Internet Appendix to

Comparing factor models under regime switching

(not for publication)

Shamim Ahmed[*] Cesare Robotti[†] Daniel Tsvetanov[‡] Xiaoxia Ye[§¶]

JEL Classification: C51; C52; C53; G11; G12

Keywords: Regime switching; Asset pricing models; Model comparison; State uncertainty; Out-of-sample analysis

[*]University of Liverpool Management School, University of Liverpool, Liverpool L69 7ZH, United Kingdom; shamim.ahmed@liverpool.ac.uk.

[†]Corresponding author. Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom; cesare.robotti@wbs.ac.uk.

[‡]Norwich Business School, University of East Anglia, Norwich NR4 7TJ, United Kingdom; d.tsvetanov@uea.ac.uk.

[§]Nottingham University Business School, University of Nottingham, Nottingham NG8 1BB, United Kingdom; xiaoxia.ye@nottingham.ac.uk.

[¶]We thank Lukas Schmid (the department editor), an anonymous associate editor, two anonymous referees, and Ziwen Bu for useful comments and suggestions.

IA.1 Proofs

This section provides proofs of the theoretical results in Section 2 of the main paper, including the asymptotic distributions of the pairwise model-comparison tests. Before presenting the proofs, we first outline the necessary assumptions and establish several preliminary results that will be used throughout the derivations. Consistent with the notation in the main paper, throughout most of the expositions we condition on a generic realization $s_t = \ell$ and index regime-specific parameters by ℓ for notational convenience. However, in expressions where the time variation of the state variable is essential, we retain the notation s_t to emphasize its stochastic nature.

IA.1.1 Preliminaries

We require that the process $S = \{s_t\}_{t\in\mathbb{N}}$ satisfies the following assumption:

Assumption IA1. *The process S is a finite, homogeneous, irreducible, and aperiodic Markov chain, with one-step transition probability matrix denoted by $P := (p_{\iota\ell})$, where $p_{\iota\ell} = \mathcal{P}(s_t = \ell \mid s_{t-1} = \iota)$, for $\iota, \ell = 1, \ldots, L$. The unique stationary distribution for P, denoted by $\pi^\star := (\pi_\ell)$, exists and satisfies $\pi_\ell = \lim_{\tau\to\infty} p^\tau_{\iota\ell}$. Furthermore, $\mathcal{P}(s_t = \ell) = \pi_\ell$ for all ℓ, t.*

Remark. The assumption of irreducibility guarantees that all states have a nonzero probability of occurring in the steady state. The stationary distribution satisfies $P'\pi^\star = \pi^\star$ and P is such that $P\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is a conformable vector of ones.

The following lemma generalizes the results of Timmermann (2000) and shows that the unconditional moments of returns are not simple averages of their regime-specific counterparts. In particular, while the unconditional mean is a probability-weighted average of state means, the unconditional variance includes an additional term reflecting the risk associated with transitions between regimes (see also Ang and Timmermann, 2012).

Lemma IA1. *Under Assumption IA1, the unconditional first and second moments of y_t, as defined in Eq. (5) of the main paper, are given by*

$$\widetilde{\mu}_y = \sum_{\ell=1}^{L} \mu_{y,\ell}\pi_\ell, \tag{IA1}$$

$$\widetilde{V}_y = \bar{V}_y + \overline{M}_y, \tag{IA2}$$

where

$$\bar{V}_y = \sum_{\ell=1}^{L} V_{y,\ell}\pi_\ell, \tag{IA3}$$

$$\overline{M}_y = \sum_{\ell=1}^{L} \left(\mu_{y,\ell} - \widetilde{\mu}_y\right)\left(\mu_{y,\ell} - \widetilde{\mu}_y\right)'\pi_\ell. \tag{IA4}$$

Proof: Let η_t be a $(N+K)$-vector that is *i.i.d.* $N(0_{N+K}, I_{N+K})$, where 0_{N+K} is a $(N+K)$-vector of zeros and I_{N+K} is the identity matrix. η_t is independent of s_t. We can then express y_t in Eq. (5) of the main paper as

$$y_t = \mu_{y,s_t} + V_{y,s_t}^{1/2}\eta_t. \tag{IA5}$$

From the law of iterated expectations, we then have

$$E\left(y_t\right) = E\left[E\left(y_t \mid s_t\right)\right] = \sum_{\ell=1}^{L} E(\mu_{y,\ell} + V_{y,\ell}^{1/2}\eta_t)\pi_\ell = \sum_{\ell=1}^{L} \mu_{y,\ell}\pi_\ell. \tag{IA6}$$

Similarly,

$$\begin{aligned} E\left[\left(y_t - \widetilde{\mu}_y\right)\left(y_t - \widetilde{\mu}_y\right)'\right] &= E\left\{E\left[\left(y_t - \widetilde{\mu}_y\right)\left(y_t - \widetilde{\mu}_y\right)' \mid s_t\right]\right\} \\ &= \sum_{\ell=1}^{L} E\left[\left(\mu_{y,\ell} + V_{y,\ell}^{1/2}\eta_t - \widetilde{\mu}_y\right)\left(\mu_{y,\ell} + V_{y,\ell}^{1/2}\eta_t - \widetilde{\mu}_y\right)'\right]\pi_\ell \\ &= \sum_{\ell=1}^{L} \left[V_{y,\ell} + \left(\mu_{y,\ell} - \widetilde{\mu}_y\right)\left(\mu_{y,\ell} - \widetilde{\mu}_y\right)'\right]\pi_\ell. \end{aligned} \tag{IA7}$$

This completes the proof of Lemma IA1. ■

In what follows, we define the estimators of the parameters of the Markov-switching specification for the vector f_t. Recall that we denote by f_t the set of factors of any two competing models. Consistent with the model in Eq. (5) of the main paper, we express f_t as

$$f_t = \left(\xi_t^{\star\prime} \otimes I_K\right)\left(\mu_f^\star + V_f^{\star 1/2} u_t\right), \tag{IA8}$$

where $\mu_f^\star = (\mu_{f,1}', \ldots, \mu_{f,L}')'$, $V_f^\star = (V_{f,1}, \ldots, V_{f,L})'$, $\xi_t^\star = (\xi_{1t}, \ldots, \xi_{Lt})'$ is a random L-vector whose ℓth element is equal to unity if $s_t = \ell$ and zero otherwise, and u_t is defined in Lemma IA2 below. The notation $\otimes$ represents the Kronecker product (also called the tensor product). It is well-known that $\mathrm{E}(\xi_t^\star) = \pi^\star$, $\mathrm{E}(\xi_t^\star|\xi_{t-1}^\star) = P'\xi_{t-1}^\star$, and $\xi_t^{\star\prime}\mathbf{1} = 1$ (see, for example, Hamilton,

1994).

Let $\theta^\star = (\theta_1', \ldots, \theta_L')'$, where $\theta_\ell = (\mu_{f,\ell}', \text{vec}(V_{f,\ell})')'$, and $\text{vec}(\cdot)$ denotes the vectorization operator that stacks the columns of the argument matrix. The problem of interest is to estimate $\theta^\star$ together with the vector of transition probabilities $\varrho := \text{vec}(P)$ from the observed sequence $\{f_t\}_{t=1}^T$. Set $\vartheta = (\theta^{\star\prime}, \varrho')'$. A commonly used estimator of ϑ is the maximizer of the logarithmic conditional likelihood function $\mathcal{L}(\vartheta)$ for ϑ, where

$$\mathcal{L}(\vartheta) = \sum_{t=1}^{T} \ln \mathfrak{g}(f_t | F_{t-1}; \vartheta). \tag{IA9}$$

In Eq. (IA9), $\mathfrak{g}(f_t|F_{t-1};\vartheta)$ represents the conditional probability density function of f_t given the history $F_{t-1} = \{f_0, \ldots, f_{t-1}\}$. It is determined by the probability density function of u_t, which is specified in Lemma IA2 below. The estimates $\hat{\theta}^\star$ of the parameters $\theta^\star$ can be obtained from the solution to the following first-order condition:

$$\frac{\partial \mathcal{L}(\vartheta)}{\partial \theta^\star} = \sum_{t=1}^{T} \sum_{\ell=1}^{L} \frac{\partial \ln \mathfrak{g}(f_t | s_t = \ell, F_{t-1}; \theta^\star)}{\partial \theta^\star} \mathcal{P}(s_t = \ell | F_T; \vartheta), \tag{IA10}$$

and an estimator of the elements of ϱ can be defined as (see, for example, Hamilton, 1990)

$$\begin{aligned} \widehat{p}_{\iota\ell} &= \frac{\sum_{t=1}^{T} \mathcal{P}(s_t = \ell, s_{t-1} = \iota | F_T; \widehat{\vartheta})}{\sum_{t=1}^{T} \mathcal{P}(s_{t-1} = \iota | F_T; \widehat{\vartheta})} \\ &= \left(\sum_{t=1}^{T} \hat{\xi}_{\ell t-1|T} \right)^{-1} \left(\sum_{t=1}^{T} \hat{\xi}_{\iota\ell t|T} \right), \end{aligned} \tag{IA11}$$

where $\hat{\vartheta}$ is a vector with the maximum likelihood estimates of ϑ, $\hat{\xi}_{\iota\ell t|T} := \mathcal{P}(s_t = \ell, s_{t-1} = \iota \mid F_T; \widehat{\vartheta})$, and $\hat{\xi}_{\ell t-1|T} := \mathcal{P}(s_{t-1} = \ell \mid F_T; \widehat{\vartheta})$.

Let the set to which ϑ belongs be denoted by Θ and let ϑ_0 be the true parameter value. Under some general regularity conditions, which we state in the lemma below, it can be shown (see, for example, Kiefer, 1978, Theorem 1) that for large enough T, the maximum likelihood estimator, $\widehat{\vartheta}$, is consistent for ϑ_0, and $\sqrt{T}(\widehat{\vartheta} - \vartheta_0)$ is asymptotically normally distributed with mean zero and variance-covariance matrix, $\mathcal{I}(\vartheta_0)^{-1}$, where $\mathcal{I}(\vartheta_0)$ is the Fisher information matrix. Cavicchioli (2014) provides a matrix expression for the maximum likelihood estimator of $\theta^\star$ and derives an explicit form for the associated asymptotic variance-covariance matrix, which is block-diagonal. However, while Cavicchioli (2014) derives the asymptotic distribution of the inverse covariance matrix, we focus on the covariance matrix itself. This alternative parametrization leads to a different variance representation involving a commutation matrix. We therefore state the result explicitly and provide its proof for completeness in the following lemma.

