Breaks and Trends in Factor Premia*

Liyuan Cui Guanhao Feng Jianxin Ma Yinan Su

First version: Jun. 2025; This version: Jan. 2026

Abstract

We estimate a cross-sectional asset pricing model with structural breaks and regime-dependent trends in factor premia. The method incorporates total-variation (TV) regularization into a characteristic-based factor model, nesting the panel predictive regression and the Fama-MacBeth procedure as extreme cases. It produces piecewise linear trend lines of varying frequencies, capturing either long-term or fine-grained factor momentum trends depending on the TV penalty, which is tuned in a machine learning framework. Large-sample theoretical properties of the estimators include consistency and asymptotic normality. Out-of-sample trend-following and factor-timing strategies yield superior performance compared to buy-and-hold or simple rolling-window momentum strategies. Considering breaks and trends improves the SDF construction compared to static counterparts.

Keywords: factor models, machine learning, return predictability; investments, factor premia; momentum trends; structural breaks; time-varying sparsity.

JEL Codes: C1, G11, G12

*We thank Raymond Kan, Bryan Kelly, Jia Li, Yan Liu, Stefan Nagel, Markus Pelger, Nikolai Roussanov, Allan Timmermann the seminar and conference participants at the Global AI Finance Research Conference, Hong Kong Conference for Fintech, AI, and Big Data in Business, HKUST IAS-SBM Joint Workshop on Financial Econometrics, Tsinghua Machine Learning Application in Financial Economics Conference, and FMA Asia/Pacific Conference for their valuable comments. Cui (E-mail: liyuan.cui@cityu.edu.hk) and Feng (E-mail: gavin.feng@cityu.edu.hk) are at the City University of Hong Kong; Ma (E-mail: jianxin.ma@warwick.ac.uk) is at the University of Warwick; Su (E-mail: ys@jhu.edu) is at Johns Hopkins University.

1 Introduction

Risk-return modeling of the stock returns mainly relies on factor models. While significant progress has been made in using firm characteristics to estimate dynamic factor loadings, estimating dynamic factor premia remains challenging due to limitations in time-series data.[1] Estimating dynamic factor premia is inherently challenging, as underlying shifts are often obscured by the high-frequency volatility of realized factor returns. However, identifying structural breaks and trends in factor premia is crucial for constructing the Stochastic Discount Factor (SDF) and for factor-based investments. New advances in this area are critically needed.[2]

From an economic perspective, there are many reasons why factor premia might change, and in particular, change in a break-and-trend fashion. Prior studies show that factor premia can decay or vanish after the publication of a factor's construction, can be influenced by factor-level trading flows, and may undergo abrupt changes in response to structural shifts in the underlying economic conditions, such as the Global Financial Crisis and the COVID-19 pandemic.[3] The variety of underlying reasons suggests that breaks and trends can emerge in complex ways, and that modeling these economic mechanisms individually may be insufficient to capture the full picture. The full picture is likely a blend of short- and long-term (or high- and low-frequency) variations. Therefore, a data-driven approach can be well-suited to tease out the breaks and trends from the stock return data.[4] We propose a new method that accomplishes these goals and demonstrate its general usefulness in both asset pricing and investment applications.

[1] Fama and French (1993) link expected returns to factor betas, and Daniel and Titman (1997) examine firm characteristics. Recent latent factor models that integrate firm characteristics include Kelly et al. (2019), Chen et al. (2025), Avramov et al. (2025).

[2] Dangl and Halling (2012), Smith and Timmermann (2022), Patton and Weller (2022), and Chib et al. (2023), among others, contribute to this area. Methodological differences are discussed below.

[3] For example, Pástor and Stambaugh (2001) identify structural breaks in the equity premium; Daniel and Moskowitz (2016) document momentum crashes during market rebounds; McLean and Pontiff (2016) show anomaly premia decline post-publication.

[4] As an alternative to the "data-driven" approach, one could also study time-varying risk premium by modeling its association with observable economic variables, for example An et al. (2025) associate factor premia with trading-flow-driven quantity fluctuation; Gagliardini et al. (2016) model risk premia as a linear function of lagged instruments such as the default spread.

Our method is called Total-Variation Panel Regression (TV-PR). It is designed with two components that facilitate the identification of dynamic factor premia and distinguish it from other approaches for the same purpose. First, we leverage the vast amount of data in the *panel* of stock returns and characteristics to improve the estimation of factor premia, rather than relying solely on time-series data on factor realizations. We rely on the aforementioned factor literature linking firm characteristics to factor loadings to focus on estimating factor premia. Then, estimation focuses on factor premia (γ_t), which, when multiplied by the loadings (characteristics), predict individual stock return ($r_{i,t+1}$).[5] As a result, the estimation boils down to a return *predictive panel regression*, where time-varying predictive coefficients are the time-varying factor premia. TV-PR extends the Fama-MacBeth method (Fama and MacBeth, 1973), which uses cross-sectional regressions to estimate static factor premia, by replacing arbitrary rolling-window averages with data-driven, piecewise-constant segments.

Second, we employ a machine learning technique known as total variation (TV) regularization to capture trends and breaks in factor premia.[6] In a nutshell, it imposes regularization on the *variation* between each pair of factor premia coefficients adjacent in time, $|\gamma_t - \gamma_{t+1}|$, also known as fused Lasso. Thereby, it encourages piecewise-constant estimates of factor premia over time. Specifically, when break signals are weak, the regularization "fuses" a sequence of γ's together, capturing a constant trend. Otherwise, it allows for abrupt changes in the γ sequence to capture breaks.

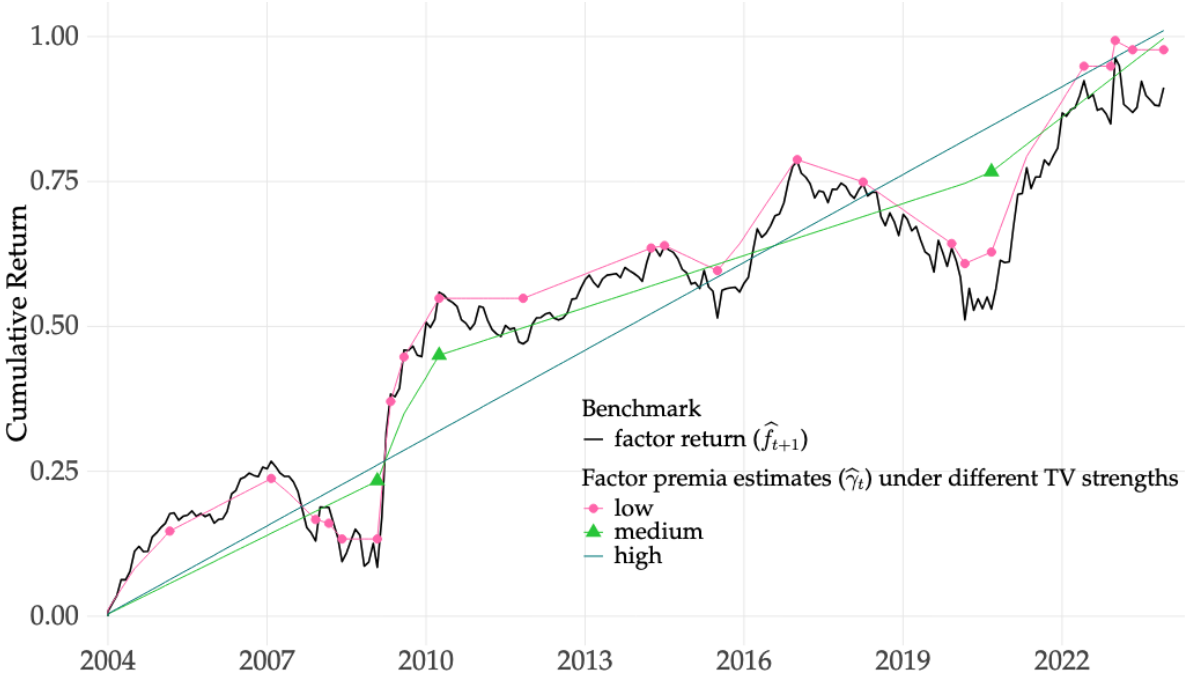
Figure 1 illustrates TV-PR's estimation of breaks and trends by plotting the cumulative sums of factor returns and factor premia estimates (where cumulative sums of piecewise constant γ estimates appear as piecewise linear segments). TV-PR effectively fits linear trend lines to match the overall direction of the factor performance.

[5] We adopt a cross-sectional "slope factor" framework, where observable firm characteristics serve as factor loadings, and the resulting cross-sectional regression coefficients are interpreted as the realized returns to characteristic-managed portfolios. This setup relates to the MSCI Barra model. For simplicity, we write γ_t in the Introduction instead of $\gamma_{p,t}$, where p indexes factors. The main model is a multi-factor framework, detailed in Section 2.

[6] TV regularization is a widely used method in statistics and machine learning. Early contributions include Rudin et al. (1992) and Tibshirani et al. (2005). More recent developments include Li, Qian, and Su (2016), which introduces an adaptive fused-LASSO method to detect multiple structural breaks in panel data with interactive fixed effects. In asset pricing, Tang et al. (2022) apply a TV penalty to the cross-section in each period.

Figure 1: Breaks and trends in factor premia, different TV strengths/smoothness levels

Note: The cumulative returns are plotted. Black: slope factor return ($\hat{f}_{t+1}$); Red/green/blue: factor premium estimates ($\widehat{\gamma}_t$) under low/medium/high TV regularization strength, respectively. In each constant γ segment, the cumulative sum is linear, indicating a trend. Higher TV regularization leads to fewer detected breaks and smoother estimated trends, and vice versa. This illustration shows the value factor (BE/ME); see the other factors and multi-factor results below. A detailed construction of every curve is presented in Section 4.



Drawing trend lines on factor performance charts is an intuitive idea, but is often considered too simplistic because of the arbitrariness of choosing the change points. TV-PR formalizes this intuitive idea by selecting breaks with total-variation regularization in stock-return predictive regressions.

TV-PR allows for the flexibility to capture breaks of varying frequencies, or, in other words, trends of varying scales. These choices are controlled by the TV regularization parameter. High regularization strength smooths trends (see the blue line in Figure 1), and in the extreme, TV-PR simplifies to fixed-coefficient panel predictive regressions (e.g., Gu et al., 2020). On the other hand, with extremely low regularization, the model reduces to period-by-period cross-sectional regressions, and the γ_t estimate approaches the slope factor realization itself (the red jagged trend line approaching the black factor return curve). Thus, the TV regularization parameter opens up a spec-

trum of breaks and trends of varying frequencies/smoothness. Neither extreme, both of which are already studied in the literature and nested within our model, yields optimal performance across various investment and asset pricing applications, as detailed further below. Instead, we find "sweet spots" that lie between these two extremes, where the model admits just the right amount of breaks and trends to best suit the specific applications.[7] Such "sweet spots" are identified by tuning the TV regularization parameter using standard machine learning techniques.

TV-PR captures trends of varying smoothness adaptively, rather than using prespecified trend lengths. As illustrated in Figure 1, the TV regularization parameter controls the overall scale of the trends, rather than directly fixing a uniform trend segment length. The lengths of the trend segments are adaptive to the data, subject to break estimation. This adaptability enables the detection of structural breaks more promptly, rather than relying on a fixed look-back window that may include data from a different regime. Fixed look-back windows are adopted in momentum trading strategies and in rolling-window re-estimates of the same model.[8] Empirical results reported further below support TV-PR's superior performance.

As such, TV-PR incorporates breaks and trends into an otherwise canonical framework, making it compatible with many existing empirical asset pricing techniques and applications. An important aspect of our analysis concerns the selection of factors. It is known that many characteristics in the "factor zoo" are redundant, and that using a LASSO-type thresholding on the factor premia can effectively improve predictive and pricing performance (e.g., Feng et al., 2020). We take the message of this literature and integrate factor selection into TV-PR, by not only encouraging trend ($\gamma_t = \gamma_{t+1}$) but also sparsity ($\gamma_t = 0$) in selected factors. As a result, TV-PR achieves *time-varying sparsity* in factor selection, allowing the set of active factors to change over time. Thereby, we construct an SDF whose factor composition is not only selected but also varies over

[7] Alternatively, Bandi and Tamoni (2022) and Bandi and Su (2025) rely on spectral methods to decompose returns into different frequency components, and show the intermediate components of business cycle lengths have particular relevance in asset pricing.

[8] Zhu and Zhou (2009) link moving-average indicators to investors' priors. Neely et al. (2014) use trend-type technical indicators to forecast the equity premium. Fama and French (2020) incorporate dynamic characteristics but estimate the model on a period-by-period basis.

time, reflecting changing market conditions and evolving investor preferences.

TV-PR estimates the set of varying risk premia in one step by solving an optimization problem. The problem can be efficiently solved using standard numerical algorithms for LASSO-like optimization. We employ an Alternating Direction Method of Multipliers (ADMM) algorithm, which yields an estimate in approximately five minutes for the largest model we consider (with over 60 characteristics/factors and around 1 million stock-month observations).

We establish the theoretical properties of the TV-PR estimator. First, the estimator is consistent, ensuring that the estimated factor premia converge to their true regime-dependent values as data accumulate. Second, it is asymptotically normal, permitting valid inference on time-varying risk premia across regimes. Finally, it satisfies the oracle property of break detection, accurately identifying both the relevant predictors and the locations of structural breaks without prior knowledge. In finite sample simulations, TV-PR averages out factor innovations within each regime, resulting in lower estimation errors and higher predictive R^2 values than a range of time-invariant and time-varying benchmark methods.

Empirical findings. We find that breaks and trends are pervasive features of factor premia, and that accounting for these features using our method improves asset pricing and investment performance, as well as the economic interpretability of factor models. This overall finding is consistent across different model configurations – using factors one at a time or jointly in a multi-factor model – and applies to both firm-characteristic and past-return factor constructions.

Specifically, we perform two groups of empirical analyses. The first uses 61 firm characteristics to form factors, either individually or jointly in a multi-factor setting. The second narrows the focus to past returns across different horizons as characteristics to examine the time-varying premia of momentum, reversal, and seasonality factors. This empirical setting follows Marrow and Nagel (2024), who approach this issue using Gaussian process regressions.

We begin by showing that TV-PR estimates breaks and trends in a coherent and

economically meaningful way, suggesting that breaks and trends are pervasive and important features of factor premia. We find that the estimated time-varying factor premia tend to exhibit distinct regime shifts around major market events such as the Global Financial Crisis and the COVID-19 pandemic. The patterns in the estimated time-varying premia align with the intuitive trend lines in Figure 1 and with visualizations of the other factors. The identified break points are stable across different rolling estimation windows (provided that the window has adequate margins before and after the break point).[9] For past return-based factors, TV-PR identifies rich decay and recovery patterns in momentum, reversal, and seasonality factors that are consistent with documented findings such as post-publication decay. It further supports the view that return-based anomalies are not permanent features of the cross-section of expected stock returns (Marrow and Nagel, 2024).

While the estimates conform to the intuitive notion of trends and breaks in factor premia, the next set of results confirms that the trends and breaks identified by TV-PR reflect genuine shifts in the underlying factor premia process, rather than ex post fitting of factor realization shocks that happen to align over a continuous period. The distinction is that true factor premia changes are persistent and contain predictive power. The following investment and asset pricing exercises investigate these persistence and predictability properties, which establish the presence of breaks and trends in factor premia and their general usefulness.

In the investment exercises, we devise an adaptive factor-timing strategy that uses time-varying factor premia estimates as trading signals. The rationale is that the most recent factor premia trend should persist and continue into future periods, thereby serving as a timing signal for next month's factor returns. Granted, TV-PR might require several months of observation since the last break to form a reliable estimate of the newest trend, and "real-time" break detection is always challenging. Nonetheless, as long as the underlying factor premia trend persists after detection, the timing

[9]The TV-PR estimates do not require long margins from the end of the estimation window to detect a new break point, suggesting the potential for "real-time" detection of breaks (Smith and Timmermann, 2021). This observation also lays the foundation for using trend estimates as factor-timing signals, which is formally detailed and further explained below.

strategy can still be effective. Note that the TV regularization parameter allows us to explore a spectrum of trend scales ranging from historical averages to the one-month factor momentum signal. Both extremes are degenerate cases of TV-PR and are well understood. The optimal regularization strength lies between these extremes, identifying an intermediate scale that yields superior trend-following signals.

We find a hump-shaped relationship between investment performance and TV regularization, indicating a "sweet spot" of trend scales for which the factor-timing strategy is most effective. The peak of the hump shape is stable and dominates the performance of static moving-average-style factor momentum strategies across various window lengths. This result shows that factor premia exhibit persistent trends and that the break-and-trend model is particularly suitable for the data. It also shows that properly accounting for this structure with TV-PR has significant economic value compared to ad hoc constructions such as momentum.

The specific strategy design involves monthly rolling-window re-estimates of TV-PR, where the last (end-of-window) factor premia are used as the trading signal for the next month. With this setup, we perform standard portfolio formation, including single-factor timing strategies and multi-factor long–short strategies, and conduct standard backtesting and out-of-sample (OOS) evaluations of investment performance, such as Sharpe ratios. The improvement in investment performance is consistent across different configurations.

Finally, we evaluate the asset pricing properties of the estimated model, which incorporates breaks and trends in factor premia. The two main asset pricing findings are, first, that TV-PR can forecast individual stock returns at the monthly frequency. Statistically, TV-PR can be interpreted as a time-varying coefficient panel predictive regression model, and this finding suggests that the predictive regression framework is working as intended. From an asset pricing perspective, this suggests that factor premia indeed vary with breaks and trends, and that multiplying the premia by the observed factor loadings (or characteristics) can explain expected stock returns across the cross section.

Second, TV-PR enables time-varying factor selection, resulting in an adaptive SDF construction whose factor composition evolves over time. That means not only the factor premia change, but also the composition of the priced factors changes across regimes, reflecting shifts in the sources of systematic risk. We show that the within-regime SDF composition exhibits higher Sharpe ratios compared to a range of classical factor models that remain constant over time (Barillas et al., 2020). It suggests the relevant drivers of cross-sectional returns tend to emerge and fade with economic cycles. Factor models of fixed factor composition are likely mis-specified, and allowing for time-varying factor selection can lead to a more accurate representation of the SDF.

