in the main text.

The same notation yields a useful expression for the cross-sectional regression slope. Let

$$p_t := X_t m_t^{(K)}.$$

Then the OLS slope in the regression of next-month returns on the SDF-weighted signal is

$$b_{t+1} = \frac{p_t' y_{t+1}}{p_t' p_t}.$$

Using the lag decomposition

$$p_t = \sum_{k=1}^{K} m_{t,k}^{(K)} \widetilde{R}_{\cdot,t+1-k}$$

and writing $\widetilde{\mathrm{PE}}_{\cdot,t+1-k} := m_{t,k}^{(K)} \widetilde{R}_{\cdot,t+1-k}$, bilinearity of the inner product yields

$$b_{t+1} = \frac{\sum_{k=1}^{K} \langle \widetilde{\mathrm{PE}}_{\cdot,t+1-k}, y_{t+1} \rangle}{\sum_{k=1}^{K} \sum_{\ell=1}^{K} \langle \widetilde{\mathrm{PE}}_{\cdot,t+1-k}, \widetilde{\mathrm{PE}}_{\cdot,t+1-\ell} \rangle}.$$

This is the pricing-error form underlying the discussion in Section 5. It shows that the momentum payoff is high when past demeaned pricing errors are aligned with the next return cross-section and when that alignment is concentrated in a stable cross-sectional direction.

G Why ex-ante SDF-based prediction avoids the finite-T critique

This subsection formalizes the idea used in Section 5 that the predictive use of the lag-state SDF is less exposed to the finite-sample critique that applies to in-sample pricing tests.

Let $(\mathcal{F}_t)_{t \geq 0}$ be the filtration generated by the data. For each asset i at date $t+1$, let $R_{i,t+1}$ denote the realized return and let $x_{i,t}$ denote the vector of lagged variables known at time t. Let the ideal history statistic be

$$P_{i,t}^{*} := x_{i,t}' m_t^{(K)},$$

and let its feasible estimated counterpart be

$$\widehat{P}_{i,t} := x_{i,t}' \widehat{m}_t^{(K)}.$$

By construction, $\widehat{P}_{i,t}$ is $\mathcal{F}_t$-measurable.

Assume the one-step-ahead innovation

$$u_{i,t+1} := R_{i,t+1} - E[R_{i,t+1} \mid \mathcal{F}_t]$$

satisfies

$$E[u_{i,t+1} \mid \mathcal{F}_t] = 0.$$

Lemma A.1. For any $\mathcal{F}_t$-measurable instrument $Z_{i,t}$,

$$E[Z_{i,t} \, (R_{i,t+1} - E[R_{i,t+1} \mid \mathcal{F}_t])] = 0.$$

Proof. By iterated expectations,

$$E[Z_{i,t}\,(R_{i,t+1} - E[R_{i,t+1} \mid \mathcal{F}_t])] = E[E[Z_{i,t}\,(R_{i,t+1} - E[R_{i,t+1} \mid \mathcal{F}_t])|\mathcal{F}_t]]\,.$$

Since $Z_{i,t}$ is $\mathcal{F}_t$-measurable,

$$= E[Z_{i,t}\,E[R_{i,t+1} - E[R_{i,t+1} \mid \mathcal{F}_t]|\mathcal{F}_t]] = 0.$$

□

Proposition A.1. Consider the date-$t+1$ cross-sectional regression

$$R_{i,t+1} = a_{t+1} + \beta_{t+1}\widehat{P}_{i,t} + \varepsilon_{i,t+1}.$$

If $\widehat{P}_{i,t}$ is $\mathcal{F}_t$-measurable, then

$$E[\widehat{P}_{i,t}\varepsilon_{i,t+1}] = 0$$

under the standard martingale-difference restriction on innovations.

Proof. Write

$$R_{i,t+1} = E[R_{i,t+1} \mid \mathcal{F}_t] + u_{i,t+1}.$$

The regression coefficient is the linear projection coefficient of $R_{i,t+1}$ on $\widehat{P}_{i,t}$ and an intercept, so the regression residual is orthogonal to the regressor if $\widehat{P}_{i,t}$ is predetermined with respect to $u_{i,t+1}$. But this follows directly from Lemma A.1 with $Z_{i,t} = \widehat{P}_{i,t}$. Therefore the predictive regression uses ex-ante variation in the estimated history statistic rather than reusing the same realization that generated the SDF estimate. □

The important point is that generated-regressor noise may still attenuate the predictive coefficient, but this is a measurement-error problem rather than the in-sample finite-T pricing-bias problem emphasized in the main text.

H Bias–variance decomposition for pricing errors

For any candidate SDF $\hat{m}_t$, define the average pricing error on test asset i as

$$\bar{e}_i(\hat{m}) = \frac{1}{T_i^{\text{eval}}} \sum_{t \in \mathcal{T}_i} \hat{m}_t R_{i,t} - 1.$$

Then

$$E[\bar{e}_i(\hat{m})^2] = (E[\bar{e}_i(\hat{m})])^2 + \text{Var}(\bar{e}_i(\hat{m})).$$

Writing

$$\text{bias}_i(\hat{m}) := E[\bar{e}_i(\hat{m})],$$

we obtain

$$E[\bar{e}_i(\hat{m})^2] = \text{bias}_i(\hat{m})^2 + \text{Var}(\bar{e}_i(\hat{m})).$$

Under weak dependence and a standard long-run variance approximation,

$$\text{Var}(\bar{e}_i(\hat{m})) \approx \frac{\text{Var}(\hat{m}_t R_{i,t})}{T_i^{\text{eval}}}.$$

Hence any comparison based on cross-sectional MSE must reflect both systematic pricing mistakes and the sampling variation of discounted returns. This is why a benchmark with $\hat{m}_t \equiv 1$ can appear attractive in short evaluation windows even if it contains much less economically meaningful time variation than a nontrivial SDF.

I Additional implementation details

Two implementation points are worth recording explicitly.

First, in the momentum application, the signal is standardized cross-sectionally each month:

$$Z_{i,t} = \frac{P_{i,t}^{\text{SDF}} - \overline{P}_t^{\text{SDF}}}{\text{sd}_t(P_{i,t}^{\text{SDF}})}.$$

This standardization does not change the cross-sectional ranking implied by the SDF-weighted history statistic. It simply makes the portfolio sorts and regression coefficients comparable across months.

Second, in the pricing application, portfolio i contributes to the evaluation of a candidate SDF only over the months in which both the estimated SDF and the test-asset return are observed. That is why the effective evaluation horizon

$$T_i^{\text{eval}} := |\mathcal{T}_i|$$

can vary substantially across portfolios, especially in the international sample. This variation is one reason why the finite-sample noise of pricing-error metrics is first-order in that application.