Lemma IA2. *Let Assumption IA1 hold. Assume further that (i) u_t in Eq. (IA8) is an i.i.d. standard normal random vector that is independent of s_t; (ii) ϑ_0 is an interior point of Θ, and Θ is compact; (iii) the family of mixtures of at most L Gaussian components for f_t is identifiable in the sense that, for at least one pair of components (ι, ℓ) with $\iota \neq \ell$, the mean vector or covariance matrices must be distinct: $\mu_{f,\iota} \neq \mu_{f,\ell}$ or $V_{f,\iota} \neq V_{f,\ell}$; and (iv) the estimator of the elements of ϱ defined in Eq. (IA11) and the maximum likelihood estimates of $\mu_{f,\ell}$ and $V_{f,\ell}$ that are characterized by the following formulas:*

$$\widehat{\mu}_{f,\ell} = \left(\sum_{t=1}^{T} \hat{\xi}_{\ell t|T} \right)^{-1} \left(\sum_{t=1}^{T} f_t \hat{\xi}_{\ell t|T} \right), \tag{IA12}$$

$$\widehat{V}_{f,\ell} = \left(\sum_{t=1}^{T} \hat{\xi}_{\ell t|T} \right)^{-1} \left(\sum_{t=1}^{T} \left(f_t - \widehat{\mu}_{f,\ell}\right) \left(f_t - \widehat{\mu}_{f,\ell}\right)' \hat{\xi}_{\ell t|T} \right), \tag{IA13}$$

for $\iota, \ell = 1, \ldots, L$, are consistent. Then, as $T \to \infty$,

(i) the asymptotic distribution of the estimator $\hat{\theta}_\ell = (\hat{\mu}'_{f,\ell}, vec(\hat{V}_{f,\ell})')'$ is

$$\sqrt{T}\left(\widehat{\theta}_\ell - \theta_\ell\right) \Longrightarrow N\left(0_{K(K+1)}, \Omega_\ell\right), \tag{IA14}$$

for all ℓ in $\{1, \ldots, L\}$, and

$$\Omega_\ell = \begin{bmatrix} \pi_\ell^{-1} V_{f,\ell} & 0_{K \times K^2} \\ 0_{K^2 \times K} & \pi_\ell^{-1} \left(I_{K^2} + \mathcal{K}_K\right) \left(V_{f,\ell} \otimes V_{f,\ell}\right) \end{bmatrix}, \tag{IA15}$$

where $\mathcal{K}_K$ is a $K^2 \times K^2$ commutation matrix, which is defined as $\mathcal{K}_K vec(V_{f,\ell}) = vec(V'_{f,\ell})$.

(ii) Also, for all $\ell \in \{1, \ldots, L\}$, the estimator $\widehat{\pi}_\ell := T^{-1} \sum_{t=1}^{T} \hat{\xi}_{\ell t|T}$ is consistent for π_ℓ.

Proof: For (ii), we have (see Cavicchioli, 2014)

$$\text{plim}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \hat{\xi}_{\ell t|T} = E\left(\xi_{\ell t}\right) = \pi_\ell.$$

To prove (i), we calculate the asymptotic information matrix for the estimator $\hat{\theta}_\ell$ as

$$\begin{aligned} \mathcal{I}(\theta_\ell) &= E \begin{pmatrix} \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \mu_{f,\ell}}\right) \cdot \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \mu_{f,\ell}}\right)' & \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \mu_{f,\ell}}\right) \cdot \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \text{vec}\left(V_{f,\ell}\right)}\right)' \\ \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \text{vec}\left(V_{f,\ell}\right)}\right) \cdot \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \mu_{f,\ell}}\right)' & \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \text{vec}\left(V_{f,\ell}\right)}\right) \cdot \left(\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \text{vec}\left(V_{f,\ell}\right)}\right)' \end{pmatrix} \\ &= E \begin{pmatrix} A & B \\ B' & C \end{pmatrix}. \end{aligned} \tag{IA16}$$

Note that under Assumption IA1,

$$\ln \mathfrak{g}(f_t|s_t = \ell, F_{t-1}; \theta^\star) \propto -\frac{1}{2}\ln|V_{f,\ell}| - \frac{1}{2}\left(f_t - \mu_{f,\ell}\right) V_{f,\ell}^{-1}\left(f_t - \mu_{f,\ell}\right)'. \tag{IA17}$$

Thus,

$$\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \mu_{f,\ell}} = \sum_{t=1}^{T} V_{f,\ell}^{-1/2} u_t \hat{\xi}_{\ell t|T} \tag{IA18}$$

and

$$\frac{\partial \ln \mathcal{L}(\vartheta)}{\partial \mathrm{vec}\left(V_{f,\ell}\right)} = \frac{1}{2}\sum_{t=1}^{T}\left[\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right)(u_t \otimes u_t) - \mathrm{vec}\left(V^{-1}\right)\right]\hat{\xi}_{\ell t|T}. \tag{IA19}$$

It then follows that

$$\begin{aligned} A &= E\left(\sum_{t=1}^{T}\sum_{\tau=1}^{T} V_{f,\ell}^{-1/2} u_t u_\tau' \hat{\xi}_{\ell t|T}\hat{\xi}_{\ell \tau|T} V_{f,\ell}^{-1/2}\right) \\ &= T V_{f,\ell}^{-1/2} E\left(\frac{1}{T}\sum_{t=1}^{T} u_t u_t' \hat{\xi}_{\ell t|T}^2\right) V_{f,\ell}^{-1/2} \\ &= \pi_\ell T V_{f,\ell}^{-1} \end{aligned} \tag{IA20}$$

as $\xi_t^\star$ is independent of u_t and $E(\xi_t^\star \xi_t^{\star\prime}) = \mathrm{diag}(\pi_1, \ldots, \pi_L)$. Similarly, for B, we have

$$\begin{aligned} B' &= E\left(\frac{1}{2}\sum_{t=1}^{T}\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right)(u_t \otimes u_t) u_t' \hat{\xi}_{\ell t|T}^2 V_{f,\ell}^{-1/2}\right) \\ &= 0_{K^2 \times K}. \end{aligned} \tag{IA21}$$

Finally, we derive C as

$$\begin{aligned} C = E\Bigg(&\frac{1}{4}\sum_{t=1}^{T}\Big[\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right)\left(u_t u_t' \otimes u_t u_t'\right)\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right) \\ &-\mathrm{vec}\left(V_{f,\ell}^{-1}\right)\mathrm{vec}\left(V_{f,\ell}^{-1}\right)'\Big]\hat{\xi}_{\ell t|T}^2\Bigg) \\ = \frac{T}{4}&\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right) E\left(\frac{1}{T}\sum_{t=1}^{T}\left[\left(u_t u_t' \otimes u_t u_t'\right) - \mathrm{vec}\left(I_K\right)\mathrm{vec}\left(I_K\right)'\right]\hat{\xi}_{\ell t|T}^2\right) \\ &\times\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right) \\ = \pi_\ell \frac{T}{4}&\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right)\left(I_{K^2} + \mathcal{K}_K\right)\left(V_{f,\ell}^{-1/2} \otimes V_{f,\ell}^{-1/2}\right) \\ = \pi_\ell \frac{T}{4}&\left(V_{f,\ell}^{-1} \otimes V_{f,\ell}^{-1}\right)\left(I_{K^2} + \mathcal{K}_K\right). \end{aligned} \tag{IA22}$$

The third equality in the above equation follows from Theorem 4.1 in Magnus and Neudecker (1979). Putting all the derived blocks together, the resulting information matrix takes the form

$$\mathcal{I}(\theta_\ell) = T \begin{pmatrix} \pi_\ell V_{f,\ell}^{-1} & 0_{K\times K^2} \\ 0_{K^2\times K} & \pi_\ell \frac{1}{4}\left(V_{f,\ell}^{-1}\otimes V_{f,\ell}^{-1}\right)(I_{K^2}+\mathcal{K}_K) \end{pmatrix}. \tag{IA23}$$

Next, let $\mathcal{M} = 1/2(I_{K^2} + \mathcal{K}_K)$. Note that $\mathcal{M}$ is symmetric and idempotent, so it is straightforward to verify that its Moore-Penrose pseudo-inverse is $\mathcal{M}$ itself. Therefore, the asymptotic variance of $\hat{\theta}_\ell$, obtained as the inverse of the asymptotic information matrix in Eq. (IA23), is given by

$$\mathcal{I}(\theta_\ell)^{-1} = \frac{1}{T} \begin{pmatrix} \pi_\ell^{-1} V_{f,\ell} & 0_{K\times K^2} \\ 0_{K^2\times K} & \pi_\ell^{-1}\left(I_{K^2}+\mathcal{K}_K\right)\left(V_{f,\ell}\otimes V_{f,\ell}\right). \end{pmatrix}. \tag{IA24}$$

This completes the proof of Lemma IA2. ■

Remark. For notational simplicity, we omit the subscript '0' from θ_ℓ. The smoothed regime probabilities, $\xi_{\ell t|T}$, can be calculated using the algorithm described in Kim (1994). The necessary conditions needed to ensure that *(iv)* is satisfied are discussed in Leroux (1992) and Bickel et al. (1998). Eqs. (IA11), (IA12), and (IA13) suggest an iterative algorithm for finding the maximum likelihood estimates. A detailed description of this algorithm can be found in Cavicchioli (2014).

IA.1.2 Proofs of the main results

Proof of Proposition 1

Rewrite the elements of $\widetilde{\mu}_y$ and $\widetilde{V}_y$ as

$$\widetilde{\mu}_y = \begin{pmatrix} \widetilde{\mu}_r \\ \widetilde{\mu}_f \end{pmatrix}, \quad \widetilde{V}_y = \begin{pmatrix} \widetilde{V}_{rr} & \widetilde{V}_{rf} \\ \widetilde{V}_{fr} & \widetilde{V}_{ff} \end{pmatrix}. \tag{IA25}$$

From Lemma IA1, we have $\widetilde{\mu}_r = \sum_{\ell=1}^{L}\alpha_\ell\pi_\ell + \sum_{\ell=1}^{L}\beta_\ell\mu_{f,\ell}\pi_\ell$, where $\beta = \widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}$ and $\widetilde{\mu}_f = \sum_{\ell=1}^{L}\mu_{f,\ell}\pi_\ell$. Substituting these expressions into the first line of Eq. (11) in the main paper and rearranging produces the second line of Eq. (11).