Related literature. Our work contributes to several strands of the literature. First, Smith and Timmermann (2021, 2022) also study break detection methods in factor models. A key methodological distinction we rely on is the machine learning technique of TV regularization, whereas they take a Bayesian approach. The TV specification enables flexible modeling that is closely related to existing methods, such as panel predictive regressions, the Fama-MacBeth procedure, and dynamic factor selection. Moreover, our specification allows for heterogeneous breaks across factors, thereby accommodating distinct structural dynamics that may be masked under a homogeneous specification. Existing time-varying models often impose the restrictive assumption of simultaneous structural shifts across all predictors. This assumption might overlook the possibility of distinct structural breaks in individual predictors, as noted by Gagliardini et al. (2005), Li et al. (2016), and Smith et al. (2019), thereby failing to capture the heterogeneous dynamics observed in practice. Our approach is more data-driven and less structured, allowing for breaks and trends without pre-specifying their functional forms or observable drivers.

This paper also speaks to the literature that addresses the "factor zoo" with selection (e.g., Harvey et al., 2016; Feng et al., 2020). However, most studies assume time-invariant predictive structures with global sparsity and do not focus on variation in predictor selection over time.[10] TV-PR incorporates time-varying sparsity, permitting

[10] Chinco et al. (2019) demonstrate that LASSO identifies sparse, transient predictors that enhance

predictor importance to evolve over time. We indeed find that time-varying selection improves the construction of SDFs.

This paper contributes to the literature on machine learning predictions (e.g., Gu et al., 2020; Kelly et al., 2019; Bryzgalova et al., 2020), but we present it within a time-varying predictive coefficient framework. Empirically, the predictive relationships are unstable, as Welch and Goyal (2008) highlights that the out-of-sample performance of most equity premium predictors is poor and unstable.[11] We introduce a novel dual-penalization strategy to address prediction instability.

Lastly, this paper contributes to the factor-investment literature by proposing a new factor-timing strategy derived from our estimates of time-varying risk premia. The break-and-trend strategy is closely related to slope factors[12] Fama and French (2020) show that these cross-sectional factors explain average stock returns better than traditional time-series factors. See also Chib et al. (2023) for evidence that slope factors outperform conventional sorted factors, and Kozak and Nagel (2023) for covariance conditions under which they span the SDF and factor momentum.[13]

This paper is structured as follows. We develop the model and derive its statistical properties in Sections 2 and 3. Section 4 reports the estimation results. The subsequent sections examine the method's empirical value: Section 5 analyzes investment performance, and Section 6 discusses SDF composition. Section 7 concludes. Technical proofs and implementation details are provided in the appendix.

2 Model and TV-PR Estimation

We specify a standard factor model for stock returns, with the only emphasis that the factor premia vary with breaks and trends. Let $r_{i,t}$ denote the excess return of

stock-return forecasts, while Freyberger et al. (2020) employ nonparametric adaptive group LASSO to reveal that few firm characteristics predict returns when accounting for nonlinearities.

[11] A key reason for the mixed success of these factors is that their predictive power appears to be time-varying, often concentrated in specific economic regimes or "pockets of predictability" (e.g., Rapach et al., 2010; Farmer et al., 2023; Cong et al., 2024).

[12] Since Fama (1976), the regression coefficients have been used as tradable "slope" factors.

[13] Kelly et al. (2021) show that momentum and reversal arise from time-varying risk exposures captured in conditional factor models; Ehsani and Linnainmaa (2022) demonstrate that momentum in individual stock returns relates to momentum in factor returns; and Arnott et al. (2023) find that factor momentum drives industry and stock momentum.

stock i in month t, and $z_{i,t}$ be a $p \times 1$ vector of observed stock characteristics. The factor model is

$$r_{i,t+1} = z'_{i,t} f_{t+1} + u_{i,t+1} \quad \text{with} \quad f_{t+1} = \gamma_t + \nu_{t+1}. \tag{1}$$

Importantly, the factor premia process $\gamma_t \equiv E_t[f_{t+1}]$ is time-varying with trends and breaks. Traditional approaches model time-varying factor premia as functions of a pre-specified set of conditioning variables or as a deterministic time process. In contrast, we treat $\{\gamma_t\}_{t=1}^T$ as a high-dimensional object—effectively pT parameters—and estimate the entire sequence of factor-premium realizations using a regularized framework. In particular, it follows a piecewise-constant structure: for the jth factor

$$\gamma_{1,j}, \gamma_{2,j}, \gamma_{3,j}, \cdots = \underbrace{\theta_1^j, \theta_1^j, \ldots, \theta_1^j}_{\text{the first regime with duration } |G_1^j|}, \underbrace{\theta_2^j, \theta_2^j, \ldots, \theta_2^j}_{\text{the second regime with } |G_2^j|}, \underbrace{\theta_3^j, \theta_3^j, \ldots, \theta_3^j}_{\text{the third regime with } |G_3^j|}, \ldots, \tag{2}$$

where θ_n^j and $|G_n^j|$represent the nth constant-trend segment of factor j's premium and duration, respectively. These segments are separated by breaks at unknown locations, which do not need to be aligned across factors. If the jth predictor enters Eq. (1) with a homogeneous effect, its coefficient is constant over time at a common value θ_1^j, i.e., $\gamma_{t,j} = \theta_1^j$ for all $1 \leq t \leq T$, and the associated segment spans the full sample, $|G_1^j| = T$. More generally, the jth predictor may display multiple structural breaks. In this case, $\gamma_{t,j} = \theta_n^j$ within regime n, where θ_n^j denotes the regime-specific coefficient for predictor j. Some of the regimes may have a zero premium ($\theta_n^j = 0$), indicating the time-varying sparsity in priced factors.

Figure 2a shows this break-and-trend structure in a 10-period, three-factor illustration. Section 3 details the assumptions of this data-generating process for deriving the estimator's statistical properties. We also note that the factor model Eq. (1) takes a simplified view of factor loadings, which are directly specified as characteristics ($z_{i,t}$). The purpose is to focus our attention on factor premia. From this perspective, model Eq. (1) is a cross-sectional factor model with time-varying premia.[14]

[14]Slope factors, also known as cross-sectional factors, are factors constructed as the coefficients of period-by-period return-on-characteristics regressions (Fama, 1976; Chib et al., 2023; Fama and French, 2020). They can be interpreted as characteristics-managed portfolio returns, but with improved pricing

On the one hand, this model allows time-varying factor premia, unlike most standard factor models (Giglio and Xiu, 2021). On the other hand, freely varying premia at each t would be too flexible to estimate. The break-and-trend structure, which is motivated by the economic reality, provides the necessary regularity in how factor premia vary over time, making the model estimable. In particular, the model specifies that breaks are sparse and that trends are constant.

The TV-PR estimation method is defined as, given a sample observation of excess returns $r_{i,t+1}$ and characteristics $z_{i,t}$ on a panel of $i = 1, \ldots, N$ and $t = 1, \ldots T$, the factor premia estimator $(\hat{\gamma}_t)$ solve the following optimization:[15]

$$\min_{\gamma_t \in \mathbb{R}^p} \underbrace{\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(r_{i,t+1} - z_{i,t}' \gamma_t\right)^2}_{\text{predictive regression}} + \lambda \underbrace{\sum_{j=1}^{p} \sum_{t=2}^{T} |\gamma_{t,j} - \gamma_{t-1,j}|}_{\text{total variation}} + \eta \underbrace{\sum_{j=1}^{p} \sum_{t=1}^{T} |\gamma_{t,j}|}_{\text{factor selection}}. \quad (3)$$

The first term is the standard least-squares objective for a panel predictive regression of stock returns on characteristics, except that the predictive coefficients (γ_t) are time-varying. The second term is TV-PR's key specification that uses total variation regularization to recover breaks and trends in factor premia. It regularizes changes in the γ_t sequence to induce sparse break points and piecewise-constant trend segments in the $\gamma_{t,j}$ sequence. We use the Lasso-style l_1 regularization to induce exact zero changes, also known as fused Lasso (Tibshirani et al., 2005). The third term is also a penalty term, but is used in a more traditional way. It induces sparsity in the values of γ to facilitate factor selection (rather than the changes of γ).

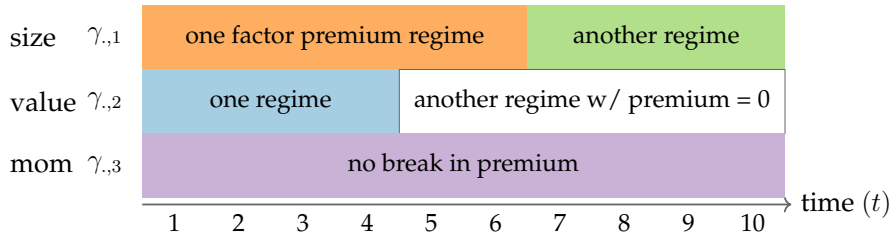
To better understand TV-PR, we examine its solution in two extreme cases of the TV regularization (while ignoring factor selection, i.e., $\eta = 0$). A very high TV penalty ($\lambda \to +\infty$) forces γ_t to remain constant over time, and the estimation reduces to the familiar pooled panel predictive regression with fixed coefficients, $r_{i,t+1} = z_{i,t}' \gamma + \epsilon_{i,t+1}$,

performances (Fama and French, 2020). More details can be found in Section 4. Direct slope factor estimation is equivalent to the TV-PR extreme case, $\lambda = 0$ (no TV regularization), as detailed further below. Alternatively, it can also be interpreted as the Barra factor model (Rosenberg, 1974) with break-and-trend factor premia. The simplification of factor loadings $\beta_{i,t}$ as $z_{i,t}$ avoids the need to estimate time-varying factor loadings, such as in IPCA, where $\beta_{i,t}$ is a linear combination $z_{i,t}$.

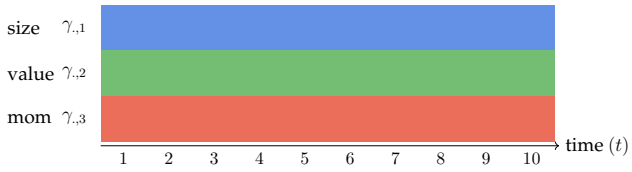
[15] Consider the true time-varying factor premia with trends and breaks in Eq. (1) is $\gamma_{o,t}$. We detail the notation formally in Section 3.

Figure 2: Factor Premia of Different Variability by TV-PR and Its Extreme Cases

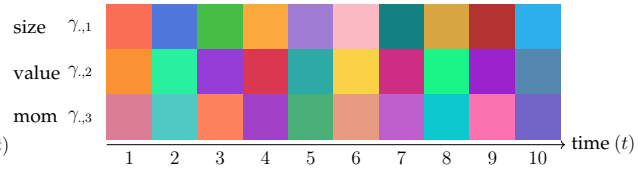
Note: Color illustrated made-up samples. Panel 2a illustrates the break-and-trend structure of a possible sample generated by the true model. Factor 1 (say size) has one structural break at $t = 6$, factor 2 has a structural break at a different time, followed by a zero premium regime, and factor 3 has no structural breaks. Panel 2b illustrates the constant factor premia model, as estimated by TV-PR extreme case $\lambda \to +\infty$. Panel 2c illustrates the period-by-period factor premia model, as estimated by TV-PR extreme case $\lambda = 0$.



(a) Breaks and trends in factor premia of the data-generating process



(b) Constant factor premia, case $\lambda \to +\infty$



(c) Period-by-period factor premia, $\lambda = 0$

with

$$\hat{\gamma}_t = \hat{\gamma} = (\sum_{i,t} z_{i,t} z_{i,t}')^{-1} \sum_{i,t} z_{i,t} r_{i,t+1}, \tag{4}$$

where the prediction error $\epsilon_{i,t+1}$ is the composite of factor realization shocks and the stock's idiosyncratic risk: $\epsilon_{i,t+1} = z_{i,t}' \nu_{t+1} + u_{i,t+1}$. Figure 2b illustrates this case. From this perspective, TV-PR with $\lambda < +\infty$ relaxes the constant coefficient assumption and is a dynamic upgrade of factor selection via predictive regressions (e.g., Freyberger et al., 2020).

On the other hand, without any TV penalty, i.e., $\lambda = 0$, the estimation decouples across t's and reduces to period-by-period cross-sectional regressions of r_{t+1} on Z_t, with

$$\hat{\gamma}_t = (Z_t' Z_t)^{-1} Z_t' r_{t+1}. \tag{5}$$

It reduces to the first stage of the Fama-MacBeth two-step procedure, where the regression coefficient is typically interpreted as an estimation of the factor realization ($\hat{f}_{t+1}$).

Fama-MacBeth second stage then averages these period-by-period $\hat{f}_{t+1}$ to obtain the factor premia estimate over a sample period, while this TV-PR extreme case directly treats factor realization $\hat{f}_{t+1}$ as the factor premia estimate ($\hat{\gamma}_t$), which is inevitably a noisy premia estimate contaminated by factor realization shocks ν_{t+1}. Figure 2c illustrates this case. From this perspective, TV-PR with $\lambda > 0$ conducts the second-stage averaging of factor realizations ($\hat{f}_{t+1}$) within adaptive segments separated by estimated break points, rather than averaging out ν_{t+1} over arbitrarily specified samples as in rolling-window Fama-MacBeth.

In summary, TV-PR innovatively bridges two well-established but distinct procedures, proposing a unified framework that adaptively interpolates between them by leveraging their combined strengths. The TV regularization allows γ_t to change over time but prevents excessively frequent shifts, thereby capturing both breaks and trends in factor premia. The balance between these two forces is controlled by the TV regularization parameter λ, which can be tuned using standard machine learning techniques, such as cross-validation. A benefit of TV-PR is that it nests and generalizes these workhorse asset pricing procedures, allowing breaks and trends to be flexibly incorporated into existing techniques, such as factor selection, as we illustrate further below.

3 Statistical Properties

This section studies the statistical properties of the proposed method. Section 3.1 develops its asymptotic theory, while Section 3.2 examines finite-sample performance through simulations and comparisons with established methods.

For the jth candidate predictor, the coefficient vector $\gamma_{\cdot,j} = (\gamma_{1,j}, \ldots, \gamma_{T,j})'$ is partitioned into B_j disjoint regimes $G_1^j, \ldots, G_{B_j}^j$ such that $G_{n_1}^j \cap G_{n_2}^j = \emptyset$ for any $n_1 \neq n_2$, and $\bigcup_{n=1}^{B_j} G_n^j = \{1, \ldots, T\}$. The total number of regimes across all predictors is denoted by $B = \sum_{j=1}^{p} B_j$. Let $|G_{\min}| = \min_{1 \leq n \leq B} |G_n|$ and $|G_{\max}| = \max_{1 \leq n \leq B} |G_n|$ denote the minimum and maximum time-varying risk premia durations. The full parameter matrix is $\Gamma = (\gamma_1, \ldots, \gamma_T)'$, and $\theta = (\theta_1, \ldots, \theta_B)' \in \mathbb{R}^B$ denotes the common parameter

values across regimes.[16] We use the subscript o to denote true parameter values θ_o.

We consider obtaining $\hat{\Gamma} = (\hat{\gamma}_1, \cdots, \hat{\gamma}_T)'$ from the following generalized double-penalized panel predictive regression procedure:

$$\min_{\gamma_t \in \mathbb{R}^p} \frac{1}{NT} \sum_{i,t} \left(r_{i,t+1} - z_{i,t}' \gamma_t\right)^2 + \sum_{j=1}^{p} \sum_{t=2}^{T} P_\lambda(|\gamma_{t,j} - \gamma_{t-1,j}|) + \sum_{j=1}^{p} \sum_{t=1}^{T} P_\eta(|\gamma_{t,j}|). \tag{6}$$

where P_η and P_λ are general folded concave penalty functions, such as the smooth clipped absolute derivatives (SCAD) by Fan and Li (2001):

$$P_\lambda(x) = \lambda|x|\mathbb{1}(0 \leq |x| < \lambda) + \frac{a\lambda|x| - (x^2 + \lambda^2)/2}{a-1}\mathbb{1}(\lambda \leq |x| \leq a\lambda) + \frac{(a+1)\lambda^2}{2}\mathbb{1}(|x| > a\lambda),$$

The same specifications apply to $P_\eta(x)$.[17] We establish the theoretical properties and implementation of the proposed method under the SCAD penalty. In Appendix D, we further elaborate on the relationship between the general folded concave penalties in Eq. (6) and the l_1-based formulation in Eq. (3).

Discussion. When econometricians have information on regime timing, they can construct a selection matrix M that encodes structural break information and links the parameter matrix to the vector of common values via $\text{vec}(\Gamma) = M\theta$.[18] Thus, the problem is simplified to detecting and estimating the nonzero components $\theta_{o\mathbb{A}} \in \mathbb{R}^q$ of $\theta_o = (\theta_{o\mathbb{A}}', \theta_{o\mathbb{Z}}')'$, where $\theta_{o\mathbb{Z}} \in \mathbb{R}^{B-q}$ contains the zero elements. Accordingly, the regime durations partition as $|G| = \{|G_1|, \ldots, |G_B|\} = \left(|G_\mathbb{A}|', |G_\mathbb{Z}|'\right)'$. Thus the estimation problem is equivalent to a weighted Lasso regularized estimator:

$$\min_{\theta \in \mathbb{R}^B} \frac{1}{NT} \|Y - \mathbf{Z}\theta\|^2 + \sum_{n=1}^{B} |G_n| P_\eta(|\theta_n|), \tag{7}$$

where the regime durations $|G_n|$ become the heterogeneous penalty weights, replacing the count of common-value occurrences in Eq. (6). Let $D_\mathbb{A} = diag(|G_{\mathbb{A},1}|^{1/2}, \ldots, |G_{\mathbb{A},q}|^{1/2})$

[16] For the jth predictor in its nth regime, we have $|G_{n'}| = |G_n^j|$ and $\theta_{n'} = \theta_n^j$, where $n' = \sum_{m=0}^{j-1} B_m + n$, where $|G_n^j|$ and θ_n^j denote the duration and common value of predictor j in regime n, respectively.