Next, using the formula for the inverse of a partitioned matrix, we obtain the following:

$$\widetilde{V}_y^{-1} = \begin{pmatrix} \left(\widetilde{V}_{rr} - \widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1} & -\left(\widetilde{V}_{rr} - \widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1}\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1} \\ -\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\left(\widetilde{V}_{rr} - \widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1} & \widetilde{V}_{ff}^{-1} + \widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\left(\widetilde{V}_{rr} - \widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1}\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1} \end{pmatrix}.$$

Applying this inverse and performing algebraic simplification yields

$$\begin{aligned}\widetilde{\mu}_y'\widetilde{V}_y^{-1}\widetilde{\mu}_y =& \widetilde{\mu}_r'\left(\widetilde{V}_{rr}-\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1}\widetilde{\mu}_r - 2\widetilde{\mu}_r'\left(\widetilde{V}_{rr}-\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1}\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{\mu}_f \\ &+ \widetilde{\mu}_f'\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\left(\widetilde{V}_{rr}-\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1}\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{\mu}_f + \widetilde{\mu}_f'\widetilde{V}_{ff}^{-1}\widetilde{\mu}_f \\ =& \left(\widetilde{\mu}_r-\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{\mu}_f\right)'\left(\widetilde{V}_{rr}-\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{V}_{fr}\right)^{-1}\left(\widetilde{\mu}_r-\widetilde{V}_{rf}\widetilde{V}_{ff}^{-1}\widetilde{\mu}_f\right)+\widetilde{\mu}_f'\widetilde{V}_{ff}^{-1}\widetilde{\mu}_f.\end{aligned}$$

Substituting for $\widetilde{\mu}_r$ and β gives the desired result in Eq. (12) of the main paper. This completes the proof of Proposition 1. ■

Proof of Proposition 2

Let $f_{Bt} = (f_{1t}', f_{2t}')'$ and note that

$$\mu_{B,\ell} = \begin{pmatrix}\mu_{1,\ell}\\ \mu_{2,\ell}\end{pmatrix} \text{ and } V_{B,\ell} = \begin{pmatrix}V_{11,\ell} & V_{12,\ell}\\ V_{21,\ell} & V_{22,\ell}\end{pmatrix}.$$

Similar to the proof of Proposition 1, we can show that

$$\begin{aligned}\mu_{B,\ell}'V_{B,\ell}^{-1}\mu_{B,\ell} =& \mu_{1,\ell}'V_{11,\ell}^{-1}\mu_{1,\ell} \\ &+ \left(\mu_{22,\ell}-V_{21,\ell}V_{11,\ell}^{-1}\mu_{1,\ell}\right)'\left(V_{22,\ell}-V_{21,\ell}V_{11,\ell}^{-1}V_{12,\ell}\right)^{-1}\left(\mu_{2,\ell}-V_{21,\ell}V_{11,\ell}^{-1}\mu_{1,\ell}\right).\end{aligned}$$

This completes the proof of Proposition 2. ■

Proof of Proposition 3

From Lemma IA2, we have

$$\sqrt{T}\left(\widehat{\theta}_\ell-\theta_\ell\right) \Longrightarrow N\left(0_{K(K+1)}, \Omega_\ell\right),$$

for all ℓ in $\{1,\ldots,L\}$. In what follows, we omit the subscript 'ℓ' from the notation if there is no ambiguity.

Using the Delta method and the fact that $\widehat{\alpha}_{21}$ is a smooth function of $\widehat{\theta}$, we obtain

$$\sqrt{T}\left(\widehat{\alpha}_{21}-\alpha_{21}\right) \Longrightarrow N\left(0_{K_2}, G_0\right), \tag{IA26}$$

where

$$G_0 = \left[\frac{d\alpha_{21}}{d\theta'}\right]\Omega\left[\frac{d\alpha_{21}}{d\theta'}\right]'. \tag{IA27}$$

It is straightforward to obtain

$$\frac{\partial \alpha_{21}}{\partial \mu_1'} = -V_{21}V_{11}^{-1}, \quad \frac{\partial \alpha_{21}}{\partial \mu_2'} = I_{K_2}, \tag{IA28}$$

where I_{K_2} is the identity matrix. In order to obtain the derivative of α_{21} with respect to $\text{vec}(V)$, note that we can write

$$V_{21} = [0_{K_2 \times K_1}, I_{K_2}]\, V_f\, [I_{K_1}, 0_{K_1 \times K_2}]', \quad V_{11} = [I_{K_1}, 0_{K_1 \times K_2}]\, V_f\, [I_{K_1}, 0_{K_1 \times K_2}]', \tag{IA29}$$

to obtain

$$\frac{\partial \text{vec}(V_{21})}{\partial \text{vec}(V_f)'} = [I_{K_1}, 0_{K_1 \times K_2}] \otimes [0_{K_2 \times K_1}, I_{K_2}] \tag{IA30}$$

and

$$\frac{\partial \text{vec}(V_{11})}{\partial \text{vec}(V_f)'} = [I_{K_1}, 0_{K_1 \times K_2}] \otimes [I_{K_1}, 0_{K_1 \times K_2}]\,. \tag{IA31}$$

Now, with the identity

$$\begin{aligned}
\frac{\partial \text{vec}(V_{11}^{-1})}{\partial \text{vec}(V_f)'} &= \frac{\partial \text{vec}(V_{11}^{-1})}{\partial \text{vec}(V_{11})'} \frac{\partial \text{vec}(V_{11})}{\partial \text{vec}(V_f)'} \\
&= -\left(V_{11}^{-1} \otimes V_{11}^{-1}\right) \left([I_{K_1}, 0_{K_1 \times K_2}] \otimes [I_{K_1}, 0_{K_1 \times K_2}]\right) \\
&= \left[V_{11}^{-1}, 0_{K_1 \times K_2}\right] \otimes \left[-V_{11}^{-1}, 0_{K_1 \times K_2}\right],
\end{aligned} \tag{IA32}$$

we can use the product rule to obtain

$$\begin{aligned}
\frac{\partial \alpha_{21}}{\partial \text{vec}(V_f)'} &= \left(\mu_1' V_{11}^{-1} \otimes I_{K_2}\right) \frac{\partial \text{vec}(V_{21})}{\partial \text{vec}(V_f)'} + \left(\mu_1' \otimes V_{21}\right) \frac{\partial \text{vec}(V_{11}^{-1})}{\partial \text{vec}(V_f)'} \\
&= \left(\mu_1' V_{11}^{-1} \otimes I_{K_2}\right) \left([I_{K_1}, 0_{K_1 \times K_2}] \otimes [0_{K_2 \times K_1}, I_{K_2}]\right) \\
&\quad + \left(\mu_1' \otimes V_{21}\right) \left[V_{11}^{-1}, 0_{K_1 \times K_2}\right] \otimes \left[-V_{11}^{-1}, 0_{K_1 \times K_2}\right] \\
&= \left[\mu_1' V_{11}^{-1}, 0_{K_2}'\right] \otimes [0_{K_2 \times K_1}, I_{K_2}] + \left[\mu_1' V_{11}^{-1}, 0_{K_2}'\right] \otimes \left[-V_{21}V_{11}^{-1}, 0_{K_2 \times K_2}\right] \\
&= \left[\mu_1' V_{11}^{-1}, 0_{K_2}'\right] \otimes \left[-V_{21}V_{11}^{-1}, I_{K_2}\right].
\end{aligned} \tag{IA33}$$

The asymptotic variance is therefore

$$\begin{aligned} G_0 =& \pi^{-1}\left(\left[\frac{\partial \alpha_{21}}{\partial \mu_f'}\right]\Omega\left[\frac{\partial \alpha_{21}}{\partial \mu_f'}\right]'\right)+\pi^{-1}\left(\left[\frac{\partial \alpha_{21}}{\partial \text{vec}(V_f)'}\right](I_{K^2}+\mathcal{K}_K)(\Omega\otimes\Omega)\left[\frac{\partial \alpha_{21}}{\partial \text{vec}(V_f)'}\right]'\right) \\ =& \pi^{-1}\left(V_{22}-V_{21}V_{11}^{-1}V_{12}\right)+\pi^{-1}\left(\mu_1'V_{11}^{-1}\mu_1\left[V_{22}-V_{21}V_{11}^{-1}V_{12}\right]\right) \\ =& \pi^{-1}\left(1+\mu_1'V_{11}^{-1}\mu_1\right)\left(V_{22}-V_{21}V_{11}^{-1}V_{12}\right), \end{aligned} \tag{IA34}$$

where $K = K_1 + K_2$. The derivation of the second term in the second equality above uses the fact that

$$\left(\left[\mu_1'V_{11}^{-1}, 0_{K_2}'\right]\otimes\left[-V_{21}V_{11}^{-1}, I_{K_2}\right]\right)\mathcal{K}_K(\Omega\otimes\Omega)\left(\left[\mu_1'V_{11}^{-1}, 0_{K_2}'\right]\otimes\left[-V_{21}V_{11}^{-1}, I_{K_2}\right]\right)' = 0_{K_2\times K_2}.$$

This completes the proof of Proposition 3. ■

Proof of Proposition 4

For ease of exposition, we redefine the mean and the variance of the factor from model A in state ℓ as $\mu_{1,\ell}$ and $V_{11,\ell}$, respectively. Similarly, the mean and the variance of the factor from model B in state ℓ are redefined as $\mu_{2,\ell}$ and $V_{22,\ell}$. Therefore,

$$\mu_{f,\ell} = \begin{pmatrix}\mu_{1,\ell}\\ \mu_{2,\ell}\end{pmatrix}, \quad V_{f,\ell} = \begin{pmatrix}V_{11,\ell} & V_{12,\ell}\\ V_{21,\ell} & V_{22,\ell}\end{pmatrix}. \tag{IA35}$$

Again, in what follows, the subscript 'ℓ' will be omitted from the notation if there is no ambiguity.

Let $\Delta = Sh_A^2 - Sh_B^2$. Since $\widehat{\Delta}$ is a smooth function of $\widehat{\theta}$, we use the Delta method to obtain

$$\sqrt{T}\left(\widehat{\Delta}-\Delta\right) \Longrightarrow N(0, G_0), \tag{IA36}$$

where

$$G_0 = \left[\frac{d\Delta}{d\theta'}\right]\Omega\left[\frac{d\Delta}{d\theta'}\right]'. \tag{IA37}$$

As in the proof of Proposition 3, we obtain

$$\frac{\partial \Delta}{\partial \mu_1'} = 2\mu_1'V_{11}^{-1}, \quad \frac{\partial \Delta}{\partial \mu_2'} = -2\mu_2'V_{22}^{-1}, \tag{IA38}$$

and

$$\frac{\partial \Delta}{\partial \text{vec}(V_f)'} = -\text{vec}\left(\mu_1\mu_1'\right)'\left(\left[V_{11}^{-1}, 0_{K_1\times K_2}\right]\otimes\left[V_{11}^{-1}, 0_{K_1\times K_2}\right]\right) + \text{vec}\left(\mu_2\mu_2'\right)'\left(\left[0_{K_2\times K_1}, V_{22}^{-1}\right]\otimes\left[0_{K_2\times K_1}, V_{22}^{-1}\right]\right). \tag{IA39}$$

Then

$$\begin{aligned} G_0 =& \pi^{-1}\left(\left[\frac{\partial \Delta}{\partial \mu_f'}\right]\Omega\left[\frac{\partial \Delta}{\partial \mu_f'}\right]'\right) + \pi^{-1}\left(\left[\frac{\partial \Delta}{\partial \text{vec}(V_f)'}\right](I_{K^2}+\mathcal{K}_K)(\Omega\otimes\Omega)\left[\frac{\partial \Delta}{\partial \text{vec}(V_f)'}\right]'\right) \\ =& \pi^{-1}\left(4\mu_1'V_{11}^{-1}\mu_1 - 8\mu_1'V_{11}^{-1}V_{12}V_{22}^{-1}\mu_2 + 4\mu_2'V_{22}^{-1}\mu_2\right) \\ &+ 2\pi^{-1}\Big(\text{vec}\left(\mu_1\mu_1'\right)'\left[V_{11}^{-1}\otimes V_{11}^{-1}\right]\text{vec}\left(\mu_1\mu_1'\right) \\ &\quad - 2\text{vec}\left(\mu_1\mu_1'\right)'\left[V_{11}^{-1}V_{12}V_{22}^{-1}\otimes V_{11}^{-1}V_{12}V_{22}^{-1}\right]\text{vec}\left(\mu_2\mu_2'\right) \\ &\quad + \text{vec}\left(\mu_2\mu_2'\right)'\left[V_{22}^{-1}\otimes V_{22}^{-1}\right]\text{vec}\left(\mu_2\mu_2'\right)\Big) \\ =& \pi^{-1}\left(4\mu_1'V_{11}^{-1}\mu_1 - 8\mu_1'V_{11}^{-1}V_{12}V_{22}^{-1}\mu_2 + 4\mu_2'V_{22}^{-1}\mu_2\right) \\ &+ 2\pi^{-1}\left(\left[\mu_1'V_{11}^{-1}\mu_1\right]^2 - 2\left[\mu_1'V_{11}^{-1}V_{12}V_{22}^{-1}\mu_2\right]^2 + \left[\mu_2'V_{22}^{-1}\mu_2\right]^2\right). \end{aligned} \tag{IA40}$$

The second term in the second equality above arises from the following property of the commutation matrix:

$$(A\otimes B)\mathcal{K}_K = (B\otimes A),$$

for some $A\in\mathbb{R}^{1\times K}$ and $B\in\mathbb{R}^{1\times K}$.

Next, we show that $\rho_{AB}Sh_{A,}Sh_B = \mu_1'V_{11}^{-1}V_{12}V_{22}^{-1}\mu_2$. Consider the two portfolios, $\mathfrak{p}_A$ and $\mathfrak{p}_B$, that are constructed by investing in f_t, the factors from both models. Let the weights of these portfolios be given by

$$\mathfrak{w}_A = \left[\gamma_A\mu_1'V_{11}^{-1}, 0_{K_2}'\right] \tag{IA41}$$

and

$$\mathfrak{w}_B = \left[0_{K_1}', \gamma_B\mu_2'V_{22}^{-1}\right], \tag{IA42}$$

respectively, where $\gamma_A = \mathfrak{m}_A(\mu_1'V_{11}^{-1}\mu_1)^{-1}$, $\gamma_B = \mathfrak{m}_B(\mu_2'V_{22}^{-1}\mu_2)^{-1}$, and $\mathfrak{m}_i$ is the target return

for portfolio i, $i \in \{A, B\}$. It is straightforward to show that $\gamma_i = \sqrt{Var(\mathfrak{p}_i)}/Sh_i$. We then have

$$\begin{aligned} Cov\left(\mathfrak{p}_A, \mathfrak{p}_B\right) &= \mathfrak{w}_A V_f \mathfrak{w}_B' \\ &= \gamma_A \mu_1' V_{11}^{-1} V_{12} V_{22}^{-1} \mu_2 \gamma_B \\ &= \frac{\mu_1' V_{11}^{-1} V_{12} V_{22}^{-1} \mu_2}{Sh_A Sh_B} \sqrt{Var\left(\mathfrak{p}_A\right)} \sqrt{Var\left(\mathfrak{p}_B\right)}. \end{aligned} \tag{IA43}$$

Solving for $\mu_1' V_{11}^{-1} V_{12} V_{22}^{-1} \mu_2$ yields the desired result. This completes the proof of Proposition 4. ■

IA.2 Simulation evidence

IA.2.1 Numerical example showing the effect of regime shifts on the squared Sharpe ratio improvement

In this section, we use a numerical example to show that $\gamma' \widetilde{V}_{r \cdot f}^{-1} \gamma$, which we will refer to as the "bias", is likely to be sizable. Let $L = 2$, $K = 2$, $N = 8$, and set

$$\mu_{f,1} = \begin{bmatrix} 0.1 \\ 0.15 \end{bmatrix}, \quad \mu_{f,2} = -1.5\mu_{f,1}, \quad V_{ff,1} = \begin{bmatrix} 0.09, & -0.06 \\ -0.06, & 0.15 \end{bmatrix}, \quad \text{and} \quad V_{ff,2} = 1.5V_{ff,1}.$$

For β_1, we generate random drawings from a uniform distribution in the open interval $(-2, 2)$ and set $\beta_2 = 1.5\beta_1$. We also restrict the alphas to be zero in both states so that when the bias is zero, $Sh^2(f, r) = Sh^2(f)$. Finally, we set $V_{r \cdot f,1}$ to be a diagonal matrix, with elements on the main diagonal drawn from the uniform distribution, and let $V_{r \cdot f,2} = 1.5V_{r \cdot f,1}$. The bias as a function of π_1 is plotted in Fig. IA1. We see that the bias peaks around $\pi_1 = 0.55$ where the unconditional variance of the different parameters is most sizable.

IA.2.2 Monte Carlo simulations

This section evaluates the small-sample properties of our various test statistics via Monte Carlo simulations. We also analyze the effect of regime shifts on the properties of the procedure proposed by Barillas et al. (2020), by applying their tests on the full sample of simulated data. In all simulation experiments, the data generating process (DGP) is defined by the nonlinear model specified in Eq. (IA8) with $L = 2$, that is, with two states. We set the time-series sample size, T, to 600, which is equal to the actual sample size of the data used in our empirical analysis. The actual rejection rates of the tests are compared over 10,000 replications to the nominal 5% level of significance. The true variance-covariance matrices of factor returns in both

states are set equal to the maximum likelihood estimates reported in Section 3 of the main paper. To investigate the power of the tests, we set the means of the factors in state 1 equal to their maximum likelihood estimates, whereas to investigate the size properties of the tests, we constrain the means of the factors in state 2 as per Barillas et al. (2020, p.1870). We also consider the following pairs of transition probabilities:

$$(p_{11}, p_{22}) = \{(0.95, 0.95), (0.95, 0.80), (0.80, 0.95), (0.50, 0.50)\}.$$

Similar pairs of transition probabilities have also been considered by Psaradakis and Sola (1998) in their study of the finite-sample properties of the maximum likelihood estimator in autoregressive models with Markov switching. The first pair of transition probabilities, (0.95, 0.95), allows for a symmetry in the persistence of the states. The limiting probability for each state is 0.50 and each state is expected to persist for 20 months on average. The second pair of transition probabilities, (0.95, 0.80), allows state 1, in which the models perform differently from each other, to be more likely than state 2. The limiting probability of state 1 is 0.80, with a corresponding average duration of 20 months, whereas the limiting probability of state 2 is 0.20, with a corresponding average duration of 5 months. This scenario is reversed with the third pair of transition probabilities that we consider, (0.80, 0.95). State 2, in which both models perform equal to each other, is now the prevailing state, with a limiting probability of 0.80 and an average duration of 20 months, whereas state 1 is characterized by a limiting probability of 0.20 and an average duration of 5 months. The last pair of transition probabilities, (0.50, 0.50), implies that both states are equally likely, but each state has an average duration of only 2 months. Moreover, as Psaradakis and Sola (1998) point out, there is no persistence in the Markov chain under this scenario, since $p_{11} + p_{22} = 1$, and the observed value at time t is independent of the prevailing state at time $t - 1$.

We begin by examining the properties of the pairwise test of equal conditional squared Sharpe ratios for nested models. As discussed earlier, in such a case, it is sufficient to test the joint restriction of zero intercepts in a multivariate linear regression model, where the factors common to both models are included as regressors and the unique factors serve as test assets. We design the simulation setup so that in state 2 the SMB, $r_{\mathrm{I/A}}$, r_{ROE}, UMD, and $\mathrm{HML_m}$ factors of the BS6 model are exactly priced by the MKT factor (that is, the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965)), whereas in state 1 the CAPM is dominated by the BS6 model. In Panel A of Table IA1, we see that the Wald test performs well. Specifically, the results show that for $p_{11} = p_{22} = 0.95$, the rejection rates are 5.6% and 100% in state 2 and state 1, respectively. In comparison, the rejection rate of the Gibbons et al. (1989) test

for this simulated data is 87.6%. When $(p_{11}, p_{22}) = (0.95, 0.80)$, that is, state 2 is less likely than state 1, the rejection rates are 9.1% and 100% in state 2 and state 1, respectively. The rejection rate of the Gibbons et al. (1989) test in this case is 100%. On the other hand, with $(p_{11}, p_{22}) = (0.80, 0.95)$, that is, state 1 is less likely than state 2, the rejection rates are 5.9% and 99.6% in state 2 and state 1, respectively. In this case, the rejection rate of the Gibbons et al. (1989) test is only 19.7%, although the two models perform similarly most times. When $p_{11} = p_{22} = 0.50$, that is, both states are equally likely, the rejection rates are 8.8% and 100% in state 2 and state 1, respectively. On the other hand, the rejection rate of the Gibbons et al. (1989) test turns out to be 93.3%.

Next, we assess the small-sample properties of the pairwise test of equal conditional squared Sharpe ratios for nonnested models. In this experiment, we consider the FF5 model versus the BS6 model. The two models are nonnested but have common factors, MKT and SMB. Thus, the null hypothesis of equal conditional squared Sharpe ratios can hold when the MKT and SMB factors span the maximum conditional squared Sharpe ratio portfolio based on the unique factors from both models, that is, the $r_{\mathrm{I/A}}$, r_{ROE}, UMD, HML, $\mathrm{HML_m}$, RMW, and CMA factors. Again, this condition can be tested using the Wald test described in Section 2.2 of the main paper. In unreported results, we find that the rejection rates of the test are similar to the case for nested models. Thus, the asymptotic test performs quite well for empirically relevant samples sizes. Recall, however, that even if the spanning condition is rejected, the conditional squared Sharpe ratios of the factor portfolios can still be equal, and this equality can be tested using the results in Proposition 4. In this case, the simulation setup is designed such that in state 2, the true conditional squared Sharpe ratios of the FF5 and BS6 models are equal to 0.115, which is the sample estimate of the conditional squared Sharpe ratio of the BS6 model in state 2.