[17] SCAD requires $a > 2$. When employing the minimax concave penalty (MCP) in Zhang (2010), $P_\lambda'(x) = sign(x)[\lambda - |x|/a]$ if $|x| \leq a\lambda$, and 0 otherwise, where $a > 1$.

[18] $\text{vec}(\Gamma) = (\gamma_1', \ldots, \gamma_T')'$ and $M = (M_1', \ldots, M_T')'$ is the $pT \times B$ selection matrix with M_t being a $p \times B$ matrix with the (j, k')th element $M_{t,j,k'} = 1$ for $t \in G_{k'}^j$, $k' = \sum_{m=0}^{j-1} B_m + k$, and $M_{t,j,k'} = 0$ otherwise for all $t \in \{1, \ldots, T\}$.

collect the durations of nonzero risk premia regimes. We denote $Y = (r_2', \ldots, r_{T+1}')'$ as an $NT \times 1$ vector of dependent observations, where $r_t = (r_{1,t}, \ldots, r_{N,t})'$. Let $\mathbf{Z} = \tilde{Z}M = (\mathbf{Z}_1^1, \ldots, \mathbf{Z}_{B_1}^1, \ldots, \mathbf{Z}_1^p, \ldots, \mathbf{Z}_{B_p}^p)$ denote the $NT \times B$ data matrix, where $\mathbf{Z}_s^j$ is an $NT \times 1$ vector. Here, $\tilde{Z} = \text{diag}(Z_1, \ldots, Z_T)$, with $Z_t = (z_{1,t}, \ldots, z_{N,t})'$ being an $N \times p$ matrix.[19]

3.1 Large Sample Properties

We establish large-sample properties for our proposed estimation method, including common risk premia across regimes (θ) and the entire time-varying risk premia matrix (Γ).

Theorem 1 (Consistency). *Suppose Assumptions 1-5 in Appendix C.1 hold. A strict local minimizer $\tilde{\theta} = (\tilde{\theta}_{\mathbb{A}}', \tilde{\theta}_{\mathbb{Z}}')'$ of Eq. (7) exists. Suppose $|G_{\min}| = O(p^2 \log(p \vee q \vee T \vee N))$, we have $\|\tilde{\theta} - \theta_o\| = O_p(\sqrt{\frac{q \log^3(q \vee N)}{N|G_{\min}|}})$.*

Theorem 1 establishes the estimation consistency of our proposed method. The convergence rate depends on the cross-sectional dimension N, the minimum regime duration $|G_{\min}|$, and the total number of effective regimes q. These quantities are subject to a clear trade-off, as formalized in Assumption 5 in the Appendix. In particular, an increase in the number of regimes q leads to a slower convergence rate. We allow regimes to have heterogeneous durations; however, the minimum duration $|G_{\min}|$ must be sufficiently large to average out the factor innovations ν_t and to consistently estimate the factor premia. Finally, we allow the number of assets N across all regimes to be large, which makes the TV-PR framework suitable for high-dimensional settings.

The appearance of the $\log^3(p \vee N)$ term reflects a deliberate trade-off between method generality and convergence rate. Much of the econometric literature delivers sharper rates for asset-level time series by imposing strong temporal dependence restrictions, such as strong mixing or m-dependence. Under such conditions, the logarithmic factor can typically be improved from $\log^3(p \vee N)$ to $\log(p \vee N)$. We do not pursue this refinement. Enforcing a specific dependence structure in each regime would substantially narrow the scope of the applications, and these assumptions are difficult

[19] If $t \in G_s^j$, the elements from the $(\sum_{i=0}^{t-1} B_i + 1)$th to the $\sum_{i=0}^{t} B_i$th locations equal $z_{\cdot,t,j}$; otherwise, the elements are zeros.

to justify uniformly across assets and across regimes in empirical applications. We therefore accommodate a broad class of temporal dependence in each regime, accepting a modestly slower convergence rate in exchange for substantially wider applicability.

For the purpose of inference analysis, we denote variance $\Omega_{\mathbb{A}} = \Phi_{\mathbb{A}}^{-1}\Sigma_{\mathbb{A}}\Phi_{\mathbb{A}}^{-1}$, $\Phi_{\mathbb{A}} = \frac{1}{N}E[D_{\mathbb{A}}^{-1}\mathbf{Z}_{\mathbb{A}}'\mathbf{Z}_{\mathbb{A}}D_{\mathbb{A}}^{-1}]$ and $\Sigma_{\mathbb{A}} = \frac{1}{N}var[D_{\mathbb{A}}^{-1}\mathbf{Z}_{\mathbb{A}}'\boldsymbol{\epsilon}]$, where $\mathbf{Z} = (\mathbf{Z}_{\mathbb{A}}, \mathbf{Z}_{\mathbb{Z}})$ according to $\theta = (\theta_{\mathbb{A}}', \theta_{\mathbb{Z}}')'$. We construct a sample analog $\tilde{\Omega}_{\mathbb{A}}$ with its mathematical definition in Eq. (A-1) in Appendix C.1.

Theorem 2 (Asymptotic Normality). *Suppose Assumptions 1-5 and 6 (ii) in Appendix C.1 hold. For some unit vector $\zeta \in \mathbb{R}^q$, we have*

$$\frac{\zeta'\sqrt{N}D_{\mathbb{A}}(\tilde{\theta}_{\mathbb{A}} - \theta_{o\mathbb{A}})}{\sqrt{\zeta'\tilde{\Omega}_{\mathbb{A}}\zeta}} \xrightarrow{d} N(0,1),$$

with $|\zeta'\tilde{\Omega}_{\mathbb{A}}\zeta - \zeta'\Phi_{\mathbb{A}}^{-1}\Sigma_{\mathbb{A}}\Phi_{\mathbb{A}}^{-1}\zeta| = o_p(1)$. Moreover $P(\tilde{\mathbb{I}} = \mathbb{I}) \to 1$, where $\mathbb{I}$ is the support of the indexes of the nonzero components $\mathbb{I} = \{n : 1 \le n \le B, \theta_{o\mathbb{A},n} \neq 0\}$ and $\tilde{\mathbb{I}} = \{n : 1 \le n \le B, \tilde{\theta}_n \neq 0\}$.

Theorem 2 delivers sufficient, yet mild, conditions under which valid inference can be conducted for our risk-premia estimator. A key feature of our setup is that premia may persist over regimes of different lengths. This heterogeneity is reflected in the regime-n estimator through the scaling $\sqrt{N|G_n|}$, which enters the asymptotic variance via the regime-specific duration $|G_n|$. As a result, $\sqrt{N}D_{\mathbb{A}}\zeta'(\tilde{\theta}_{\mathbb{A}} - \theta_{oA})$ is asymptotically Gaussian with variance $\zeta'\Phi_{\mathbb{A}}^{-1}\Sigma_{\mathbb{A}}\Phi_{\mathbb{A}}^{-1}\zeta$. The linear functional indexed by ζ accommodates standard hypotheses of interest. For example, to test a hypothesis about the risk premium in a particular regime n, set $\zeta_n = 1$ and $\zeta_m = 0$ for all $m \neq n$. Setting $\zeta_n = 1/\sqrt{q}$ for all $n = 1, \ldots, q$ yields a test of whether the average risk premium across regimes is zero. Moreover, Theorem 2 also establishes selection consistency of TV-PR and hence facilitates the detection of breaks and trends in factor risk premia.

Theorem 3 (Oracle Property). *Suppose Assumptions 1-5 and 6 (i) in Appendix C.1 hold. A positive constant c exists whereby $\min_{1\le n\le q}|\theta_{\mathbb{A},n}| > c\eta$ and $\min_{1\le j\le p, B_j>1}\min_{n_1\neq n_2}|\theta_{on_1}^j -$*

$\theta^j_{on_2}| > c\lambda$. *A local minimizer* $\hat{\Gamma}$ *of the penalized criterion Eq.* (6) *exists such that*

$$P(\hat{\Gamma} = \tilde{\Gamma}^*) \to 1,$$

where $\tilde{\Gamma}^*$ *is the oracle estimator corresponding to* $\tilde{\theta}^* = (\tilde{\theta}^{*\prime}_{\mathbb{A}}, \mathbf{0}_{B-q})'$ *that* $vec(\tilde{\Gamma}^*) = M\tilde{\theta}^*$*, which minimizes the following criterion:* $\tilde{\theta}^*_A = \arg\min_{\theta_A \in \mathbb{R}^q} \frac{1}{NT} \sum_{t=1}^{T} \|Y - \mathbf{Z}_{\mathbb{A}}\theta_{\mathbb{A}}\|^2$.

Theorem 3 shows that, under mild conditions—namely, a sufficiently strong weakest signal and sufficient separation of signals across regimes—our estimator achieves oracle properties. In particular, although our TV-PR procedure is implemented without knowledge of the regime partition or the subset of factors with nonzero premia, it is asymptotically equivalent in probability to the infeasible oracle estimator that is computed using the true grouping and the true active set. This oracle equivalence implies that the statistical cost of learning the latent structure is asymptotically negligible, so that the feasible estimator behaves, to first order, as if the regime structure were known ex ante. Taken together, the results imply that our method consistently recovers the time-varying and individually heterogeneous grouping structure of risk-factor premia, including both the regime partition and the associated regime-specific premium levels.

3.2 Finite Sample Properties

In this section, we evaluate the finite-sample performance of the TV-PR estimator and benchmark it against established methods across varying sample sizes and cross-sectional dimensions. The analysis utilizes a calibrated data-generating process (DGP) based on empirical data.

We simulate a three-regime, three-factor specification of model (1), using data-generating parameters calibrated to empirical factor and stock returns. Specifically, we set

$$r_{i,t+1} = z^{[1]}_{i,t} f^{[1]}_{t+1} + z^{[2]}_{i,t} f^{[2]}_{t+1} + z^{[3]}_{i,t} f^{[3]}_{t+1} + u_{i,t+1}, \quad t = 1, \ldots, T, \tag{8}$$

where each factor is generated with time-varying, piecewise-constant factor premia, featuring structural breaks at $T/3$ and $2T/3$. In particular, we define $f_{t+1}^{[p]} = \gamma_t^{[p]} + \nu_{t+1}^{[p]},\ p = 1, 2, 3$, where $\nu_{t+1}^{[p]}$ varies across the three regimes.[20] We simulate both factor innovation as ν_{t+1} and idiosyncratic errors as i.i.d. Normal variables with calibrated standard deviations.[21] The predictors are generated from a VAR(1) process: $z_{i,t} = Az_{i,t-1} + e_{i,t},\ e_{i,t} \sim N(0, \Sigma)$.[22]

We experiment with four sample-size configurations to assess the finite-sample performance of TV-PR with $T = 120$ or 240 and $N = 500$ or 1000. For each configuration, we conduct 50 simulations to evaluate the estimation accuracy. To assess the finite-sample performance of our method, we compare TV-PR with several alternatives, including time-invariant model selection methods (LASSO and Ridge), a standard time-invariant estimator (OLS), and time-varying approaches based on a 120-month rolling average and a recursive average. We evaluate estimation accuracy using two performance metrics: (i) average absolute estimation error (Abs.Err.), defined as $(pT)^{-1}\sum_j^p \sum_t^T |\hat{\gamma}_{t,j} - \gamma_{o,t,j}|$; and (ii) the predictive R^2, defined as $1 - \frac{\sum_{i,t}(r_{i,t+1} - \hat{r}_{i,t+1})^2}{\sum_{i,t} r_{i,t+1}^2}$.

Figure 3 presents the true risk premia and their estimates from 50 simulations ($N = 1000$ and $T = 240$). The dashed lines represent the true risk premia, while the solid lines depict the estimates from each simulation, with three colors corresponding to the three factors, respectively.

Overall, the simulation results validate the effectiveness of our time-varying risk premia estimation framework. In most simulations, the estimated risk premia closely track the true values, suggesting the factor innovations within each regime are effectively averaged out to reveal the underlying risk premia.[23] The timing of breaks aligns

[20] Specifically, the factor premia are specified as the $\gamma_t^{[p]} = (1 + E_t^{\text{ann}}[f_{t+1}^{[p]}])^{1/12} - 1$, where $E_t^{\text{ann}}[f_{t+1}^{[p]}]$ denotes the annual expected factor return, following: $E_t^{\text{ann}}[f_{t+1}^{[1]}] = 10\%,\ E_t^{\text{ann}}[f_{t+1}^{[2]}] = 8\%,\ E_t^{\text{ann}}[f_{t+1}^{[3]}] = -3\%$, for $t = 1, \ldots, T/3$; $E_t^{\text{ann}}[f_{t+1}^{[1]}] = -10\%,\ E_t^{\text{ann}}[f_{t+1}^{[2]}] = 3\%,\ E_t^{\text{ann}}[f_{t+1}^{[3]}] = 6\%$, for $t = T/3 + 1, \ldots, 2T/3$; $E_t^{\text{ann}}[f_{t+1}^{[1]}] = 5\%,\ E_t^{\text{ann}}[f_{t+1}^{[2]}] = -5\%,\ E_t^{\text{ann}}[f_{t+1}^{[3]}] = 12\%$, for $t = 2T/3 + 1, \ldots, T$.

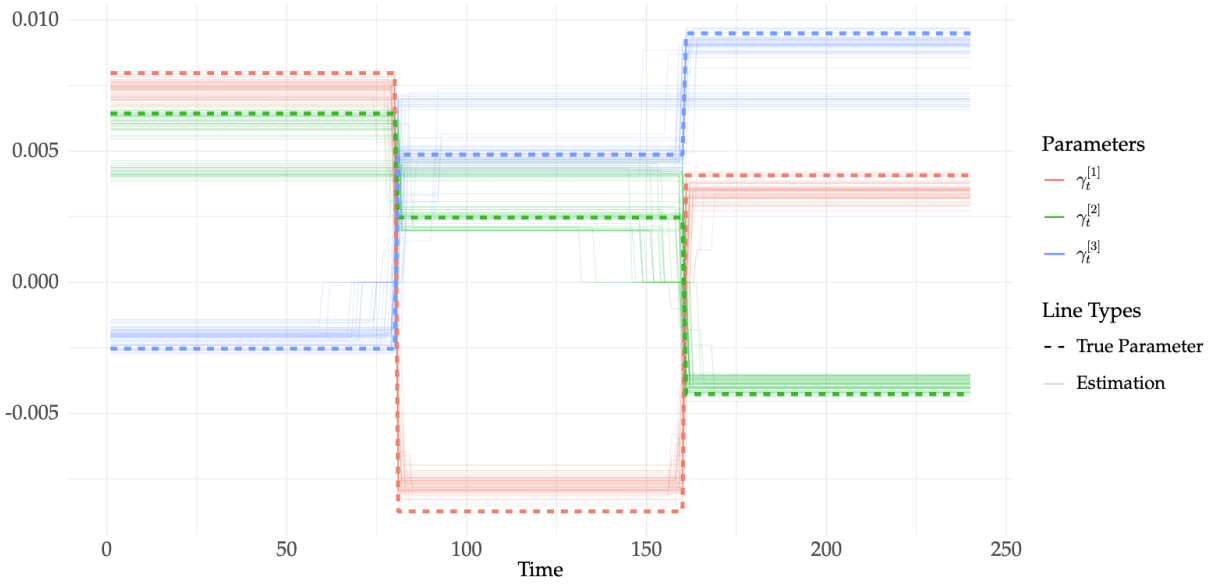
[21] $\nu_{t+1} \sim$ i.i.d. $N(0, \sigma_\nu^2)$, with $\sigma_\nu = 0.29\%$. The standard deviation of the idiosyncratic errors is calibrated such that the total R^2 of Eq. (1) is approximately 20%. The total R^2 is defined as $1 - \sum_{i,t} \varepsilon_{i,t+1}^2 / \sum_{i,t} r_{i,t+1}^2$.

[22] Matrices A and Σ are estimated with empirical characteristics. For each i, the sequence of $z_{i,t}$ is generated from a separate VAR process, discarding the first ten realizations.

[23] The Fama–MacBeth regression conducts period-by-period cross-sectional regressions and then averages the estimated coefficients over time to obtain the risk premia. This time-series averaging helps

Figure 3: Estimation of Risk Premia from 50 Simulations

Note: This figure illustrates the estimation of risk premia across 50 simulations with $N = 1000$ and $T = 240$. The transparent solid lines represent the estimated risk premia, while the dashed lines indicate their true values.



well with the true breaks, though it occasionally appears earlier or later than the true structural breaks. Nonetheless, these deviations are modest, and the method successfully recovers the true regime structure in most cases.

However, some estimation errors and biases are observed and behave as expected under TV-PR's design. First, in some simulations, the green lines across regimes 1 and 2 are fused together. This occurs as the total variation penalty fails to detect the first break point in factor 2's risk premia, given the relatively small magnitude of change. In such cases, the estimates tend to converge to an average of the true risk premia across the adjacent regimes, reflecting the difficulty of detecting weak structural breaks. Third, the risk premium estimates exhibit mild shrinkage toward zero, a direct consequence of the Lasso penalty. This regularization effect is particularly beneficial in high-dimensional settings, as it guards against overfitting and enhances out-of-sample stability, albeit at the cost of introducing slight bias. For the same reason, in a few simulations, the factor premia are erroneously set to zero in between

average out factor innovations in the estimated premia, as such innovations can induce substantial estimation errors (Giglio and Xiu, 2021). Our method requires a minimum regime duration to ensure that factor innovations are sufficiently averaged out within each regime.

two regimes where the premia switch signs (blue lines around $T = 80$ and green lines around $T = 160$).

Table 1: Average Absolute Estimation Error of Risk Premia

Note: This table compares the estimation accuracy of different methods across two time horizons ($T = 120$ and $T = 240$) and two sample sizes ($N = 500$ and $N = 1000$). Coefficient precision is evaluated using the average absolute estimation error (Abs.Err), expressed as a percentage. Results are averaged over 50 simulations, with standard deviations in parentheses.