In Panel B of Table IA1, we see that the rejection rates of the test (that is, the hypothesis of equivalent model performance is true in state 2) are 3.7% when $p_{11} = p_{22} = 0.95$, 2.3% when $(p_{11}, p_{22}) = (0.95, 0.80)$, 4.7% when $(p_{11}, p_{22}) = (0.80, 0.95)$, and 4.5% when $p_{11} = p_{22} = 0.50$. The power of the normal test is explored by setting the population values of the squared Sharpe ratios of the FF5 and BS6 models in state 1 equal to their actual sample values. The rejection rates of the test for state 1 are 92.6% when $(p_{11}, p_{22}) = (0.95, 0.95)$, 99.4% when $(p_{11}, p_{22}) = (0.95, 0.80)$, 54.2% when $(p_{11}, p_{22}) = (0.80, 0.95)$, and 92.5% when $(p_{11}, p_{22}) = (0.50, 0.50)$. These results suggest that our testing procedure continues to be reliable with good power and size properties. On the other hand, the rejection rates of the Barillas et al. (2020) test are 47.7% when $p_{11} = p_{22} = 0.95$ and the difference between the squared Sharpe ratio of the BS6 and FF5 models is 0.114, 98% when $(p_{11}, p_{22}) = (0.95, 0.80)$ and the difference between the squared

Sharpe ratio of the BS6 and FF5 models is 0.182, only 9% when $(p_{11}, p_{22}) = (0.80, 0.95)$ and the difference between the squared Sharpe ratio of the BS6 and FF5 models is 0.046, and 47.4% when $p_{11} = p_{22} = 0.50$ and the difference between the squared Sharpe ratio of the BS6 and FF5 models is 0.114.

We now assess the properties of the multiple-comparison inequality test for nonnested models. Under this test, the conditional squared Sharpe ratio for the benchmark model is at least as high as that for all other models that are being considered. Therefore, to analyze the properties of the test, the simulation setup is designed as follows. In state 2, the true conditional squared Sharpe ratios of the FF5, HXZ4, BS6, and SY4 models are equal to 0.152 (which is the sample estimate of the conditional squared Sharpe ratio of the FF5 model in state 2). In state 1, the true conditional squared Sharpe ratios of these models are set equal to their sample estimates in state 1. Also, to analyze the effect that the number of models may have on the properties of the test, we consider the following two different scenarios: (i) FF5 versus {SY4, BS6} ($q = 2$); and (ii) FF5 versus {SY4, BS6, HXZ4} ($q = 3$). The test results in Table IA2 show that the rejection rates in state 2 are: 5.7% (for $q = 2$) and 5.2% (for $q = 3$) when $p_{11} = p_{22} = 0.95$; 4.9% (for $q = 2$) and 3.9% (for $q = 3$) when $(p_{11}, p_{22}) = (0.95, 0.80)$; 6.1% (for $q = 2$) and 5.7% (for $q = 3$) when $(p_{11}, p_{22}) = (0.80, 0.95)$; and 6.1% (for $q = 2$) and 5.5% (for $q = 3$) when $p_{11} = p_{22} = 0.50$. Thus, the test is well-specified under the null hypothesis of equivalent model performance. The results also suggest that the test has a good power. In particular, the rejection rates of the test in state 1 are close to 100% for most pairs of transition probabilities and different number of competing factor models that are being considered. When $(p_{11}, p_{22}) = (0.80, 0.95)$, the rejection rates are 67.2% for $q = 2$ and 61.9% for $q = 3$. In contrast, the unconditional test results for Barillas et al. (2020) range widely from 13.7% to 99.4%.

Overall, the results from the Monte Carlo simulations indicate that our state-conditional tests are reliable for sample lengths typically encountered in empirical work. In contrast, the unconditional procedure of Barillas et al. (2020) exhibits rejection rates that depend strongly on which regime dominates the sample. When the state in which competing models differ is prevalent, rejection rates are high, whereas when the state in which competing models coincide is prevalent, rejection rates are low.

To complement the two-state design above, we also analyze the properties of the proposed conditional tests by considering a DGP with a single state. In this case, the analysis of the test size for nested models is based on the simulation setup in which the SMB, HML_m, UMD, r_{ROE}, and $r_{\text{I/A}}$ factors of the BS6 model are exactly priced by the MKT factor of the CAPM. As for the power of the test, the population squared Sharpe ratios of the FF5 model and the CAPM are

set equal to their respective sample estimates in state 1 from the two-state specification. Next, we examine the statistical properties of the pairwise test of equal conditional squared Sharpe ratios for nonnested models. In this case, the analysis of the test size is based on setting the population squared Sharpe ratios of the two competing models, FF5 and BS6, equal to 0.115, which corresponds to the sample estimate of the conditional squared Sharpe ratio of the BS6 model in state 2 under the two-state model. To study power, the population squared Sharpe ratios of the FF5 and BS6 models are instead set equal to their respective sample estimates in state 1 from the two-state specification. The above simulation designs allow us to examine the ability of the test to detect differences in conditional squared Sharpe ratios of factor models when the true DGP does not feature regime shifts.

Table IA3 reports the rejection probability rates of the alpha-based (Panel A) and direct (Panel B) tests for the pairwise comparison of factor models. The rejection rates correspond to a sample size of $T = 600$ and are computed from 10,000 replications. We see that the conditional test is oversized in both states. The high empirical size of the test under a single-state DGP is to be expected, given that the procedure conditions on the existence of two regimes; this misspecification leads to over-rejection due to the test exploiting spurious state separation. The results in Panel A show that the test has nearly 100% power in both states. We also see in Panel B that the test has a good (albeit not size-adjusted) power in both states. Thus, the size and power of our conditional test are consistent across states. The unconditional testing procedure shows good size and power properties, which is consistent with Barillas et al. (2020).

For nonnested multiple-model comparison, we adopt a simulation design similar to the one for pairwise nonnested model comparison. Under the null, the population squared Sharpe ratios of all four models (that is, FF5, HXZ4, BS6, and SY4) are set equal to 0.152, which is the sample estimate of the conditional squared Sharpe ratio of the FF5 model in state 2 of the two-state model. Under the alternative, the four population squared Sharpe ratios are set equal to their respective sample estimates in state 1. As before, the DGP consists of a single state, thereby enabling us to evaluate the size and power properties of the proposed procedure in a setting without regime dependence. We report the empirical rejection rates of the multiple-model comparison test for nonnested models in Table IA4. The conditional test exhibits good (albeit not size-adjusted) power in both states. But the test is consistently oversized in both states, which is not surprising. The unconditional test exhibits good size and power properties.

IA.3 Selecting the number of regimes

We follow a similar approach to Guidolin and Timmermann (2008) to select the number of regimes. In the first step, we conduct a regime-specification analysis, where we consider a range of values for the number of regimes, $L = 1, 2, 3$, and 4. To select the number of regimes, we consider the HQC and BIC criteria that trade off in-sample fit against parsimony.[1] Table IA5 reports HQC and BIC for a range of regime-switching models (that is, $L = 1, 2, 3$, and 4). Our results indicate that a two-state model specification emerges as the preferred one.

In the second step, we employ a likelihood ratio (LR) test to assess whether a more parsimonious model is supported by the data. Specifically, we test the null hypothesis of a single-state specification against the alternative of a two-state regime-switching model. Since the parameters of the two-state model are not identified under the null, we compute the p-value using a parametric bootstrap procedure similar to the one proposed by Carrasco et al. (2014). The bootstrap procedure is as follows:

1. Using the observed data, we estimate the parameters of both the single-state and two-state models and compute the LR statistic, denoted by LR_{obs}.

2. Given the estimated parameters of the single-state model, we generate B independent samples under the null hypothesis, assuming the true parameter values are equal to the estimates from the single-state model.

3. For each simulated sample, we re-estimate the LR statistic, denoted by $\text{LR}^{(b)}$, for $b = 1, \ldots, B$.

The bootstrap p-value is then calculated as

$$\frac{1}{B} \sum_{b=1}^{B} \mathbb{1} \left\{ \text{LR}^{(b)} > \text{LR}_{\text{obs}} \right\},$$

where $\mathbb{1}\{\cdot\}$ is the indicator function that equals 1 if the argument is true and 0 otherwise. We set $B = 10,000$.

We find that the value of the LR_{obs} statistic is 2141.713 with a p-value of 0.000. This suggests that the single-state model is strongly rejected by the LR test in favor of a two-regime specification. Note, however, that while this approach provides statistical support for the two-state specification, it is equally important to assess the practical implications of model complexity for estimation precision and numerical stability. In fact, given the number of factors considered,

[1] We abstain from using the Akaike information criterion since it tends to select overparameterized models (see Fenton and Gallant, 1996).

more complex regime-switching specifications may compromise both estimation precision and numerical robustness in model comparison. To illustrate this issue, Table IA6 reports several diagnostics for each regime across single-, two-, three-, and four-state specifications. Specifically, we report the estimated stationary probability, $\widehat{\pi}_\ell$, which reflects the long-run proportion of time the system is expected to spend in regime ℓ, and the ratio $T\widehat{\pi}_\ell/|\theta_\ell|$, which captures the effective number of observations per parameter and serves as a proxy for estimation reliability. We also include two standard measures of covariance conditioning. The first, $\text{rcond}(\Sigma_\ell)$, is an estimate of the reciprocal of the condition number of Σ_ℓ in the 1-norm; small values indicate near-singularity and potential numerical instability. The second, $\lambda_{\max,\ell}/\lambda_{\min,\ell}$, is the ratio of the largest to the smallest eigenvalue of Σ_ℓ, providing the exact condition number for symmetric positive definite matrices, with large values signaling ill-conditioning.

The diagnostic results reveal that the two-state model offers the most balanced and stable specification among the multi-regime alternatives. Both regimes have high stationary probabilities, ensuring sufficient data representation. The observation-to-parameter ratios are comfortably above one, indicating reliable parameter estimation. Moreover, the covariance conditioning measures take moderate values in both regimes, suggesting numerical stability. In contrast, the three- and four-state models show significant imbalance: one regime tends to dominate, while others suffer from low data support, and poor covariance conditioning. These issues raise concerns about overparameterization and estimation reliability in more complex specifications. Taken together, the two-state model strikes the best trade-off between capturing regime-dependent dynamics and maintaining estimation precision and numerical robustness.

IA.4 Bootstrap algorithm to account for state uncertainty

In what follows, the subscript ℓ is suppressed for notational simplicity. Let $\hat{\zeta}$ denote the value of the statistic of interest computed from the original sample. The parametric bootstrap proceeds as follows for $b = 1, \ldots, B$:

Step 1: Simulate a path of latent states, $\{s_t^{\star(b)}\}_{t=1}^T$, from the Markov chain with transition matrix $\hat{P}$ computed from the original sample.