	$N = 500, T = 120$				$N = 500, T = 240$			
	TV-PR	Lasso	Ridge	OLS	TV-PR	Lasso	Ridge	OLS
All Regimes	0.07	0.49	0.49	0.49	0.08	0.49	0.49	0.49
	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.00)	(0.00)	(0.00)
	$N = 1000, T = 120$				$N = 1000, T = 240$			
	TV-PR	Lasso	Ridge	OLS	TV-PR	Lasso	Ridge	OLS
All Regimes	0.07	0.49	0.49	0.49	0.08	0.49	0.49	0.49
	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.00)	(0.00)	(0.00)

Table 1 presents estimation accuracy across sample sizes, highlighting the robustness of our method relative to alternatives. TV-PR achieves average absolute estimation errors approximately seven times lower than those of competing methods, demonstrating superior precision in parameter estimation. In contrast, time-invariant estimation methods can only estimate the time-invariant risk premia as the average across all regimes, thereby inducing large estimation errors.

Table 2 reports the predictive R^2 across different methods. First, TV-PR consistently achieves the highest predictive R^2 across the three regimes, whereas the time-invariant methods exhibit mixed performance across regimes, highlighting their limited ability to capture dynamic relationships. This demonstrates TV-PR's reliability in producing accurate predictions under varying conditions and structural breaks. Second, the time-varying estimation methods, rolling and recursive averages, perform surprisingly worse than the time-invariant methods, as these two approaches can only capture gradual changes rather than structural breaks. However, whether rolling or breaks perform better depends on the true DGP. Determining which approach is more applicable in real-world settings is an empirical question. Later, we show that the break-based method indeed performs better, suggesting that the true DGP of financial data is characterized more by structural breaks than by gradual rolling changes.

Table 2: Predictive R^2 across Different Models

Note: This table compares model predictability across different methods for two time horizons ($T = 120$ and $T = 240$) and two sample sizes ($N = 500$ and $N = 1000$). Predictability is evaluated using the predictive R^2, reported in percentages. Results are averaged over 50 simulations, with standard deviations in parentheses.

			$N = 500, T = 120$						$N = 500, T = 240$			
	TV-PR	Lasso	Ridge	OLS	Rolling	Recursive	TV-PR	Lasso	Ridge	OLS	Rolling	Recursive
Regime 1	17.82	-0.64	-0.56	-0.61	0.21	1.04	17.64	-0.61	-0.53	-0.58	0.35	1.65
	(1.68)	(1.41)	(1.39)	(1.41)	(1.04)	(1.08)	(1.32)	(0.97)	(0.95)	(0.97)	(0.72)	(0.56)
Regime 2	16.46	1.98	1.96	1.94	-4.83	-2.72	16.28	2.06	2.04	2.02	-4.11	-1.90
	(1.17)	(1.15)	(1.14)	(1.15)	(0.82)	(0.80)	(1.02)	(0.81)	(0.80)	(0.81)	(0.56)	(0.38)
Regime 3	19.48	7.65	7.59	7.66	3.38	0.16	19.33	7.71	7.65	7.72	4.11	0.22
	(1.55)	(1.24)	(1.22)	(1.24)	(1.33)	(0.69)	(1.30)	(0.96)	(0.95)	(0.96)	(0.88)	(0.41)
All	17.96	3.05	3.05	3.05	-0.35	-0.48	17.78	3.09	3.09	3.09	0.17	0.00
	(0.81)	(0.45)	(0.45)	(0.45)	(0.67)	(0.29)	(0.71)	(0.32)	(0.32)	(0.32)	(0.42)	(0.13)
			$N = 1000, T = 120$						$N = 1000, T = 240$			
	TV-PR	Lasso	Ridge	OLS	Rolling	Recursive	TV-PR	Lasso	Ridge	OLS	Rolling	Recursive
Regime 1	17.62	-0.79	-0.71	-0.77	0.08	0.86	17.54	-0.72	-0.64	-0.69	0.30	1.61
	(1.50)	(1.19)	(1.17)	(1.19)	(0.94)	(0.94)	(1.14)	(0.75)	(0.74)	(0.75)	(0.60)	(0.44)
Regime 2	16.48	2.02	2.00	1.98	-4.71	-2.47	16.22	2.08	2.06	2.04	-4.07	-1.94
	(1.21)	(0.98)	(0.97)	(0.98)	(0.79)	(0.63)	(0.96)	(0.82)	(0.81)	(0.82)	(0.55)	(0.34)
Regime 3	19.43	7.69	7.62	7.70	3.38	0.00	19.26	7.68	7.61	7.69	4.11	0.32
	(1.58)	(1.30)	(1.28)	(1.31)	(1.20)	(0.64)	(1.14)	(0.81)	(0.80)	(0.81)	(0.65)	(0.26)
All	17.88	3.02	3.02	3.02	-0.36	-0.52	17.70	3.05	3.05	3.05	0.17	0.01
	(0.79)	(0.39)	(0.39)	(0.39)	(0.56)	(0.22)	(0.64)	(0.28)	(0.28)	(0.28)	(0.30)	(0.11)

4 Estimation Results

We begin by presenting break and trend estimates in both single- and multi-factor settings, and then introduce results for return-based factors, which offer longer histories, clearer interpretations, and greater prevalence in technical trading. In the next section, we examine the performance of investment strategies implied by trend- and break-based signals, derived from either estimated time-varying risk premia or predicted returns. Finally, we document the time-varying factor composition of the SDF and discuss its implications relative to time-invariant asset-pricing factor models.

4.1 Data

Our analysis uses a standard dataset of U.S. stock characteristics and returns. We use monthly U.S. equity data spanning 2004 to 2023. We apply data filters consistent with those used in constructing the Fama–French factors. Specifically, the sample is restricted to stocks listed on the NYSE, AMEX, or NASDAQ for at least 1 year, with CRSP share codes 10 or 11. Firms with negative book equity or lagged market equity are excluded to maintain data integrity. The monthly median and mean stock

numbers are 3,694 and 3,860, respectively. The dataset encompasses 61 firm characteristics observed monthly, categorized into eight clusters: momentum, value, investment, profitability, size, volatility, liquidity, and intangibles. These characteristics are standardized to the interval $[-1, 1]$ on a cross-sectional basis.

4.2 Estimation Results in Single-Factor Settings

The single-factor setting offers a simpler laboratory to illustrate the inner workings of how TV-PR estimations behave. This setting streamlines the model while preserving its essential time-varying structure. Notwithstanding the relative simplicity of the setting, TV-PR already shows the empirical relevance and power of modeling breaks and trends in factor premia.

Specifically, the single-factor setting uses one characteristic at a time. We append a column of ones to that characteristic. As a result, $p = 2$, and the data matrix Z_t has dimensions $N_t \times 2$, where N_t is the number of stocks in the cross-section. We combine the return panel and this Z_t matrix in the TV-PR estimation in Eq. (6). In all single-factor estimations, the factor-selection term (the third term) is turned off by setting $\eta = 0$, given that there is no need to select among multiple factors. The results below focus on the behavior of the estimation as the TV regularization parameter (λ) varies.

As discussed in Section 2, TV-PR estimates the dynamic premia of slope factors (also known as cross-sectional factors).[24] In particular, the realized slope factor return is:

$$\hat{f}_{t+1} = (Z_t'Z_t)^{-1}Z_t'r_{t+1} = \gamma_t + (Z_t'Z_t)^{-1}Z_t'\epsilon_{t+1}, \tag{9}$$

where $\epsilon_{t+1} = Z_t\nu_{t+1} + u_{t+1}$ is the composite of factor innovations and idiosyncratic shocks. The slope factor has two interpretations. First, it is the TV-PR estimation in the extreme case of no TV regularization (see Figure 2b). Specifically, $\gamma_t = \hat{f}_{t+1}$, when $\lambda = 0$. This is, inevitably, a very poor estimate of the ex ante factor premium, as it

[24] When Z_t includes a column of ones, the intercept estimate $\hat{\gamma}_{0,t}$ corresponds to a market portfolio. When each characteristic is also cross-sectionally standardized to have zero mean, this portfolio becomes equivalent to an equal-weighted long-only market portfolio. The remaining elements of $\hat{\gamma}_t$ represent slope coefficients on the characteristics. For example, $\gamma_{BM,t}$ is the coefficient on the book-to-market ratio, $Z_{BM,t}$. The slope coefficients represent returns on long-short, zero-investment portfolios (cross-sectional factors) when Z_t includes an intercept or is mean-standardized.

is highly contaminated by the factor realization shock (ν_{t+1}). Second, the properly estimated γ_t with $\lambda > 0$ aims to recover the breaks and trends in the observed slope factors f_{t+1}.

We assess whether the estimation results indeed exhibit the expected behaviors of TV-PR. In single-factor settings, we plot the second entry of $\hat{f}_{t+1}$, which corresponds to the slope factor return, while the first constant entry corresponds to a version of the market portfolio. Similarly, we also plot the second entry of $\hat{\gamma}_t$.

Figure 4 plots cumulative sums of both $\hat{f}$ and $\hat{\gamma}$. As a result, the cumulative sum of $\hat{f}$ is the familiar visualization of factor performance (price) charts, which investors rely on to visually and subjectively identify trends. On the other hand, the cumulative sum of γ visualizes TV-PR's formal estimations of breaks and trends. Figure 4 plots the different sets of γ_t estimates at three levels of TV regularization constants (λ). The six panels show six characteristic examples: Market beta (CAPM beta), Size (market equity), Value (book-to-market ratio), Momentum (12-month horizon), Investment (asset growth rate), and Profitability (operating profitability).

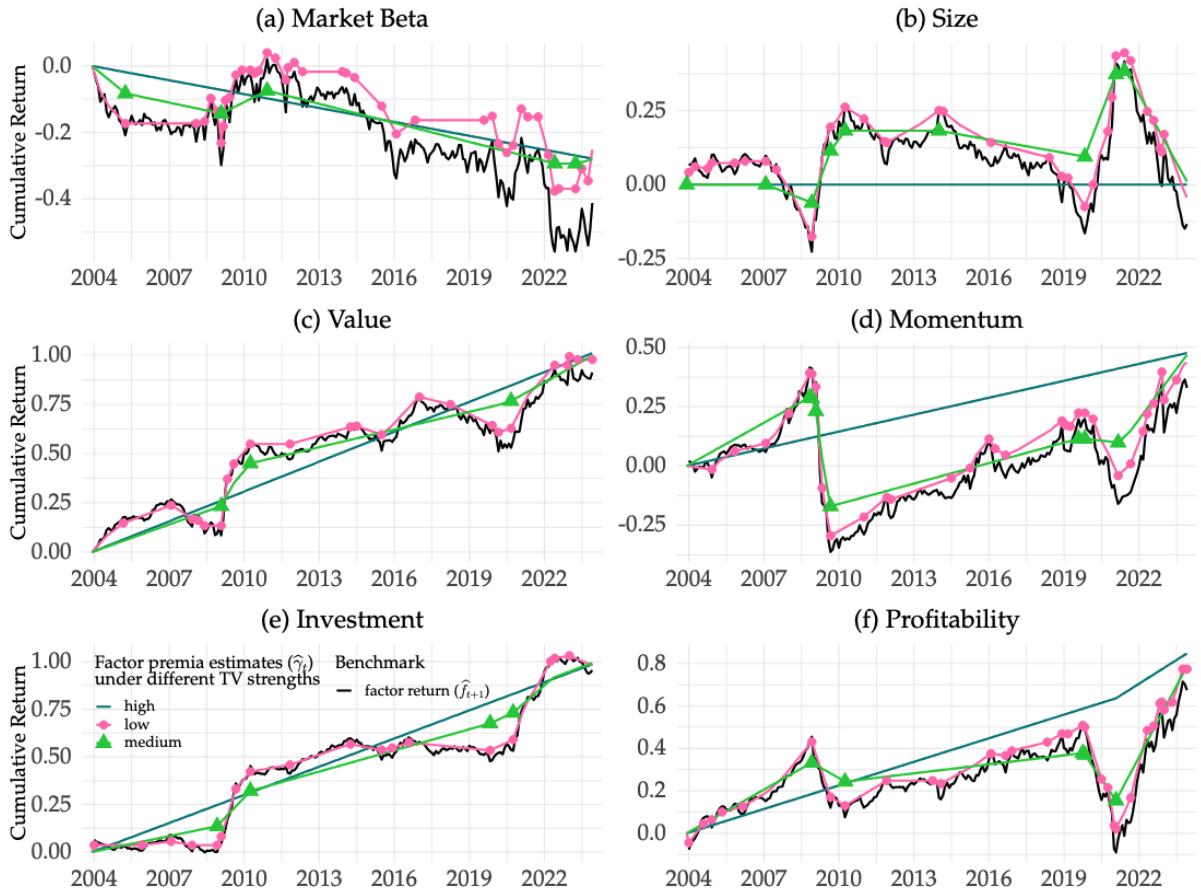
Overall, the results exhibit very intuitive trend lines. The cumulative sums of γ indeed behave in a piecewise constant pattern, where the constant segments correspond to the trend. The slopes of these segments represent factor premia estimates, while the turning points correspond to structural breakpoints. The estimated trend lines are close to the realized factor returns and closely resemble what subjective trend-fitting methods would produce.

An important aspect of identifying the underlying trends in these factor premia is TV regularization strength, which corresponds to a crucial trade-off. If coefficients change excessively, the estimator may be overwhelmed by noise introduced by frequent structural breaks. Conversely, prolonged periods of constant coefficients may obscure meaningful structural changes, leading to missed detections.

Specifically, in Figure 4, factor premia estimated with medium regularization (green line) effectively capture regime shifts, such as the Global Financial Crisis and the COVID-19 shock, while also reflecting trends in realized factor returns. This suggests

Figure 4: Cumulative Return of Time-varying Risk Premium and Slope Factors

Note: This figure presents cumulative returns for six slope factors and their corresponding time-varying factor premia estimates, evaluated under varying regularization levels. Structural shifts in the data are highlighted through estimated breakpoints. Eq. (1) is estimated using a sample spanning the period from 2004 to 2023.



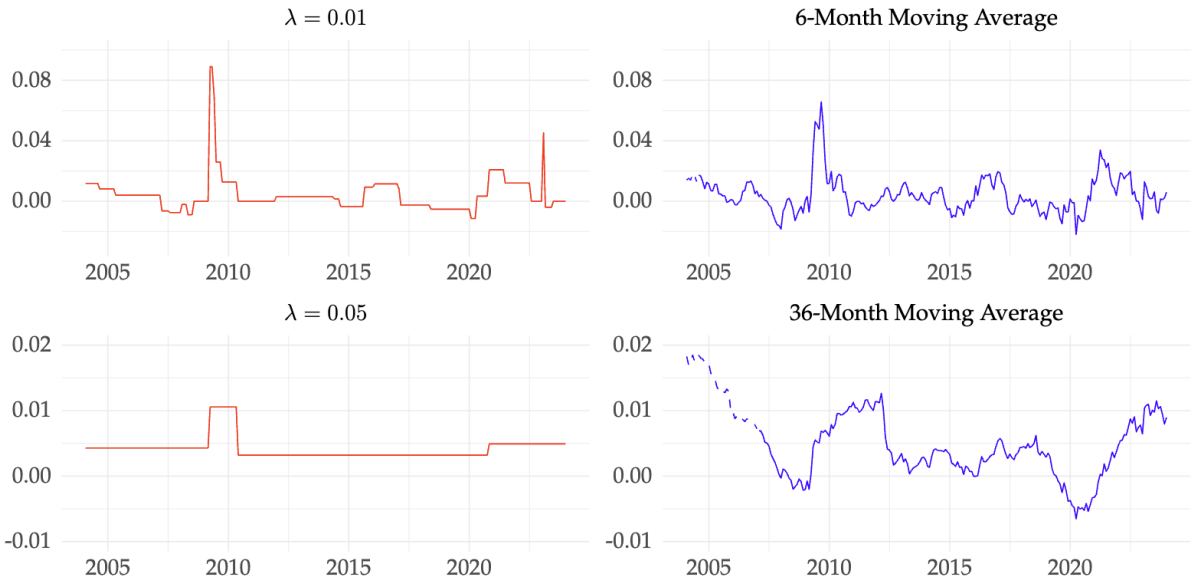
that TV-PR identifies regime-dependent factor premia. Meanwhile, low regularization (red line) yields too many breaks, causing it to track the factor estimation ($\hat{f}_{t+1}$, black line) rather than the underlying risk premium. In contrast, high regularization (blue line) yields nearly constant estimates that capture only broad trends while ignoring structural breaks. In addition to common break points associated with major economic events, individual factor premia exhibit distinct patterns. For instance, the size factor shows a unique structural break at the beginning of 2014, which is absent in other factors under medium regularization. Additionally, Panel (d) of Figure 4 highlights momentum crash periods, marked by sharp declines in profitability during market reversals. These findings demonstrate TV-PR's ability to capture both broad market

disruptions and factor-specific dynamics.

We next zoom in on the single-factor premium to show piecewise-constant patterns and structural break points. Figure 5 changes the visualization to the γ series rather than its cumulative sum and compares it to model-free moving average "trend estimates". Specifically, we have two levels of TV regulation and two moving average smoothing window lengths (6 and 36 months). We use the book-to-market ratio as an example; other factors behave similarly.

Figure 5: Time-Varying Risk Premium Estimation for BE/ME

Note: This figure illustrates the time-varying risk premium of the value factor, proxied by the book-to-market ratio. We compare four approaches: two derived via our methodology under varying regularization levels, and the 6-month and 36-month moving averages. Dotted lines denote the initial estimation sample used for computing the moving averages.