Step 2: Conditional on the simulated state path, generate a bootstrap sample, $\{f_t^{\star(b)}\}_{t=1}^T$, from the regime-specific distribution with parameters $\hat{\mu}_{s_t^{\star(b)}}$ and $\hat{V}_{s_t^{\star(b)}}$.

Step 3: Re-estimate the Markov-switching model using the bootstrap sample, $\{f_t^{\star(b)}\}_{t=1}^T$, yielding bootstrap parameter estimates $\hat{\theta}^{\star(b)}$.

Step 4: Compute the bootstrap analogue of the statistic of interest, $\zeta^{\star(b)}$, using $\hat{\theta}^{\star(b)}$.

The bootstrap distribution of $\hat{\zeta}$ is approximated by the empirical distribution of $\{\zeta^{\star(b)}\}_{b=1}^{B}$, and bootstrap p-values are obtained by comparing $\hat{\zeta}$ to this distribution. To illustrate, consider the pairwise model comparison test and let Δ denote the difference in the conditional squared Sharpe ratios between two models. The test statistic is given by

$$\hat{\zeta} = \frac{\widehat{\Delta}}{s(\widehat{\Delta})},$$

where $\widehat{\Delta}$ is the estimate of Δ from the actual sample and $s(\widehat{\Delta})$ is its standard error. The corresponding bootstrap statistic is given by

$$\zeta^{\star(b)} = \frac{\Delta^{\star(b)} - \widehat{\Delta}}{s(\Delta^{\star(b)})},$$

where $\Delta^{\star(b)}$ is the estimate of Δ from the bth bootstrap sample. An analogous centering is applied to all other test statistics, including the likelihood ratio statistic used in the multiple-model comparison test, so that the bootstrap distribution reflects the null hypothesis. The bootstrap p-value is then computed as

$$\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left\{|\zeta^{\star(b)}| > |\hat{\zeta}|\right\}.$$

By resampling the latent state sequence using the estimated transition matrix, this procedure accounts for uncertainty in regime classification. The bootstrap is applied uniformly to all test statistics described in Section 2 of the main paper, and we set $B = 10{,}000$.

IA.5 Bias adjustment in squared Sharpe ratios

In this section, we assess the robustness of our main empirical findings to small-sample bias in squared Sharpe ratios. We consider both the unconditional (single-state) and conditional (regime-specific) cases.

In the single-state setting, Jobson and Korkie (1980) show that the squared Sharpe ratio estimator is upward biased in finite samples and provide an analytical correction. The bias is of order $O(1/T)$ and therefore vanishes asymptotically, implying that it does not affect the asymptotic distribution. However, in the regime-switching setting, an analogous analytical correction is not available due to the presence of latent states and regime-dependent parameters, making it unclear whether the bias affects the asymptotic distribution. Nevertheless, a bias-adjusted esti-

mator of the conditional squared Sharpe ratio can be constructed using the parametric bootstrap procedure described in Section IA.4 (Steps 1–3). To gain insight into the asymptotic behavior of the bias, we conducted a simple simulation exercise in which we examined how the bootstrap bias varies with sample size. Specifically, for all models considered in the paper, we regressed the logarithm of the estimated bias on the logarithm of the sample size. We found that, across all models, the estimated slopes are consistently close to one, suggesting that the bias most likely decays at a rate of approximately $O(1/T)$ and is therefore asymptotically negligible and unlikely to affect the asymptotic distribution. We acknowledge that this approach is admittedly simple. A more thorough analysis of the bias in the conditional squared Sharpe ratio would be of independent interest, but lies beyond the scope of the present paper.

Taken together, these results suggest that the bias is asymptotically negligible in both the single-state and regime-switching settings. We therefore recompute all model comparison tests using the bias-adjusted squared Sharpe ratios. The results, reported in Tables IA8 and IA9, remain qualitatively unchanged. In particular, the bias-adjusted differences and corresponding test statistics across the single-state, bull, and bear regimes are very similar to the unadjusted quantities reported in Tables 4 and 5 of the main paper. Although the use of bias-adjusted statistics leads to some changes in the associated p-values, these changes do not alter the overall conclusions.

IA.6 Out-of-sample comparison of squared Sharpe ratios

Thus far, our findings on factor model performance in the main paper come from in-sample analyses. But there is extensive evidence that the in-sample performance of a factor model tends to correlate poorly with its out-of-sample performance (see, among others, Kan et al., 2024; Simin, 2008). Hence, we also compare models based on their out-of-sample squared Sharpe ratios. The out-of-sample framework for comparing factor models is as follows.

Let $f_{At} \in \mathbb{R}^{K_A}$ denote the returns on the factors in model A at time t. Given a full sample of T observations, we choose an in-sample estimation window of length $\bar{T}$. For each $\bar{t} = \bar{T}, \ldots, T-1$, using all available data up until (and including) date $\bar{t}$, we update the maximum likelihood estimates of $\mu_{f,\ell}$, $V_{f,\ell}$, and P and forecast the most likely state for period $\bar{t}+1$ as

$$\hat{s}_{\bar{t}+1} = \operatorname*{argmax}_{\ell} \mathcal{P}(s_{\bar{t}+1} = \ell \mid F_{\bar{t}}; \vartheta),$$

where, by the law of total probability,

$$\mathcal{P}(s_{\bar{t}+1} = \ell \mid F_{\bar{t}}; \vartheta) = \sum_{\iota=1}^{L} \mathcal{P}(s_{\bar{t}} = \iota \mid F_{\bar{t}}; \vartheta) \times \hat{p}_{\iota\ell}.$$

Next, conditional on the predicted state $\hat{s}_{\bar{t}+1}$, we compute the one-step-ahead return on the tangency portfolio under model A as

$$r^{A}_{\max \bar{t}+1} = \mathfrak{w}'_{\bar{t}} f_{A\bar{t}+1},$$

where the portfolio weights are given by

$$\mathfrak{w}_{\bar{t}} = \frac{\hat{V}^{-1}_{A,\hat{s}_{\bar{t}+1}} \hat{\mu}_{A,\hat{s}_{\bar{t}+1}}}{\mathbf{1}'_{K_A} \hat{V}^{-1}_{A,\hat{s}_{\bar{t}+1}} \hat{\mu}_{A,\hat{s}_{\bar{t}+1}}}.$$

We implement this procedure for each competing factor model. The resulting series of out-of-sample portfolio returns is then used to compare models. We implicitly assume that the factors span the investment opportunity set and construct tangency portfolios using only the factors included in each model, conditional on the forecasted regime at each point in time. This design ensures that model performance is evaluated in a manner consistent with its pricing assumptions, and that portfolio construction remains state-contingent. Under the null hypothesis that a model is correctly specified and satisfies the spanning condition, the resulting tangency portfolio should be mean-variance efficient. Accordingly, we compare models based on the squared Sharpe ratios of their out-of-sample portfolio returns and assess their relative performance by examining pairwise differences in the squared Sharpe ratios.

Remark. This approach does not contradict with Proposition 1, which cautions against the use of unconditional Sharpe ratios under regime switching due to the "bias" introduced by cross-regime variability in factor means and covariances. By conditioning portfolio decisions on regime forecasts, our method avoids this pitfall and yields a realized return series that reflects a dynamic, regime-aware investment strategy. While a higher out-of-sample Sharpe ratio for one model relative to another may indicate a stronger ability to span the conditional mean-variance frontier implied by the factors, it should not be interpreted as conclusive evidence of superior pricing performance. The Sharpe ratio reflects not only the model's structural specification but also the success of regime prediction and parameter estimation. As such, it provides a useful measure of predictive and investment performance under model-implied pricing restrictions, but it does not directly test for mispricing.

To determine whether the differences in squared Sharpe ratios are statistically significant,

we employ the dependent wild bootstrap proposed by Shao (2010). The bootstrap proceeds as follows. Let $R_{\max,\bar{\tau}} = (r^1_{\max,\bar{\tau}}, \ldots, r^D_{\max,\bar{\tau}})'$ for $\bar{\tau} = 1, \ldots, \widetilde{T}$, where $\widetilde{T} = T - \bar{T}$ is the number of out-of-sample observations. In words, $R_{\max,\bar{\tau}}$ is a D-vector of portfolio returns at time $\bar{\tau}$, with each element corresponding to a portfolio constructed under a particular model. We assume that $r^1_{\max,\bar{\tau}}$ corresponds to the benchmark model.

Given the sample, $\mathcal{R}_{\max} \equiv \{R_{\max,\bar{\tau}}\}_{\bar{\tau}=1}^{\widetilde{T}}$, we generate B bootstrap samples of size $\widetilde{T}$ using

$$R^{\star}_{\max,\bar{\tau}} = \widetilde{R}_{\max} + \left(R_{\max,\bar{\tau}} - \bar{R}_{\max}\right) W_{\bar{\tau}},$$

where $\bar{R}_{\max} = \widetilde{T}^{-1} \sum_{\bar{t}=1}^{\widetilde{T}} R_{\max,\bar{t}}$ is the sample mean and $\widetilde{R}_{\max} \in \mathbb{R}^D$ is a vector used to impose the null that the squared Sharpe ratios of the portfolios under the competing models are equal to the squared Sharpe ratio of the portfolio under the benchmark model. Specifically, the first element of $\widetilde{R}_{\max}$ equals the sample mean of $r^1_{\max,\bar{\tau}}$, and for $d = 2, \ldots, D$, the d-th element is set equal to $\sqrt{\widehat{Sh}_1^{\,2}} \cdot \hat{\sigma}_d$, where $\widehat{Sh}_1^{\,2}$ is the squared Sharpe ratio of the benchmark portfolio, and $\hat{\sigma}_d$ is the sample standard deviation of the d-th portfolio's returns.

The sequence $\{W_{\bar{\tau}}\}_{\bar{\tau}=1}^{\bar{T}}$ consists of $\bar{T}$ random variables that are independent of $\mathcal{R}_{\max}$. More precisely, define $\Sigma_{\psi} = (k_{BT}((\bar{\tau} - \underline{\tau})/\psi))_{\bar{\tau},\underline{\tau}=1,\ldots,\widetilde{T}}$, where $k_{BT}(\cdot)$ is the Bartlett kernel and ψ is a bandwidth parameter. Then $(W_1, \ldots, W_{\widetilde{T}})' = \Sigma_{\psi}^{1/2} \Xi_{\widetilde{T}}$, where $\Xi_{\widetilde{T}} \sim MN(0_{\widetilde{T}}, I_{\widetilde{T}})$.