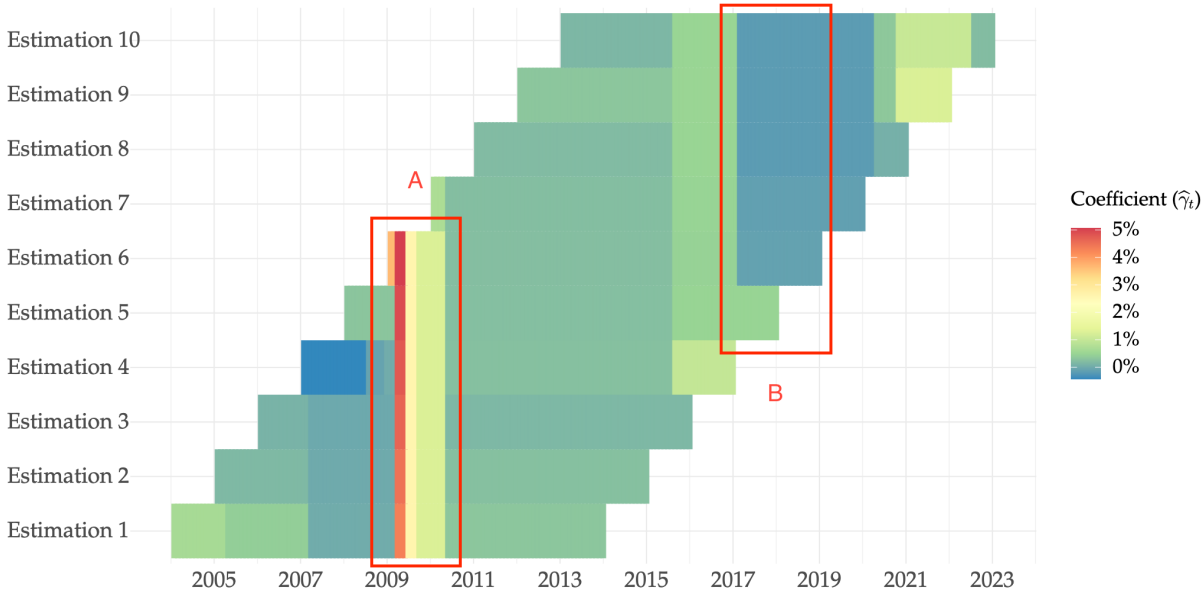
The low regularization level in our method resembles the behavior of a short-window moving average, but with distinct structural break points. As shown in the upper panel of Figure 5, our estimation ($\lambda = 0.01$) aligns closely with a 6-month moving average. However, there are two important differences between the TV-PR estimates and the moving average. First, the piecewise-constant part in factor premia clearly captures the trend shown in Figure 4 when viewed through cumulative returns. In contrast, the moving-average estimates continuously evolve, capturing both factor premia and factor innovations, and thus lead to a factor–return rather than a fac-

tor–premia estimation. Second, TV-PR highlights structural changes through abrupt breaks rather than gradual ones, helping us identify shifts in regimes. These differences are more salient at a medium regularization level ($\lambda = 0.05$), where TV-PR captures major peaks and dips, such as those around 2010 and 2020, as illustrated in the left lower panel of Figure 5. On the contrary, although a 36-month moving average reflects these changes, TV-PR introduces minor fluctuations that obscure clear breakpoints. Overall, TV-PR provides a more interpretable representation of time-series changes by emphasizing sparse structural shifts. By avoiding overly smooth transitions, it facilitates sharper insights into underlying trends and breaks, offering a valuable alternative to traditional smoothing techniques.

Next, we show that factor premia break estimates are consistent across rolling estimation windows and that TV-PR can detect breaks in near-real time. These properties form the foundation for rolling-window-based investment strategies discussed further below. Figure 6 plots the gamma estimates from a series of 10 rolling-window estimates. Each estimation window is 10 years long and is rolled forward by 1 year.

Figure 6: Stability of Break Identification for BE/ME

Note: This figure presents 10-year rolling estimates of the dynamic risk premium for the value factor, proxied by the book-to-market ratio. To ensure time-series independence, estimates are computed with a one-year gap between windows. The regularization parameter is set to $\lambda = 0.05$, balancing model flexibility and control over overfitting.



The rolling-window estimation consistently identifies structural break points, as shown in Figure 6. First, focusing on rectangular A, Estimations 1 through 6 consistently detect a pronounced break around 2010, accompanied by an increase in the estimated risk premium, highlighted in red. Notably, this break remains identifiable even when the sample period begins immediately prior to the event, as shown in Estimations 5 and 6. This evidence suggests that TV-PR can detect breakpoints with relatively few observations surrounding the event. Turning to rectangular B, Estimations 6 through 10 identify a structural break in 2017. Estimation 5, which ends in 2018, does not detect this break; however, all subsequent estimations (6–10) retrospectively identify a break in 2017. This pattern suggests that TV-PR can detect structural breaks in near real time with minimal delay, similar to the findings of Smith and Timmermann (2021). If the most recent trend persists, near-real-time detection remains valuable for OOS prediction and trading strategies. Indeed, the later rolling-window analysis shows that TV-PR rapidly identifies regime shifts, enabling actionable OOS trading signals.

Moreover, the estimated risk premia within each regime are highly stable across sample periods. For example, risk premia estimates between 2011 and 2015 are consistently close to 0.5% across all 10 estimations. Overall, these findings highlight TV-PR's robustness in detecting shifts in risk premia across diverse sample periods.

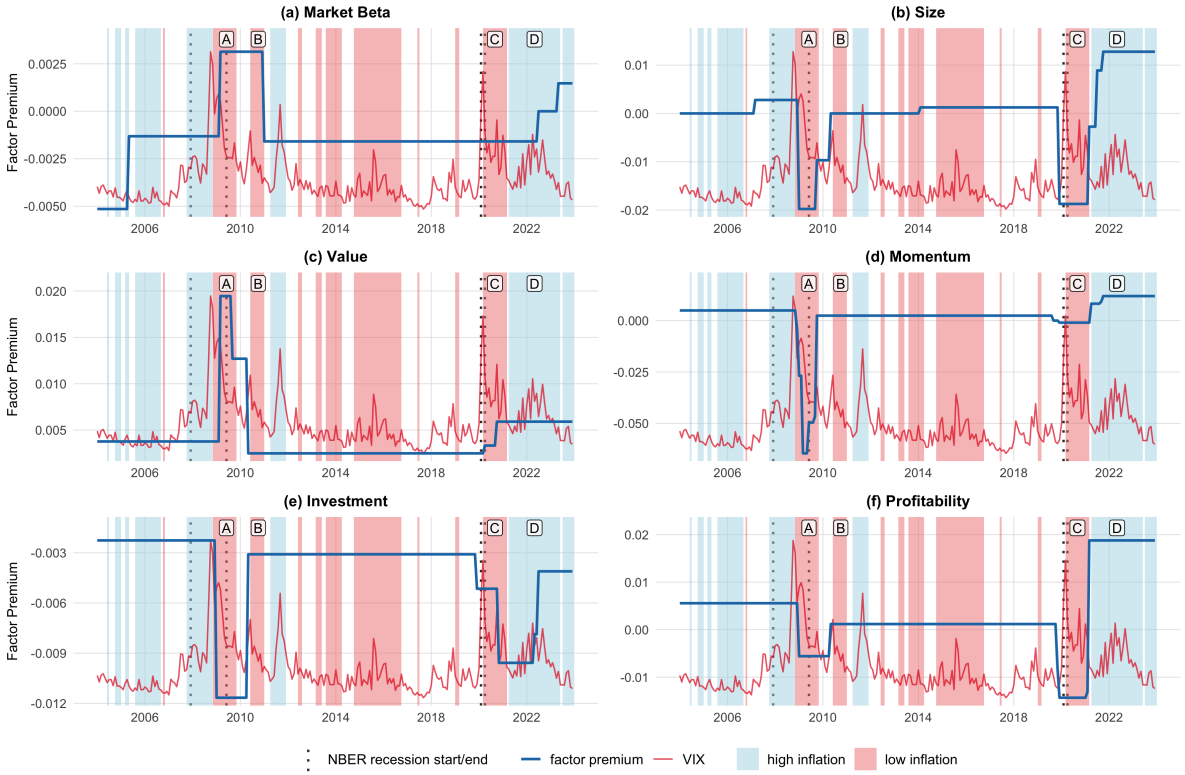
4.3 Factor Premia Breaks and Macroeconomic Regimes

To validate the economic relevance of our TV-PR estimates, we examine whether the identified structural breaks coincide with fundamental shifts in market conditions. Figure 7 plots the time-varying factor premia estimates alongside NBER recession periods, inflation regimes, and market volatility (VIX). This comparison allows us to assess whether the statistical breaks detected by our model reflect genuine macroeconomic state variables rather than idiosyncratic noise.

Overall, the estimated structural breaks in factor premia align closely with major macroeconomic changes, suggesting that shifts in the economic environment are systematically reflected in factor pricing dynamics. First, the breaks in estimated factor

Figure 7: Factor Premia Breaks and Macroeconomic Regimes

Note: This figure presents time-varying factor premia estimates from Figure 4, together with NBER recessions, rescaled daily VIX, and year-over-year CPI inflation. High- and low-inflation regimes are defined as the top and bottom 30% of inflation observations within the sample.



premia correspond closely to transitions across inflation regimes. In particular, for the low-inflation Regimes A and B, the estimated break points occur near both the onset and the termination of these regimes. For example, the market beta exhibits one break close to the start of Regime A and another precisely at the end of Regime B, and other factor premia display a similar pattern.

The transition from the low-inflation Regime C to the high-inflation Regime D is also associated with pronounced breaks in several factor premia. In particular, the size, momentum, and profitability factors exhibit breaks that closely align with the shift from Regime C to Regime D, indicating that the change in the inflation environment is accompanied by a reconfiguration of factor risk compensation.

Second, the estimated breaks are closely aligned with major changes in market uncertainty, as proxied by the VIX. All six factor premia estimates exhibit breaks near

the first VIX peak at the beginning of Regime A, while some factors, such as investment, size, momentum, and profitability, also display breaks that coincide with the rapid decline following the VIX peak. This pattern suggests that sharp fluctuations in aggregate risk sentiment play an important role in driving changes in factor premia.

Finally, the largest changes in factor premium estimates coincide with NBER-defined recessions, particularly during the Global Financial Crisis, and nearly all estimates exhibit pronounced shifts around the COVID-19 recession. Together, these findings indicate that factor premia dynamics are tightly linked to business-cycle conditions, highlighting the economic significance of macroeconomic downturns for asset pricing.

4.4 Estimation Results in Multi-Factor Settings

We estimate the multi-factor model using 61 firm characteristics to capture time-varying premia in Eq. (1). Given that the data matrix Z_t has dimension $N_t \times 61$, we account for both time variation and variable selection by employing a combination of total variation and LASSO regularization. The strengths of the two penalties are governed separately by λ and η, as specified in Eq. (6).

The optimal change frequency should strike a balance between the ability to detect structural breaks and the need to construct a parsimonious, piecewise-constant representation of the underlying trend. We implement a machine–learning–based tuning procedure to optimize the regularization levels λ and η. Specifically, the hyperparameters associated with the penalties are optimized through OOS cross-validation to maximize predictive accuracy, measured by OOS predictive R^2. The detail is in Appendix D.3. Using a rolling window approach, we train the model on a ten-year period and forecast the subsequent one year.

Table 3 reports stock return predictability across size-based groups for TV–PR, using the pair of hyperparameters that yields a positive overall OOS predictive R^2. The results show that TV–PR delivers relatively strong OOS performance for mega and large firms, while predictive accuracy declines for small and micro stocks. For the nano group, TV–PR produces a negative OOS R^2, indicating that returns of extremely

Table 3: Out-of-Sample Stock Return Predictability

Note: This table reports OOS predictive $R^2(\%)$ for monthly stock returns. Stocks are grouped into five non-overlapping size categories: Mega, Large, Small, Micro, and Nano Cap, with breakpoints based on the market equity of NYSE stocks at the end of each month. All models are trained on 10 years of historical data and generate forecasts for the following year. Each model is updated annually.

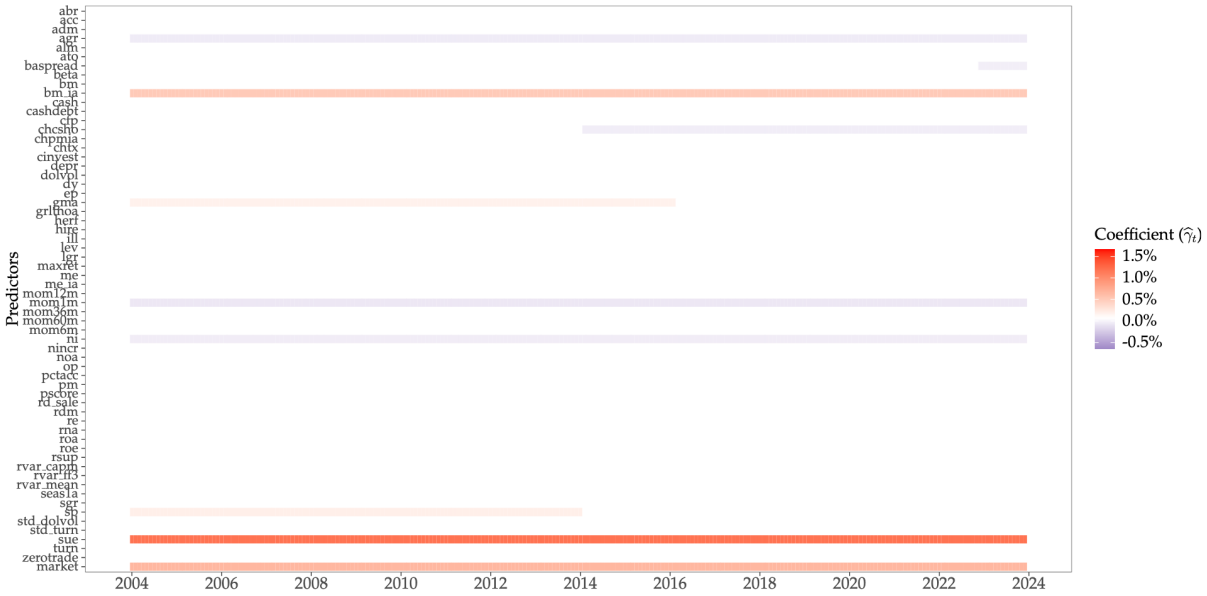
	All Stocks	Size Groups				
		Mega	Large	Small	Micro	Nano
Rolling OLS	0.04	0.35	0.40	0.20	0.33	-0.86
Rolling LASSO	0.07	0.82	0.77	0.44	0.31	-1.10
TV-PR	0.14	0.73	0.76	0.51	0.42	-1.05

small firms are not reliably predictable.

In comparison, OLS performs worse than TV–PR across all size groups, except for nano stocks, where predictive performance remains negative. LASSO performs slightly better for mega stocks and comparably for large stocks, but it does not outperform TV–PR across the remaining size groups and delivers overall predictive performance approximately half that of TV–PR. Overall, TV–PR shows economically meaningful predictive power for individual stock returns.

Figure 8: Time-Varying Factor Premia

Note: This figure depicts time-varying premia linked to slope factors, estimated from the dynamic coefficients in Eq. (1), which map factor returns to firm-specific attributes. The color bar illustrates the magnitude and thresholds of key firm characteristics, highlighting their relative importance. The intercept represents the equally weighted market portfolio, serving as the benchmark.



Next, we focus on the full-sample analysis of time-varying risk premia to explore

the persistence and variability of each characteristic's impact on expected returns. In Figure 8, among the characteristics analyzed, standardized unexpected earnings (`sue`) emerge as a robust predictor of cross-sectional stock returns, with a monthly risk premium exceeding 1.1%. This finding suggests the importance of earnings surprises in explaining equity returns. The industry-adjusted book-to-market ratio (`bm_ia`) also contributes positively, emphasizing the relevance of sector-neutral valuation measures in asset pricing. Other characteristics, such as sales-to-price (`sp`), gross profitability (`gma`), and change in shares outstanding (`chcsho`), exhibit positive risk premia, although their contributions fluctuate over time.

4.5 Breaks and Trends of Momentum, Reversal, and Seasonality Anomalies

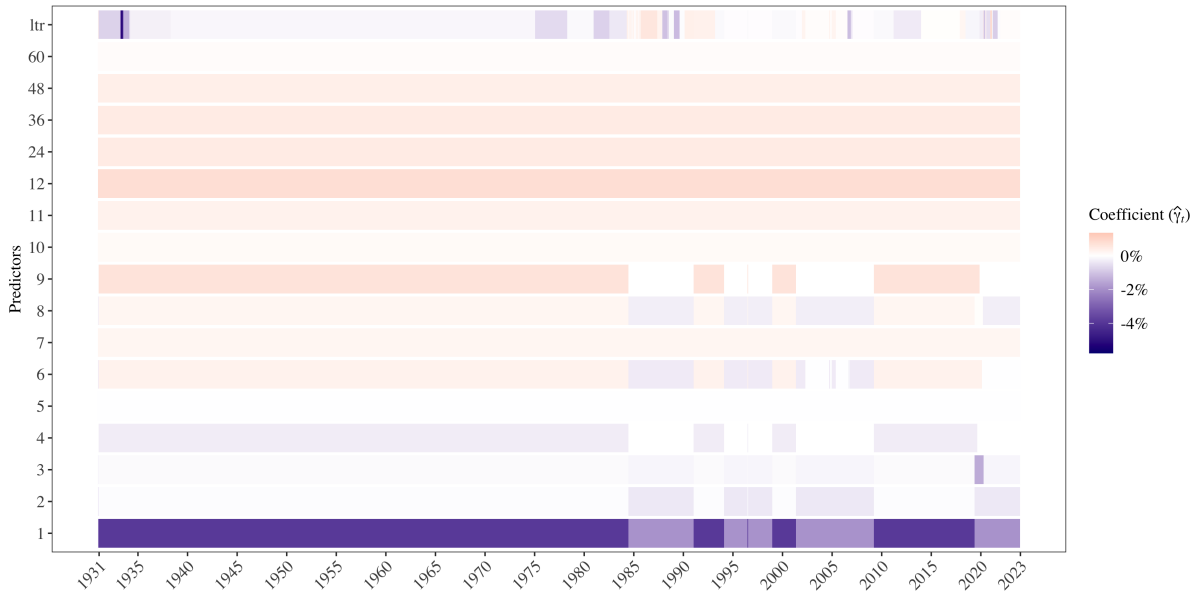
In this section, we examine the evolution of risk premia associated with return-based predictors. Return-based predictors are particularly suited for this analysis due to their well-defined nature, long history, and systematic applicability both before and after publication. Specifically, we analyze lagged returns over short-term horizons (one to twelve months) to capture reversal and momentum effects. To detect seasonality patterns, we examine lagged returns at intervals of 24, 36, 48, and 60 months. Additionally, cumulative returns over months 13–60 are analyzed to identify long-term reversal effects.[25]

Figure 9 plots the estimated time-varying coefficients of return-based predictors from 1931 to 2023. The long-term reversal effect, as documented by De Bondt and Thaler (1985), is prominently visible in the `ltr` predictor during the early 20th century but weakens substantially over time. By the late 20th century, this effect became volatile and exhibited occasional opposite signs. Similarly, the 2–12-month momentum premium identified by Jegadeesh and Titman (1993) undergoes notable evolution. Its profile shifts structurally in subsequent decades, with the effect concentrating in the 7–12-month window and attenuating at shorter lags, consistent with the findings in Novy-Marx (2012). In contrast, the one-month reversal documented by Jegadeesh

[25]Following Marrow and Nagel (2024), we exclude firms with market capitalization below the 20th percentile of the prior month-end.

Figure 9: Time-Varying Coefficients of Return-Based Predictors

Note: This figure depicts dynamic premia tied to return-based anomalies, estimated via dynamic coefficients from Eq. (1). Predictors include lagged returns over 1-12, 24, 36, 48, and 60 months. `ltr` captures long-term reversal effects based on cumulative returns from months 13-60. The color bar reflects predictor magnitudes and thresholds, emphasizing their role in explaining return anomalies.



(1990) demonstrates remarkable persistence throughout the sample period. Despite minor variations in strength, its negative coefficient remains consistently strong across decades. The results indicate that the seasonality effect at the 24-, 36-, 48-, and 60-month lags, as identified by Heston and Sadka (2008), exhibits a consistent pattern with no changes. Overall, the results emphasize that, while anomalies may appear ex ante before publication, their predictive power often diminishes after academic discovery.

From a technical trading perspective, time-varying factor premia are reflected by observable trends in factor returns. TV-PR formalizes this concept by incorporating investors' prior beliefs about the smoothness of these trends and the frequency of market regime changes. A higher penalty reflects an assumption of smooth, long-term trends, producing piecewise-linear trend lines that highlight stable market regimes with major structural breaks—similar to technical traders using long-term moving averages. Conversely, a lower penalty captures more volatile trends, enabling the model to identify subtle, short-term shifts in market dynamics.

This approach translates the varying time horizons and trend-smoothness assumptions inherent in technical analysis into a disciplined statistical model. Additionally, by applying a second Lasso penalty for factor selection, the framework accounts for the transient relevance of factors, allowing predictive signals to adapt dynamically. This provides a robust method for systematically identifying regime-dependent momentum trends and structural breaks based on specified investor beliefs about market dynamics. TV-PR offers an alternative to Bayesian shrinkage methods, such as those proposed by Marrow and Nagel (2024), by integrating investor assumptions directly into the model's structure.

5 The Economic Value of Time-Varying Premia

Previous results suggest that the estimated factor premia exhibit economically meaningful patterns in both sign and magnitude. The piecewise-constant nature of our estimates provides a framework for predicting OOS risk premia, assuming that γ_t and γ_{t+1} remain unchanged within a given regime. To evaluate the investment performance of these estimates, we implement *factor timing* and *factor sorting* strategies, both based on the sign and magnitude of the estimated premia. These strategies assume that regime-dependent momentum in factor premia persists over time. By detecting structural changes with minimal delay, our approach aims to capture the profitability associated with these dynamics. Moreover, we also consider a forecast-implied investment strategy based on $\hat{r}_{i,t+1}$ from TV–PR to assess the economic value of TV–PR's predictive ability.

5.1 Factor Timing and Sorting Strategies

This section shows that the most recent trend estimates of factor premia can be used as trading signals for future factor returns, enabling the formation of various factor timing and sorting strategies with superior investment performance. These investment strategy results show that the break-and-trend modeling is relevant with potential economic value. They also show that trend estimates are not just fitting factor realization shocks that happen to align in the sample. Instead, it suggests TV-PR

recovers underlying factor premia dynamics that persist into the future. It further attests to TV-PR's ability to detect trends in near-real time. Otherwise, if break detection is too "slow" the detected signal would become obsolete for predicting future returns.

Specifically, Eq. (9) provides a basis for a timing strategy: invest in the slope factor based on the sign of the estimated $\hat{\gamma}_t$. Previous evidence suggests that these estimates can detect structural breaks and capture trends in factor returns, making them valuable for generating trading signals.

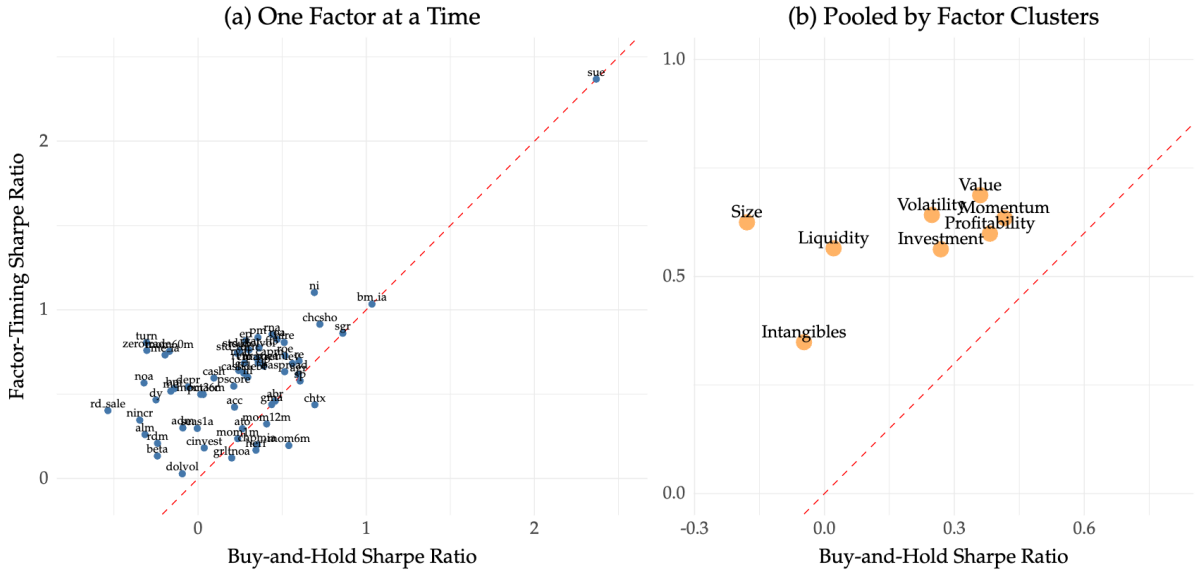
The trading strategies are constructed as follows. The investment target is the slope-factor, as defined in Eq. (9), which is treated as a tradable portfolio. For every month t, TV-PR is estimated in a look-back window that ends at month t. The most recent trend estimate is $\hat{\gamma}_{t-1}$ (which is the fitted expectation for the last observed month t). We take the assumption that there is no break between month t and the future month $t+1$: $\hat{\gamma}_{t-1} = \hat{\gamma}_t$. Therefore, we have the trading signal for the factor portfolio, $\hat{f}_{t+1}$.

To implement the *factor timing* strategy, we take a long position in the slope factor ($\hat{f}_{t+1}$) when the estimated time-varying risk premium for a given characteristic is positive ($\hat{\gamma}_t > 0$), and a short position when it is negative. As a benchmark, we calculate the buy-and-hold cross-sectional factor return as $(Z_t'Z_t)^{-1}Z_t'r_{t+1} \times \text{sgn}(\text{chars})$, where $\text{sgn}(\text{chars})$ equals 1 or -1, depending on the trading direction specified in the literature for each characteristic. In the first factor timing experiment, we consider each of the 61 factors individually. We use a ten-year rolling window of historical data and applies the latest estimate of γ_t to forecast next-month returns. The training sample spans ten years, and model estimates are updated monthly. The OOS testing period covers 120 months from 2014 to 2023.

The timing strategy demonstrates superior performance compared to the buy-and-hold benchmark across most characteristics, as indicated by the majority of points lying above the 45-degree line in Panel (a) of Figure 10. Panel (b) categorizes the slope factors into eight groups, revealing that firms within the size category achieve the highest returns under the factor-timing strategy. Improvements in Sharpe ratios

Figure 10: Single Factor Timing Performance on Time-Varying Premia

Note: This figure displays annualized Sharpe ratios for out-of-sample factor timing strategies versus buy-and-hold across 61 slope factors. Timing strategies, rebalanced monthly, adjust exposures based on the sign of estimated time-varying risk premia. For comparison, we report average Sharpe ratios across eight factor categories.



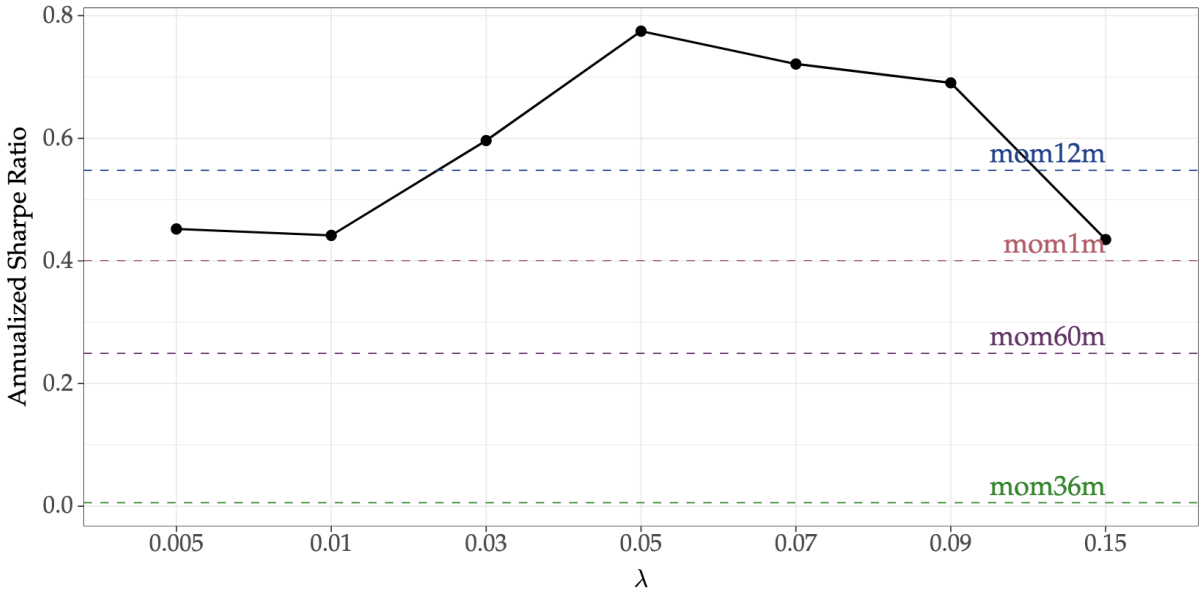
vary across characteristics, reflecting heterogeneity in factor premia. Liquidity-related characteristics, such as share turnover (`turn`) and the number of zero-trading days (3-month rolling, `zerotrade`), exhibit the largest gains. This finding highlights that cross-sectional factors tied to liquidity benefit substantially from timing strategies. Only a small subset of characteristics shows performance that is similar to or slightly below the buy-and-hold benchmark. Overall, the results show that the TV-PR risk-premia estimates have clear economic meaning and confirm that the identified trends and breaks reflect genuine shifts in the underlying factor-premia process.

Next, we use both the sign and the magnitude of the time-varying risk premia estimates to implement a *factor sorting* strategy. Specifically, we construct the factor-sorting portfolio by ranking the values of $\hat{\gamma}_t$ at each month. The strategy takes long positions in the five slope factors with the highest estimated premia and short positions in the five with the lowest estimated premia.

The performance of factor sorting strategies also depends on the frequency of parameter estimation. We present the investment performance at various regularization

Figure 11: Factor Sorting Strategy Performance based on Time-Varying Premia

Note: This figure reports annualized Sharpe ratios for out-of-sample factor rotation strategies. We construct long-short portfolios using 61 slope factors, estimate time-varying premia, and rebalance them monthly. Strategy performance is analyzed under varying levels of fused penalty regularization and benchmarked against long-short portfolios based on factor momentum constructed from cumulative past factor returns.



levels. In Figure 11, the optimal parameter achieves the highest Sharpe ratio of 0.78, demonstrating a balance between flexibility and performance. More importantly, the investment performance confirms the economic meaning of the identified breaks and trends in factor premia. The Sharpe ratio remains relatively stable across the hyperparameter spectrum, exhibiting a hump-shaped pattern, indicating that the strategy is robust to different smoothing choices.

The horizontal lines in the figure represent the annualized Sharpe ratios of factor momentum strategies constructed using the same bucket sizes as previously described. The sorting signal is the cumulative past return of each factor, defined as $\sum_{s=t-h}^{t} \hat{f}_{s+1}$, where h denotes the momentum window. We consider momentum signals based on return horizons of 1, 12, 36, and 60 months. None of these strategies surpasses TV-PR, reassuring the importance of piecewise-constant patterns in estimating risk premia. Momentum strategies, by contrast, depend on an ad hoc selection of window length, which lacks guidance for optimal specification. TV-PR, on the other hand, identifies piecewise-constant patterns in risk premia and consistently out-

performs both over-smoothed (e.g., 1-month factor momentum) and under-smoothed (e.g., 60-month factor momentum) alternatives, demonstrating strong robustness and adaptability.

5.2 Forecast-Implied Portfolio Investment Strategy

To examine whether explicitly modeling time variation can yield consistent and economically significant OOS investment performance, we construct investment portfolios based on predicted returns ($\hat{r}_{i,t+1}$) as signals for portfolio allocation in subsequent periods. The model is re-estimated annually using a rolling window of the past ten years. By explicitly accounting for structural shifts in market dynamics, our methodology adapts to evolving conditions, offering a practical framework for improving asset allocation decisions. Meanwhile, we employ period-by-period regression models as a benchmark to capture maximum flexibility in our analysis.

Forecast-implied portfolios take long positions in stocks in the top decile of predicted returns and short positions in stocks in the bottom decile. Both equal-weighted and value-weighted approaches are employed to allocate positions.

Table 4: Out-of-Sample Forecast-Implied Investment Performance

Note: This table presents the OOS performance of forecast-implied portfolios, constructed using equal-weighted and value-weighted long-short strategies. The analysis employs rolling regressions, including rolling OLS, rolling LASSO, and TV-PR methods. Each panel reports five key performance metrics: monthly average return (Avg, %), standard deviation (Std, %), annualized Sharpe ratio (SR), Jensen's alpha (%), and maximum drawdown (MDD, %).

	Avg	Std	SR	Alpha	MDD
	Panel A: Equal-Weighted				
Rolling OLS	2.61	3.87	2.34	2.87^{***}	16.54
Rolling LASSO	2.04	3.78	1.87	2.31^{***}	17.67
TV-PR	2.46	3.40	2.50	2.60^{***}	6.41
	Panel B: Value-Weighted				
Rolling OLS	0.22	4.64	0.16	0.57	52.06
Rolling LASSO	0.05	4.19	0.04	0.44	47.75
TV-PR	0.43	3.41	0.44	0.70^{**}	29.06

Table 4 summarizes the OOS performance of forecast-implied portfolio strategies. In Panel A, the equal-weighted long-short portfolio based on TV-PR demonstrates strong performance, with an annualized Sharpe ratio of 2.5 and a monthly alpha of

approximately 2.6%. TV-PR also exhibits a significantly lower maximum drawdown compared to the period-by-period regression approach. Similarly, Panel B shows that the value-weighted long-short portfolio constructed using TV-PR achieves a higher Sharpe ratio and a statistically significant alpha, while maintaining a considerably smaller maximum drawdown than the period-by-period regression method. These results underscore the model's predictive accuracy and economic value.

The forecast-implied performance demonstrates that introducing structural breaks at appropriate frequencies enhances the model's ability to capture dynamic changes. While period-by-period regression updates effectively adapt to frequent structural changes, they lead to increased variance, as evidenced by a lower Sharpe ratio and a significant maximum drawdown. By comparison, the TV-PR model achieves an optimal balance in the bias-variance trade-off while maintaining an interpretable factor structure that enhances investment performance.

6 Time-Varying Factor Composition of the SDF

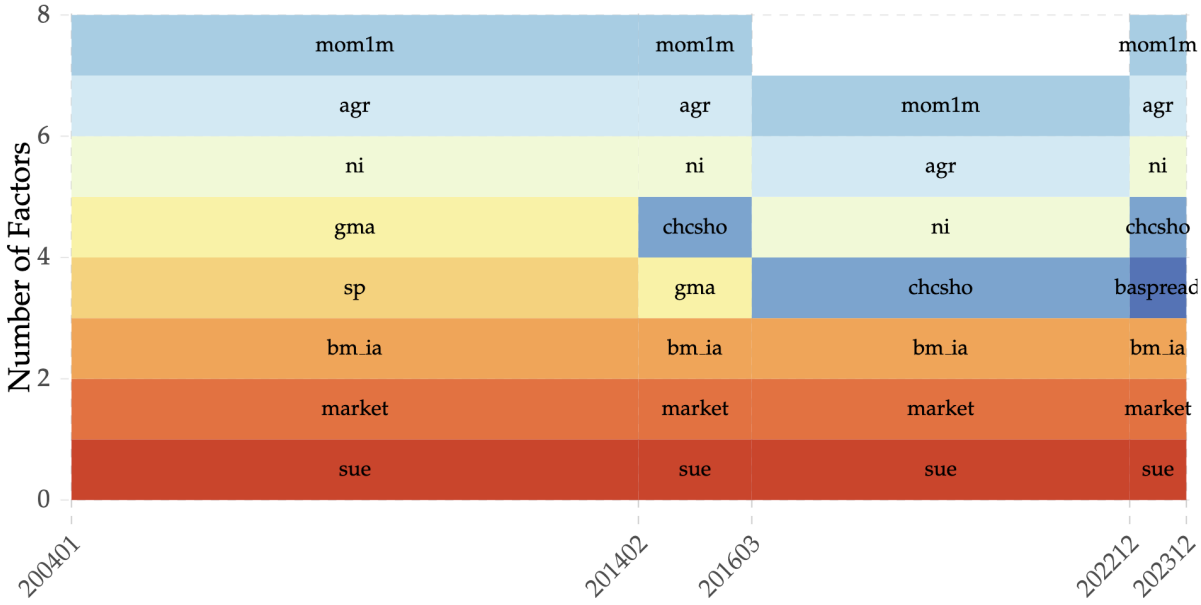
One important question in empirical asset pricing is determining when and which factors are useful for pricing other assets, or, equivalently, when these factors are relevant to asset-pricing factor models. To address this question, we analyze the dynamic structure of slope-factor risk premia and compare time-varying and time-invariant factor models.