For each bootstrap sample $b = 1, \ldots, B$, we compute the squared Sharpe ratios for portfolios under each model based on the resampled returns, $\{R^{\star}_{\max,\bar{\tau}}\}_{\bar{\tau}=1}^{\widetilde{T}}$. We then calculate the differences in squared Sharpe ratios between each model $d = 2, \ldots, D$ and the benchmark model, by forming the bootstrap statistics

$$\Delta^{\star}_d = \widehat{Sh}_1^{\star 2} - \widehat{Sh}_d^{\star 2}, \qquad d = 2, \ldots, D.$$

The two-sided p-value for each d is then computed as

$$\frac{1}{B} \sum_{b=1}^{B} \mathbb{1}\left\{ |\Delta_d^{\star(b)}| > |\hat{\Delta}_d| \right\},$$

where $\hat{\Delta}_d = \widehat{Sh}_1^{\,2} - \widehat{Sh}_d^{\,2}$ is the observed squared Sharpe ratio difference from the original sample.

Note that in the above procedure, each model is compared against multiple competitors. This raises the issue of multiple testing: low p-values may arise simply by chance (or by capitalising on random noise), meaning that a model could appear to outperform the benchmark due to "luck" rather than genuine superiority. To address this issue, we use the bootstrap procedure described above to adjust for spurious rejections arising from multiple-model comparisons and data-snooping biases, in the spirit of the "reality check" method of White (2000). In particular,

we obtain a p-value for the smallest observed test statistic, so that the null and the alternative hypotheses are specified as

$$H_0 : \min_{d=2,\ldots,D} \Delta_d \geqslant 0$$

and

$$H_1 : \min_{d=2,\ldots,D} \Delta_d < 0,$$

respectively. The null hypothesis states that no competing model outperforms the benchmark, while the alternative is that at least one competitor has a squared Sharpe ratio exceeding the benchmark. This motivates the following choice of test statistic:

$$\mathcal{J}_{\widetilde{T}} = \min_{d=2,\ldots,D} \widehat{\Delta}_d. \tag{IA44}$$

To implement the test, we employ the resampled returns, $\{R^{\star}_{\max,\bar{\tau}}\}_{\bar{\tau}=1}^{\widetilde{T}}$, to form the bootstrap statistic, $\mathcal{J}^{\star}_{\widetilde{T}}$, and calculate the p-value as

$$\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left\{\mathcal{J}^{\star}_{\widetilde{T}} \leqslant \mathcal{J}_{\widetilde{T}}\right\}.$$

Following Avramov et al. (2023) and Barillas and Shanken (2017), the initial in-sample estimation window is chosen to be one-half of the monthly data (that is, $T/2$). This choice offers a trade-off between estimation precision and out-of-sample evaluation length: it ensures a sufficiently large sample for reliable parameter estimation, while preserving enough observations for meaningful out-of-sample evaluation.[2] We also set $B = 10,000$ and $\psi = 10$.

Table IA10 reports the differences in the out-of-sample squared Sharpe ratios of factor models, along with the p-values. The results show that most pairwise differences are small and statistically insignificant (that is, most p-values are well above 0.05). Moreover, the reality check results suggest that after accounting for multiple-model comparisons, no model outperforms the others, with statistical significance at the 5% nominal level. This aligns with our earlier in-sample findings that no model dominates in the bear regime. Also, the apparent in-sample performance gains seen in the single or bull states disappear out of sample.

[2] In a recursive framework, where the in-sample window expands gradually, starting with a shorter window may result in imprecise early estimates. These early estimation errors can adversely affect the initial out-of-sample performance, especially before the sample size becomes large enough to stabilize the regime-switching parameter estimates.

IA.7 Marginal contributions to the squared Sharpe ratios

In this section, we complement our analysis (in the main paper) on why some models outperform others in the bull regime, while all models perform similarly in the bear regime. To do so, we compute the marginal contribution of a factor to the squared Sharpe ratio of a given model. Based on the identity in Eq. (2) of the main paper, Fama and French (2018) define the marginal contribution of a factor to a model as the increase in $Sh^2(f)$ when that factor is added to the model's other factors. In particular,

$$\alpha_k^2/\text{RSE}_k^2 = Sh^2(f) - Sh^2(f_{-k}),$$

where α_k is the intercept from projecting the kth element of f onto the remaining elements f_{-k}, and RSE_k is the residual standard error from this regression. Accordingly, a factor's marginal contribution is small if the factor's α_k is close to zero, its variation not explained by other factors is large, or both (Fama and French, 2018). Note, however, that the marginal contributions of all K factors do not sum up to $Sh^2(f)$.

In Table IA11, we report the marginal contribution of each pricing factor k to the squared Sharpe ratio of the corresponding model. The results for the single regime in Panel A indicate that the market excess return factor, MKT, is the top marginal contributor to the squared Sharpe ratios of the FF5, FF6, and FF6_C models. The cash-based profitability factor, RMW_C, is by far the most important marginal contributor to the squared Sharpe ratio of the FF5_C model, followed by the MKT factor. On the other hand, the profitability factor, r_ROE, appears to be the top marginal contributor to the squared Sharpe ratios for both the HXZ4 and BS6 models. For the SY4 model, it can be seen that the management-related mispricing factor, MGMT, is the most important marginal contributor to the squared Sharpe ratio, followed by the MKT factor. We also find that the size-related factors, SMB, r_ME, and SMB_SY, the value factor, HML, and the investment factor, $r_\text{I/A}$ (in the BS6 model), do not contribute much to the squared Sharpe ratios of the corresponding models. The contributions of the remaining factors to the squared Sharpe ratios of the models are small.

Panel B summarizes the results for the bull regime, which are qualitatively similar to those of the single regime. One notable exception is the improvement in the marginal contribution of the size factor, SMB, to the squared Sharpe ratios of the FF5_C and FF6_C models. Another exception is the reduction in the marginal contribution of the investment factor, CMA, to the squared Sharpe ratios of the corresponding models. We report the results for the bear regime in Panel C of Table IA11. It can be seen that most factors do not contribute much to the squared

Sharpe ratios of their corresponding models. The notable exceptions are the investment factors, CMA and $r_{\mathrm{I/A}}$, and the management-related mispricing factor, MGMT. The relatively large marginal contributions of these factors are actually generating statistically significant squared Sharpe ratios for their corresponding models in the bear regime (see also Table 3 of the main paper). Overall, Table IA11 shows that marginal contributions vary substantially across regimes, highlighting the importance of regime-aware model evaluation. Certain factors (that is, MKT, r_{ROE}, $\mathrm{RMW}_{\mathrm{C}}$, and MGMT) tend to perform well in the bull market, driving model performance. But in the bear regime, models and most factors converge in performance, suggesting limited differentiation.

IA.8 Sharpe ratio decomposition

In this section, we present a decomposition of the Sharpe ratio for a tangency portfolio, which allows us to identify the key drivers of model performance. Consider a portfolio $\mathfrak{p}$ with weights $\mathfrak{w}$ that are given by the solution to the problem

$$\max_{\mathfrak{w}} Sh(f),$$

where

$$Sh(f) = \frac{\mu_{\mathfrak{p}}}{\sigma_{\mathfrak{p}}} = \frac{\mathfrak{w}'\mu_f}{\sqrt{\mathfrak{w}'V_f\mathfrak{w}}}.$$

Note that the gradient of $Sh(f)$ is zero at the optimal weights, that is,

$$\nabla_{\mathfrak{p}} Sh(f) = \frac{\mu_f\sigma_{\mathfrak{p}}^2 - \mu_{\mathfrak{p}}V_f\mathfrak{w}}{\sigma_{\mathfrak{p}}^3} = 0,$$

which yields

$$\mu\sigma_{\mathfrak{p}}^2 = \mu_{\mathfrak{p}}V_f\mathfrak{w}. \tag{IA45}$$

Taking the k-th component of Eq. (IA45) gives

$$\begin{aligned}\mu_k\sigma_{\mathfrak{p}}^2 &= \mu_{\mathfrak{p}}(V_f\mathfrak{w})_k \\ &= \mu_{\mathfrak{p}}Cov(f_{kt}, r_{\mathfrak{p}t}),\end{aligned}$$

where $r_{\mathfrak{p}t} = f_t\mathfrak{w}$. Since $Cov(f_{kt}, r_{\mathfrak{p}t}) = \rho_{k\mathfrak{p}}\sigma_k\sigma_{\mathfrak{p}}$, it follows that

$$\frac{\mu_{\mathfrak{p}}}{\sigma_{\mathfrak{p}}} = \frac{\mu_k}{\rho_{k\mathfrak{p}}\,\sigma_k}, \tag{IA46}$$

for every asset k. Defining $Sh(f_k) = \mu_k/\sigma_k$, Eq. (IA46) implies

$$Sh(f) = \frac{1}{\rho_{k\mathfrak{p}}} Sh(f_k). \tag{IA47}$$

Eq. (IA47) shows that the Sharpe ratio of any asset is proportional to the Sharpe ratio of the tangency portfolio, with proportionality coefficient given by its correlation with the portfolio. Furthermore,

$$\sigma_{\mathfrak{p}} = \sum_{k=1}^{K} \mathfrak{w}_k \rho_{k\mathfrak{p}} \sigma_k, \tag{IA48}$$

which follows directly from Euler's homogeneous function theorem (see also Darolles et al., 2015). Eq. (IA48) tells us that the total volatility of a portfolio is just the sum of each asset's (in our case, each factor's) weighted contribution to that volatility, where each contribution is equal to the product of the proportion of the portfolio invested in factor k, the factor's individual volatility, and the correlation of the factor with the tangency portfolio.

References

Ang, A., Timmermann, A., 2012. Regime changes and financial markets. Annual Review of Financial Economics 4, 313–337.

Avramov, D., Cheng, S., Metzker, L., Voigt, S., 2023. Integrating factor models. Journal of Finance 78, 1593–1646.

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

Barillas, F., Shanken, J., 2017. Which Alpha? Review of Financial Studies 30, 1316–1338.

Bickel, P. J., Ritov, Y., Ryden, T., 1998. Asymptotic normality of the maximum-likelihood estimator for general hidden Markov models. Annals of Statistics 26, 1614–1635.

Carrasco, M., Hu, L., Ploberger, W., 2014. Optimal test for Markov switching parameters. Econometrica 82, 765–784.

Cavicchioli, M., 2014. Analysis of the likelihood function for Markov-switching VAR(CH) models. Journal of Time Series Analysis 35, 624–639.

Darolles, S., Gouriéroux, C., Jay, E., 2015. Robust portfolio allocation with systematic risk contribution restrictions. In: Jurczenko, E. (Ed.), Risk-based and Factor Investing. Elsevier, Oxford, pp. 123–146.

Fama, E. F., French, K. R., 2018. Choosing factors. Journal of Financial Economics 128, 234–252.