Based on the results above, TV-PR provides a framework for detecting changes in factor model specifications across distinct economic regimes ($\hat{\theta}_n^j \neq 0$). The specification of the factor model across different regimes is crucial for understanding the relationship between SDF composition and asset returns. Figure 12 illustrates the regime-specific factor model specification. The model initially included eight factors in early 2004 and maintained that count through 2014. By early 2016, the number of factors had decreased slightly to seven, remained stable through late 2022, and then returned to eight.

The composition of pricing factors evolves significantly over time, even as their

Figure 12: Regime-Specific Factor Model Specification

Note: The figure illustrates the number of selected factors used in asset pricing models within each regime identified by structural breaks. These factors are ranked by their estimated risk premia in descending order, starting from the bottom of the plot.



overall number remains relatively stable. A core set of factors, including `sue`, `market`, `bm_ia`, and `mom1m`, consistently demonstrates pricing power throughout the two-decade span. In contrast, other factors exhibit transience. For example, `sp`, which was relevant in 2004, is replaced by `chcsho` starting in 2014. The emergence of the `baspread` factor in late 2022 further highlights the dynamic nature of firm characteristics. The results suggest that the attributes influencing asset pricing change over time, reflecting shifts in market conditions and investor preferences.

To evaluate the relative performance of time-varying versus time-invariant factor model specifications, we employ differences in squared Sharpe ratios (SR^2), following the non-nested model comparison framework proposed by Barillas et al. (2020). This methodology enables a rigorous evaluation of how effectively each specification captures evolving market dynamics. The findings underscore the effectiveness of time-varying models in adapting to shifts in market conditions.

The performance of factor models varies across distinct market regimes. As shown in Table 5, during the first regime (2004:01–2014:01), the time-varying specification de-

Table 5: Regime Specific Factor Model Comparison

Note: This table reports squared differences in Sharpe ratios between time-varying and static factor models across market regimes ($yymm$). Static models include the CAPM, Fama-French 3-factor, 5-factor, and 6-factor (FF3, FF5, FF6), Hou et al. (2021) 5-factor ($q5$), Daniel et al. (2020) 3-factor (DHS3), and Barillas and Shanken (2018) 6-factor (BS6) models. p-values are reported in parentheses.

	0401-1401	1402-1602	1603-2211	2212-2312
CAPM	1.48	-0.10	-0.09	-0.18
	(0.00)	(0.54)	(0.17)	(0.40)
FF3	1.49	-0.07	-0.07	-0.05
	(0.00)	(0.68)	(0.29)	(0.84)
FF5	1.41	-0.13	-0.08	0.04
	(0.00)	(0.60)	(0.30)	(0.87)
FF6	1.42	-0.15	-0.08	0.13
	(0.00)	(0.51)	(0.33)	(0.57)
$q5$	1.40	-0.55	-0.12	-0.04
	(0.00)	(0.07)	(0.21)	(0.88)
DHS3	1.45	-0.26	-0.10	-0.20
	(0.00)	(0.26)	(0.21)	(0.56)
BS6	1.45	-0.23	-0.08	0.12
	(0.00)	(0.32)	(0.34)	(0.59)

livers both statistically and economically superior performance relative to benchmark models. Specifically, the results show substantial SR^2 improvements, including a 1.48 increase over the CAPM, as well as similarly pronounced enhancements relative to FF3, FF5, and other models. These findings are supported by p-values close to zero, indicating statistical significance.

In contrast, the incremental value added by the time-varying model diminishes considerably in post-2014 regimes. Across most specifications and subperiods from 2014 onward, SR^2 differences are economically modest and statistically insignificant. Furthermore, distinctions among time-invariant models become less pronounced in later regimes, highlighting the increased difficulty in differentiating between competing factor models during these periods. This observation applies not only to comparisons between time-varying and time-invariant models but also to distinctions among the time-invariant models themselves.

7 Conclusion

This paper develops a regime-based predictive regression framework that accommodates both structural breaks and time-varying sparsity in the relationship between

firm characteristics and expected returns. By combining panel information with total variation and Lasso penalties, the proposed TV-PR estimator delivers piecewise constant risk premia and dynamically selects predictors without relying on restrictive assumptions about coefficient evolution. This design allows the model to adjust flexibly to changes in underlying economic conditions while maintaining parsimony.

Our theoretical results establish the large-sample properties of the estimator under both known and unknown structural break settings, providing a foundation for inference on dynamic relationships. Empirically, the framework reveals economically meaningful shifts in factor premia, underscoring the transient nature of return predictors. Factor-timing and sorting strategies based on TV-PR's estimated premia outperform both the buy-and-hold benchmark and the factor momentum benchmark. Time-varying SDF composition identified by TV-PR yields a better dynamic factor model than time-invariant ones.

By incorporating regime-dependent dynamics and time-varying sparsity, TV-PR addresses key complexities in return prediction. These findings contribute to the literature on dynamic asset pricing models and offer practical insights for portfolio management. TV-PR treats firm characteristics as observable factor loadings, while focusing on the dynamics of factor premia. A joint estimation of time-varying loadings and premia is an interesting direction for future research.

References

An, Y., Y. Su, and C. Wang (2025). Quantity, risk, and return. Technical report, Johns Hopkins University.

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

Avramov, D., G. Feng, J. He, and S. Xiao (2025). Schrodinger's sparsity in the cross section of stock returns. Technical report, City University of Hong Kong.

Bandi, F. M. and Y. Su (2025). Conditional spectral methods. *Journal of Econometrics 248*, 105863.

Bandi, F. M. and A. Tamoni (2022). Spectral financial econometrics. *Econometric Theory 38*(6), 1175–1220.

Barillas, F., R. Kan, C. Robotti, and J. Shanken (2020). Model comparison with sharpe ratios. *Journal of Financial and Quantitative Analysis 55*(6), 1840–1874.

Barillas, F. and J. Shanken (2018). Comparing asset pricing models. *Journal of Finance 73*(2), 715–754.

Bryzgalova, S., M. Pelger, and J. Zhu (2020). Forest through the trees: Building cross-sections of stock returns. Technical report, London Business School.

Chen, Q., N. Roussanov, and X. Wang (2025). Semiparametric conditional factor models in asset pricing. Technical report, The Chinese University of Hong Kong, Shenzhen.

Chib, S., Y. C. Lin, K. Pukthuanthong, and X. Zeng (2023). Slope factors outperform: evidence from a large comparative study. Technical report, Washington University in St. Louis.

Chinco, A., A. D. Clark-Joseph, and M. Ye (2019). Sparse signals in the cross-section of returns. *The Journal of Finance 74*(1), 449–492.

Cong, L. W., G. Feng, J. He, and Y. Wang (2024). Mosaics of predictability. Technical report, City University of Hong Kong.

Dangl, T. and M. Halling (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics 106*(1), 157–181.

Daniel, K., D. Hirshleifer, and L. Sun (2020). Short-and long-horizon behavioral factors. *Review of Financial Studies 33*(4), 1673–1736.

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

Daniel, K. and S. Titman (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance 52*(1), 1–33.

De Bondt, W. F. and R. Thaler (1985). Does the stock market overreact? *Journal of Finance 40*(3), 793–805.

Ehsani, S. and J. T. Linnainmaa (2022). Factor momentum and the momentum factor. *Journal of Finance 77*(3), 1877–1919.

Fama, E. (1976). Foundations of finance: Portfolio decisions and securities prices.

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

Fama, E. F. and K. R. French (2020). Comparing cross-section and time-series factor models. *Review of Financial Studies 33*(5), 1891–1926.

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

Fan, J. and R. Li (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association 96*(456), 1348–1360.

Farmer, L. E., L. Schmidt, and A. Timmermann (2023). Pockets of predictability. *Journal of Finance 78*(3), 1279–1341.

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

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

Gagliardini, P., E. Ossola, and O. Scaillet (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica 84*(3), 985–1046.

Gagliardini, P., F. Trojani, and G. Urga (2005). Robust GMM tests for structural breaks. *Journal of Econometrics 129*(1-2), 139–182.

Giglio, S. and D. Xiu (2021). Asset pricing with omitted factors. *Journal of Political Economy 129*(7), 1947–1990.

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

Harvey, C. R., Y. Liu, and H. Zhu (2016). ... and the cross-section of expected returns. *Review of Financial Studies 29*(1), 5–68.

Heston, S. L. and R. Sadka (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics 87*(2), 418–445.

Hou, K., H. Mo, C. Xue, and L. Zhang (2021). An augmented q-factor model with expected growth. *Review of Finance 25*(1), 1–41.

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

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

Kelly, B. T., T. J. Moskowitz, and S. Pruitt (2021). Understanding momentum and reversal. *Journal of Financial Economics 140*(3), 726–743.

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

Kozak, S. and S. Nagel (2023). When do cross-sectional asset pricing factors span the stochastic discount factor? Technical report, National Bureau of Economic Research.

Li, D., J. Qian, and L. Su (2016). Panel data models with interactive fixed effects and multiple structural breaks. *Journal of the American Statistical Association 111*(516), 1804–1819.

Marrow, B. and S. Nagel (2024). Real-time discovery and tracking of return-based anomalies. Technical report, National Bureau of Economic Research.

McLean, R. D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *Journal of Finance 71*(1), 5–32.

Neely, C. J., D. E. Rapach, J. Tu, and G. Zhou (2014). Forecasting the equity risk premium: the role of technical indicators. *Management science 60*(7), 1772–1791.

Novy-Marx, R. (2012). Is momentum really momentum? *Journal of Financial Economics 103*(3), 429–453.

Pástor, L. and R. F. Stambaugh (2001). The equity premium and structural breaks. *Journal of Finance 56*(4), 1207–1239.

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

Rapach, D. E., J. K. Strauss, and G. Zhou (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies 23*(2), 821–862.

Rosenberg, B. (1974). Extra-market components of covariance in security returns. *Journal of Financial and Quantitative Analysis 9*(2), 263–274.

Rudin, L. I., S. Osher, and E. Fatemi (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena 60*(1-4), 259–268.

Smith, S. C. and A. Timmermann (2021). Break risk. *Review of Financial Studies 34*(4), 2045–2100.

Smith, S. C. and A. Timmermann (2022). Have risk premia vanished? *Journal of Financial Economics 145*(2), 553–576.

Smith, S. C., A. Timmermann, and Y. Zhu (2019). Variable selection in panel models with breaks. *Journal of Econometrics 212*(1), 323–344.

Tang, X., X. Tang, and G. Zhou (2022). Heterogeneous responses in financial markets: Insights from machine learning. Technical report, University of Texas at Dallas.

Tibshirani, R., M. Saunders, S. Rosset, J. Zhu, and K. Knight (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology 67*(1), 91–108.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Annals of Statistics 38*(2), 894–942.

Zhu, Y. and G. Zhou (2009). Technical analysis: An asset allocation perspective on the use of moving averages. *Journal of Financial Economics 92*(3), 519–544.

Appendices

A Additional Results

Figure A.1: Factor Realizations from One Simulation

Note: This figure illustrates the realizations of the three factors ($f_{t+1}^{[p]} = \gamma_t^{[p]} + \nu_{t+1}^{[p]},\ p = 1, 2, 3$) from a single simulation ($N = 1000$, $T = 240$).

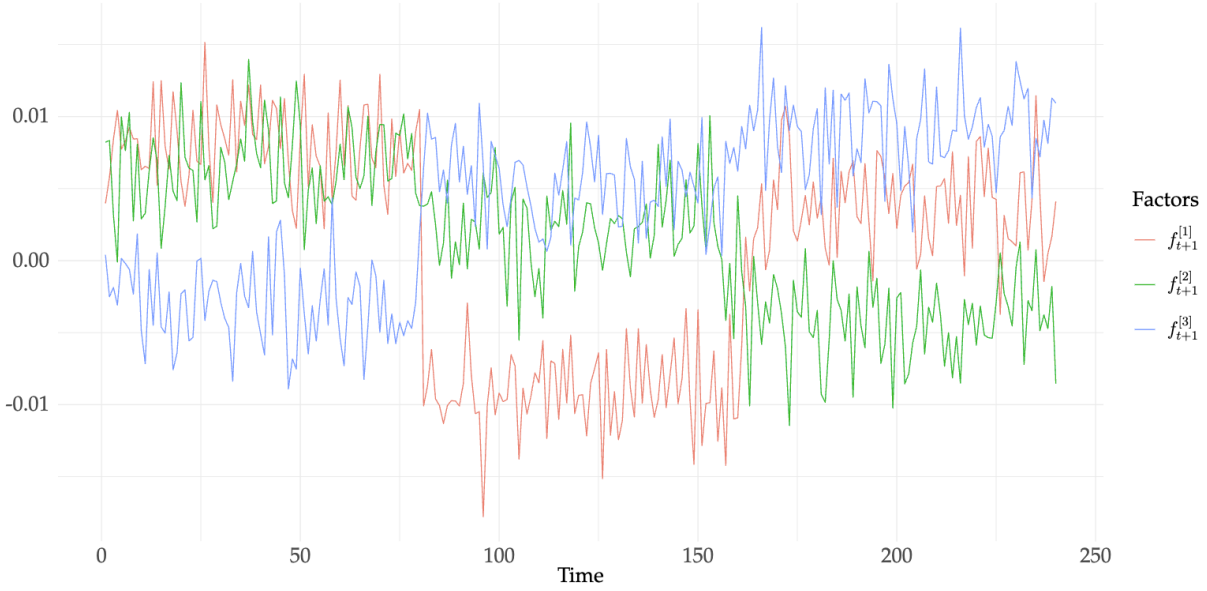


Figure A.1 displays the realizations of all factors across regimes from a single simulation draw. The first factor, $f_{t+1}^{[1]}$, exhibits a decline in expected returns, shifting from positive in the first regime to negative in the second, before returning to positive in the third regime. The second factor, $f_{t+1}^{[2]}$, shows a continuous decline, whereas the third factor, $f_{t+1}^{[3]}$, steadily increases from the first to the final regime. Factor innovations introduce noise into risk premia estimation; thus, to obtain clean estimates, researchers need a methodology to average out these innovations within each regime.

Figure A.2: Sharpe Ratios of Factor-Timing Strategies with Different Hyperparameters

Note: This figure reports annualized Sharpe ratios for the factor-timing strategy versus buy-and-hold strategy using extreme hyperparameter values (largest and smallest) over 2004–2023. The dot denotes the performance of the optimal factor-timing strategy, while the 45-degree line represents the outcome associated with the smallest or largest value of λ. The results indicate that the two extremes – frequently adjusting premia each period and maintaining constant premia – underperform the factor-timing strategy. This finding suggests that intermediate points between these extremes may deliver superior investment performance, as documented in the main results.

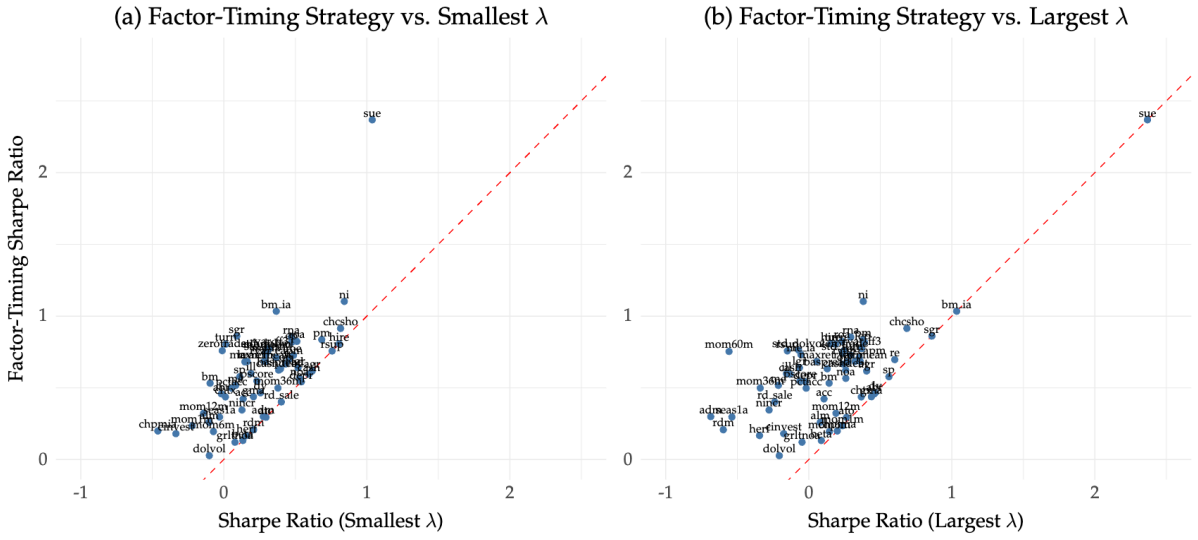
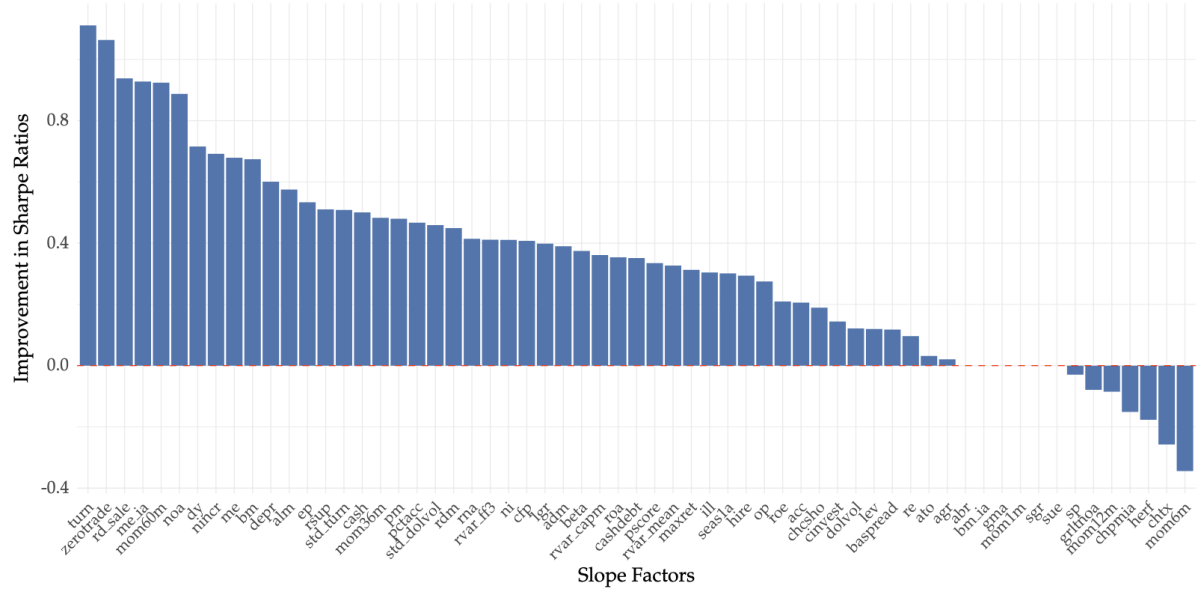


Figure A.3: Investment Improvement of Factor Timing Strategies

Note: This figure reports the improvement in annualized Sharpe ratios of the out-of-sample factor-timing strategy relative to the buy-and-hold benchmark from 2004 to 2023. The factor-timing approach leverages the sign of estimated time-varying risk premia, with monthly portfolio rebalancing.



B Predictor Descriptions

Table A.1: **Equity Characteristics**

No.	Acronym	Description	Category
1	abr	Cumulative abnormal returns around earnings announcement dates	Momentum
2	acc	Operating Accruals	Investment
3	adm	Advertising Expense-to-market	Intangibles
4	agr	Asset growth	Investment
5	alm	Quarterly Asset Liquidity	Intangibles
6	ato	Asset Turnover	Profitability
7	baspread	Bid-ask spread rolling 3m	Liquidity
8	beta	Beta rolling 3m	Volatility
9	bm	Book-to-market equity	Value
10	bm_ia	Industry-adjusted book to market	Value
11	cash	Cash holdings	Value
12	cashdebt	Cash flow to debt	Value
13	cfp	Cash flow to price ratio	Value
14	chcsho	Change in shares outstanding	Investment
15	chpmia	Industry-adjusted change in profit margin	Profitability
16	chtx	Change in tax expense	Momentum
17	cinvest	Corporate investment	Investment
18	depr	Depreciation / PP&E	Momentum
19	dolvol	Dollar trading volume	Liquidity
20	dy	Dividend yield	Value
21	ep	Earnings-to-price	Value
22	gma	Gross profitability	Profitability
23	grltnoa	Growth in long-term net operating assets	Investment
24	herf	Industry sales concentration	Intangibles
25	hire	Employee growth rate	Intangibles
26	ill	Illiquidity rolling 3m	Liquidity
27	lev	Leverage	Value
28	lgr	Growth in long-term debt	Investment
29	maxret	Maximum daily returns rolling 3m	Volatility
30	me	the market equity	Size
31	me_ia	Industry-adjusted size	Size
32	mom12m	Momentum rolling 12m	Momentum
33	mom1m	Momentum	Momentum
34	mom36m	Momentum rolling 36m	Momentum

Table A.1: Equity Characteristics (Continued)

No.	Acronym	Description	Category
35	mom60m	Momentum rolling 60m	Momentum
36	mom6m	Momentum rolling 6m	Momentum
37	ni	Net Stock Issues	Investment
38	nincr	Number of earnings increases	Momentum
39	noa	(Changes in) Net Operating Assets	Investment
40	op	Operating profitability	Profitability
41	pctacc	Percent operating accruals	Investment
42	pm	profit margin	Profitability
43	pscore	Performance Score	Profitability
44	rd_sale	R&D to sales	Intangibles
45	rdm	R&D Expense-to-market	Intangibles
46	re	Revisions in analysts earnings forecasts	Momentum
47	rna	Quarterly Return on Net Operating Assets	
48	roa	Return on assets	Profitability
49	roe	Return on equity	Profitability
50	rsup	Revenue surprise	Momentum
51	rvar_capm	Residual variance - CAPM rolling 3m	Volatility
52	rvar_ff3	Residual variance - ff3 rolling 3m	Volatility
53	seas1a	Seasonality	Intangibles
54	sgr	Sales growth	Value
55	sp	Sales-to-price	Value
56	std_dolvol	Std. of dollar trading volume rolling 3m	Volatility
57	std_turn	Std. of Share turnover rolling 3m	Volatility
58	sue	Unexpected quarterly earnings	Momentum
59	rvar_mean	return variance rolling 3m	Volatility
60	turn	Shares turnover	Liquidity
61	zerotrade	Number of zero-trading days rolling 3m	Liquidity

C Technical Details and Proofs

C.1 Assumptions

We will first discuss the asymptotic guarantees of our TV-PR estimator, considering both scenarios: with and without prior knowledge of structural breaks. We introduce some additional notations. Let $\lambda_{\min}(D)$, $\lambda_i(D)$ and $\lambda_{\max}(D)$ denote the minimum, the ith, and the maximum eigenvalue of a $K \times K$ matrix D. For a $K_1 \times K_2$ matrix D, denote $\|D\|_{2,\infty} = \max_{1\le k\le K_1} \|D_{k,}\|$ with $\|D_{k,}\| = \sqrt{\sum_{1\le j\le K_2} D_{kj}^2}$. Let $\mathbb{S}_G = \{\Gamma \in \mathbb{R}^{T\times p} : \gamma_{j,t} = \gamma_{j,\tau}, \text{for}\, t,\tau \in G_n^j, 1 \le n \le B_j, j = 1,\dots,p\}$. Denote $\epsilon_{t+1} = Z_t'\nu_{t+1} + u_{t+1}$ and define $\boldsymbol{\epsilon} = (\epsilon_2',\dots,\epsilon_{T+1}')$, which is a $NT \times 1$ vector. We can partition the associated regime duration vector into two subvectors $|G| = \{|G_1|,\dots,|G_N|\} = (|G_{\mathbb{A}}|', G_{\mathbb{Z}}|')'$. Let $D_{\mathbb{A}} = diag(|G_{\mathbb{A},1}|^{1/2},\dots,|G_{\mathbb{A},q}|^{1/2})$. We denote $\mathbf{Z}_{\mathbb{A},i}$ as the ith row of $\mathbf{Z}_{\mathbb{A}}$ and $\mathbf{Z}_{\mathbb{A},\cdot,j}$ as the jth column of $\mathbf{Z}_{\mathbb{A}}$. Similar notation applies to $\mathbf{Z}_{\mathbb{Z},i}$ and $\mathbf{Z}_{\mathbb{Z},\cdot,k}$ with $i = 1,\dots,NT$, $j = 1,\dots,q$, and $k = 1,\dots,B-q$. Let $\omega_T = \min_{1\le j\le p, B_j>1} \min_{n_1\neq n_2} |\theta_{on_1}^j - \theta_{on_2}^j|$, denoting the minimum difference of risk premia across different regimes.

Assumption 1. *The penalty function $P_\lambda(x)$ is symmetric, nondecreasing, and concave in $x \in [0,\infty)$, and $P_\lambda(0) = 0$. For some $a > 0$, $P_\lambda(x)$ is a constant for $x > a\lambda$. $P_\lambda'(x)$ is increasing in $\lambda \in [0,\infty)$ and $P_\lambda'(x)$ is nonincreasing in $x \in [0,\infty)$ with $\lambda^{-1}P_\lambda'(0+) \ge a_1$ and $P_\lambda'(x) \ge a_1\lambda$ for $x \in (0, a_2\lambda]$. The same requirements hold for $P_\eta(x)$.*[a]

Assumption 2. *Assume $\{(z_{i,1}',\dots,z_{i,T}',u_i')\}_{i=1}^N$ is an independent sequence. Suppose that for all $i = 1,\dots,N$, $j = 1,\dots,p$, and $t = 1,\dots,T$, $E[u_{i,t}|r_{i,t-1},\dots,r_{i,1},z_{i,t},\dots,z_{i,1}] = 0$ and $E[\nu_t|r_{i,t-1},\dots,r_{i,1},z_{i,t},\dots,z_{i,1}] = 0$.*

Assumption 3. *Suppose there exist constants $C_1, C_2 > 0$ such that, for every $\kappa \ge 0$, $t = 1,\dots,T$, and $i = 1,\dots,N$, $\sup_{\|a\|\le 1} P(|a'\nu_t| \ge \kappa) \le C_1 e^{-C_2\kappa^2}$ and $\sup_{\|b\|\le 1} P(|b'u_t| \ge \kappa) \le C_1 e^{-C_2\kappa^2}$. Moreover, $z_{i,t}$ are uniformly subgaussian that $\sup_{\|e\|\le 1} P(|e'z_{i,t}| > \kappa) \le C_1 e^{-C_2\kappa^2}$.*

Assumption 4. *Suppose (i) the eigenvalues of the $q \times q$ matrix $E[\frac{1}{N}\mathbf{D}_{\mathbb{A}}^{-1}\mathbf{Z}_{\mathbb{A}}'\mathbf{Z}_{\mathbb{A}}\mathbf{D}_{\mathbb{A}}^{-1}]$ are bounded below by $C_3 > 0$ and bounded above by $C_4 > 0$. (ii) There exist positive constants*

[a] For SCAD, $a_1 = a_2 = 1$, $P_\lambda'(x) = \text{sign}(x)\lambda$ if $|x| \le \lambda$, and $\text{sign}(x)\frac{(a\lambda-|x|)_+}{a-1}$ otherwise. For MCP in Zhang (2010), $P_\lambda'(x) = \text{sign}(x)\,[\lambda - |x|/a]$ if $|x| \le a\lambda$, and 0 otherwise, where $a_1 = 1 - 1/a$ and $a_2 = 1$.

C_5, C_6, and C_7 such that for all $i = 1, \ldots, NT$, and $j = 1, \ldots, q$, $\sup_i \|\mathbf{Z}_{\mathbb{A},i}\|^2 \leq C_5 N|G_i|$, $\sup_j \|\mathbf{Z}_{\mathbb{Z},\cdot,j}\|^2 \leq C_6 q$, and $\|\mathbf{Z}'_{\mathbb{Z}}\mathbf{Z}_{\mathbb{A}}\|_{2,\infty} = C_7 NT$.

Assumption 5. *Suppose (i) $q\log^3(q \vee N) = o(N|G_{\min}|)$ and $qp^2\log^4(p \vee q \vee |G_{\max}|) = o(N)$. (ii) $\log(B) = O((N|G_{\min}|)^{\alpha})$ for some $\alpha \in (0, 1/2)$. (iii) $P'_\eta(0+) \geq \max\{[\sqrt{\frac{q\log^3(q\vee N)}{N|G_{\min}|}} + \sqrt{\frac{qp^2\log^4(p\vee q\vee|G_{\max}|)}{N}}]/T|G_{\min}|, \sqrt{\log(N|G_{\min}|)/N|G_{\min}|^{1-\alpha}}/|G_{\min}|\}$. (iv) $\mathbf{N}_0 = \{\Theta \in \mathbb{R}^q : \|\Theta - \theta_{o\mathbb{A}}\| \leq d_T\}$, for $\theta \in \mathbf{N}_0$, $\max_n P''_\eta(\Theta_n) = o\big(|G_{\min}|/|G_{\max}|T\big)$. (v) $\lambda \geq c(\sqrt{q(\log(q \vee N))^3/NT} + \sqrt{qp^2\log^4(q \vee p)/NT} + 2\eta/|G_{\min}| + \eta|G_{\min}|)$ for some positive constant c.*

Assumption 6. *Suppose $d_T = \min_{1\leq n\leq q}|\theta_{\mathbb{A},n}|/2$ as half of the minimum signal with $d_T \gg \sqrt{\frac{q\log^3(q\vee N)}{N|G_{\min}|}} + \sqrt{\frac{qp^2\log^4(p\vee q\vee|G_{\max}|)}{N}}$. (i) $P'_\eta(d_T) = O(\sqrt{p^2(\log(q \vee p \vee |G_{\max}|)^4/NT^2)})$. (ii) $P'_\eta(d_T) = o\big(\sqrt{p^2\log^4(q \vee p \vee |G_{\max}|)/qNT^2}\big)$.*

Assumption 1 states formal requirements for penalty functions, enabling us to circumvent irrepresentable conditions and accommodate correlated candidate covariates. Assumption 2 imposes independence across individuals $i = 1, \ldots, N$, a common condition in panel data analysis.[b] We relax the assumption on temporal dependence, such as strong mixing conditions, widely used in the econometric literature, but at a cost of involving $\log^3(p \vee N)$, rather than $\log(p \vee N)$ in the penalty terms η and λ. Importantly, we do not assume identical error distributions across individuals or time. Additionally, conditional heteroskedasticity in the error terms is permitted. Assumption 3 assumes that both the error terms and predictors are uniformly subgaussian. The general applicability of this assumption in high-dimensional panel data regression models has been illustrated by Kock and Tang (2019). Assumption 4 pertains to the restricted eigenvalue condition discussed in Bickel et al. (2009), a standard assumption in the variable selection literature. This condition is naturally satisfied when Φ is positive definite. Assumptions 5 and 6(i) ensure estimation consistency by regulating the relative rates at which q, N, and $|G|$ diverge to infinity. Assumption 6(ii) further strengthens Assumption 6(i) to establish asymptotic normality. Similar assumptions are used in Ke et al. (2015) to characterize homogeneity among regressors.

[b] However, this assumption can be relaxed to allow for cross-sectional dependence by incorporating latent structures, as demonstrated in Su and Wang (2017).

Now, we elaborate the construction of a feasible consistent covariance estimator involved in Theomre 2 that

$$\tilde{\Omega}_{\mathbb{A}} = \tilde{\Phi}_{\mathbb{A}}^{-1}\tilde{\Sigma}_{\mathbb{A}}\tilde{\Phi}_{\mathbb{A}}^{-1}, \tag{A-1}$$

where we further explain the construction of the sample analogues $\tilde{\Phi}_{\mathbb{A}}$ and $\tilde{\Sigma}_{\mathbb{A}}$ below.

First, given the separation of $\mathbf{Z} = (\mathbf{Z}_{\mathbb{A}}, \mathbf{Z}_{\mathbb{Z}})$ according to $\theta = (\theta'_{\mathbb{A}}, \theta'_{\mathbb{Z}})'$, we denote $\Sigma_{\mathbb{A}} = \frac{1}{N} var[D_{\mathbb{A}}^{-1}\mathbf{Z}'_{\mathbb{A}}\boldsymbol{\epsilon}]$, where $\Sigma_{\mathbb{A},n'n'} = \frac{1}{N|G_n^j|}\sum_{i=1}^{N}\sum_{t\in G_n^j} E[(z_{i,t,j}\epsilon_{i,t+1})^2]$ and $\Sigma_{\mathbb{A},n'v'} = \frac{1}{N\sqrt{|G_n^j||G_v^w|}}\sum_{i=1}^{N}\sum_{t\in G_n^j\cap G_v^w} E[z_{i,t,j}z_{i,t,w}\epsilon_{i,t+1}^2]$, where $n' = \sum_{l=1}^{j-1} B_l + n$ and $v' = \sum_{l=1}^{w-1} B_l + v$ with $1 \leq n \leq B_j$ and $1 \leq v \leq B_w$ for $1 \leq j, w \leq p$. We construct a feasible sample counterpart of $\Sigma_{\mathbb{A}}$ as $\tilde{\Sigma}_{\mathbb{A}}$ such that $\tilde{\Sigma}_{\mathbb{A},n'n'} = \frac{1}{N|G_n^j|}\sum_{i=1}^{N}\sum_{t\in G_n^j}\tilde{\epsilon}_{i,t+1}^2 z_{i,t,j}^2$ and $\tilde{\Sigma}_{\mathbb{A},n'v'} = \frac{1}{N\sqrt{|G_n^j||G_v^w|}}\sum_{i=1}^{N}\sum_{t\in G_n^j\cap G_v^w}\tilde{\epsilon}_{i,t+1}^2 z_{i,t,j}z_{i,t,w}$, where $\tilde{\epsilon}_{i,t+1} = r_{i,t+1} - z'_{i,t}\tilde{\gamma}_t$, where $\tilde{\Gamma} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_T)$ such that $vec(\tilde{\Gamma}) = M\tilde{\theta}$ with $\tilde{\theta}$ that solves Eq. (7).

Given $\Phi_{\mathbb{A}} = \frac{1}{N}E[D_{\mathbb{A}}^{-1}\mathbf{Z}'_{\mathbb{A}}\mathbf{Z}_{\mathbb{A}}D_{\mathbb{A}}^{-1}]$, we consider a feasible sample counterpart of $\Phi_{\mathbb{A}}$ as $\tilde{\Phi}_{\mathbb{A}}$, where $\tilde{\Phi}_{\mathbb{A},n'v'} = 0$ if $G_n^j \cap G_v^w = \emptyset$, $\tilde{\Phi}_{\mathbb{A},n'n'} = \frac{1}{N|G_n^j|}\sum_{i=1}^{N}\sum_{t\in G_n^j} z_{i,t,j}^2$, and $\tilde{\Phi}_{\mathbb{A},n'v'} = \frac{1}{N\sqrt{|G_n^j||G_v^w|}}\sum_{i=1}^{N}\sum_{t\in G_n^j\cap G_v^w} z_{i,t,j}z_{i,t,w}$.