Fenton, V. M., Gallant, A. R., 1996. Qualitative and asymptotic performance of SNP density estimators. Journal of Econometrics 74, 77–118.

Gibbons, M. R., Ross, S. A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 57, 1121–1152.

Guidolin, M., Timmermann, A., 2008. Size and value anomalies under regime shifts. Journal of Financial Econometrics 6, 1–48.

Hamilton, J. D., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics 45, 39–70.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press.

Jobson, D. J., Korkie, B., 1980. Estimation for Markowitz efficient portfolios. Journal of the American Statistical Association 75, 544–554.

Kan, R., Wang, X., Zheng, X., 2024. In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models. Journal of Financial Economics 155, 103837.

Kiefer, N. M., 1978. Discrete parameter variation: Efficient estimation of a switching regression model. Econometrica 46, 427–434.

Kim, C.-J., 1994. Dynamic linear models with Markov-switching. Journal of Econometrics 60, 1–22.

Leroux, B. G., 1992. Maximum-likelihood estimation for hidden Markov models. Stochastic Processes and their Applications 40, 127–143.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47, 13–37.

Magnus, J. R., Neudecker, H., 1979. The commutation matrix: Some properties and applications. Annals of Statistics 7, 381–394.

Psaradakis, Z., Sola, M., 1998. Finite-sample properties of the maximum likelihood estimator in autoregressive models with Markov switching. Journal of Econometrics 86, 369–386.

Shao, X., 2010. The dependent wild bootstrap. Journal of the American Statistical Association 105, 218–235.

Sharpe, W. F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance 19, 425–442.

Simin, T., 2008. The poor predictive performance of asset pricing models. Journal of Financial and Quantitative Analysis 43, 355–380.

Timmermann, A., 2000. Moments of Markov switching models. Journal of Econometrics 96, 75–111.

White, H., 2000. A reality check for data snooping. Econometrica 68, 1097–1126.

Table IA1
Properties of the pairwise model comparison tests

The table reports the rejection probability rates of the alpha-based (Panel A) and direct (Panel B) tests for the pairwise comparison of factor models. The rejection rates correspond to a sample size of $T = 600$ and are computed from 10,000 replications. The simulation setup for the alpha-based test is designed such that in state 2 the SMB, HML_m, UMD, r_{ROE}, and $r_{\text{I/A}}$ factors of the BS6 model (see also notes to Tables 1 and 2 in the main paper) are exactly priced by the MKT factor of the CAPM, whereas in state 1 the true squared Sharpe ratios of the models are equal to their sample values in state 1 and thus the CAPM is dominated by the BS6 model. The true squared Sharpe ratios of the CAPM and BS6 models in state 1 are 0.058 and 0.418, respectively. For the direct test, the simulation setup is designed such that in state 2, the true conditional squared Sharpe ratios of the FF5 and BS6 models are equal to 0.115 (the sample estimate of the conditional squared Sharpe ratio of the BS6 model in state 2), whereas in state 1, the true conditional squared Sharpe ratios of the models are equal to their sample values in state 1. To analyze the effect that the level of persistence in the Markov chain may have on the properties of the tests, four different pairs of transition probabilities are considered: $(p_{11}, p_{22}) = \{(0.95, 0.95), (0.95, 0.80), (0.80, 0.95), (0.50, 0.50)\}$. The table also reports the rejection rates of the unconditional testing procedure of Barillas et al. (2020) applied to the simulated data.

(p_{11}, p_{22})	Conditional tests		Unconditional tests
	State 1	State 2	
Panel A: Alpha-based tests			
$(0.95, 0.95)$	1.000	0.056	0.876
$(0.95, 0.80)$	1.000	0.091	1.000
$(0.80, 0.95)$	0.996	0.059	0.197
$(0.50, 0.50)$	1.000	0.088	0.933
Panel B: Direct tests			
$(0.95, 0.95)$	0.926	0.037	0.477
$(0.95, 0.80)$	0.994	0.023	0.980
$(0.80, 0.95)$	0.542	0.047	0.090
$(0.50, 0.50)$	0.925	0.045	0.474

Table IA2
Properties of the multiple-model comparison test

The table reports the rejection probability rates of the multiple-model comparison test for nonnested models. The simulation setup is designed such that in state 2, the true conditional squared Sharpe ratios of the FF5, HXZ4, BS6, and SY4 models are equal to 0.152 (which is the sample estimate of the conditional squared Sharpe ratio of the FF5 model in state 2), whereas in state 1, the true conditional squared Sharpe ratios of these models are set equal to their sample estimates in state 1. For a description of the factor models, see notes to Table 2 in the main paper. To analyze the effect that the number of models may have on the properties of the test, two different scenarios are considered: (i) FF5 versus {SY4, BS6} ($q = 2$); and (ii) FF5 versus {SY4, BS6, HXZ4} ($q = 3$). See also the notes to Table IA1.

(p_{11}, p_{22})	q	Conditional tests		Unconditional tests
		State 1	State 2	
$(0.95, 0.95)$	2	0.964	0.057	0.617
	3	0.953	0.052	0.585
$(0.95, 0.80)$	2	0.999	0.049	0.994
	3	0.998	0.039	0.992
$(0.80, 0.95)$	2	0.672	0.061	0.150
	3	0.619	0.057	0.137
$(0.50, 0.50)$	2	0.966	0.061	0.617
	3	0.959	0.055	0.583

Table IA3
Properties of the pairwise model comparison tests under the single-state DGP

The table reports the empirical rejection rates of the alpha-based (Panel A) and direct (Panel B) tests for the pairwise comparison of factor models under the single-state DGP. The rejection rates correspond to a sample size of $T = 600$ and are computed from 10,000 replications. To analyze the size of the alpha-based test, the simulation is designed such that the SMB, HML_m, UMD, r_{ROE}, and $r_{\text{I/A}}$ factors of the BS6 model are exactly priced by the MKT factor of the CAPM. To analyze the power of the alpha-based test, the population squared Sharpe ratios of the FF5 model and the CAPM are set equal to their respective sample estimates in state 1 from the two-state specification. To analyze the size of the direct test, the simulation is designed such that the population squared Sharpe ratios of the FF5 and BS6 models are equal to 0.115, which corresponds to the sample estimate of the conditional squared Sharpe ratio of the BS6 model in state 2 under the two-state model. As for the power of the direct test, the population squared Sharpe ratios of the FF5 and BS6 models are set equal to their respective sample estimates in state 1 from the two-state specification. The table also reports the rejection rates of the unconditional testing procedure of Barillas et al. (2020) applied to the simulated data.

Properties	Conditional tests		Unconditional tests
	State 1	State 2	
Panel A: Alpha-based tests			
Size	0.681	0.682	0.049
Power	0.999	0.999	1.000
Panel B: Direct tests			
Size	0.233	0.238	0.045
Power	0.800	0.792	0.999

Table IA4
Properties of the multiple-model comparison test under the single-state DGP

The table reports the empirical rejection rates of the multiple-model comparison test for nonnested models under the single-state DGP. The simulation design is adopted for the multiple-model comparison among the FF5, HXZ4, BS6, and SY4 models. Under the null, the population squared Sharpe ratios of all four models are set equal to 0.152, which is the sample estimate of the conditional squared Sharpe ratio of the FF5 model in state 2 under the two-state model. Under the alternative, the population squared Sharpe ratios are set equal to their respective sample estimates in state 1. See also the notes to Table IA3.

Properties	q	Conditional tests		Unconditional tests
		State 1	State 2	
Size	2	0.200	0.204	0.055
	3	0.183	0.191	0.050
Power	2	0.880	0.878	1.000
	3	0.866	0.864	1.000

Table IA5
Diagnostics for model specification

The table reports diagnostics for the multivariate regime-switching models in Eq. (5) of the main paper. $\mathcal{L}(\vartheta)$ is the log-likelihood function, HQC is the Hannan-Quinn information criterion, and BIC is the Bayesian information criterion. The sample period is from January 1972 to December 2021.

Model	Number of parameters	$\mathcal{L}(\vartheta)$	HQC	BIC
1-state	119	−15969.78	32381.25	32700.80
2-state	240	−14898.93	30688.65	31333.12
3-state	363	−14739.61	30826.54	31801.31
4-state	488	−14632.75	31076.79	32387.21

Table IA6
Diagnostics for numerical stability

The table reports diagnostics for each regime ℓ in the single-state, two-state, three-state, and four-state models. $\hat{\pi}_\ell$ is the estimated stationary probability of the regime, which can also be interpreted as the long-run proportion of time the system spends in regime ℓ. The ratio, $T\hat{\pi}_\ell/|\theta_\ell|$, reflects the effective number of observations per parameter in regime ℓ, where $|\theta_\ell|$ is the cardinality of θ_ℓ, representing the number of parameters estimated in that regime. rcond(Σ_ℓ) is an estimate of the reciprocal of the condition number of Σ_ℓ in the 1-norm and $\lambda_{\max,\ell}/\lambda_{\min\ell}$ is the ratio of the largest to the smallest eigenvalue of the covariance matrix Σ_ℓ, indicating the degree of ill-conditioning. The sample period is from January 1972 to December 2021.

State	$\hat{\pi}_\ell$	$T\hat{\pi}_\ell/\lvert\theta_\ell\rvert$	rcond(Σ_ℓ)	$\lambda_{\max,\ell}/\lambda_{\min,\ell}$
Single-state model				
1	1.000	5.042	1.72×10^{-3}	291.614
Two-state model				
1	0.784	3.955	2.23×10^{-3}	221.390
2	0.216	1.087	1.06×10^{-3}	494.813
Three-state model				
1	0.105	0.528	1.89×10^{-4}	2454.863
2	0.755	3.805	2.39×10^{-3}	217.622
3	0.141	0.710	8.15×10^{-4}	629.736
Four-state model				
1	0.074	0.372	8.04×10^{-4}	567.581
2	0.119	0.602	6.03×10^{-4}	695.401
3	0.723	3.646	2.97×10^{-3}	225.960
4	0.084	0.421	3.58×10^{-4}	1365.387

Table IA7
Correlations

The table reports the correlations between the pricing factors in the single state (Panel A), bull state (Panel B), and bear state (Panel C). The sample period is from January 1972 to December 2021. See also the notes to Table 1 in the main paper for a description of the factors.

Panel A: Single state

	SMB	HML	HML_m	UMD	RMW	CMA	r_{ME}	r_{ROE}	$r_{I/A}$	RMW_C	SMB_{SY}	MGMT	PERF
MKT	0.250	−0.218	−0.047	−0.176	−0.209	−0.374	0.255	−0.215	−0.341	−0.369	0.240	−0.482	−0.262
SMB		−0.003	0.080	−0.074	−0.379	−0.026	0.973	−0.413	−0.098	−0.462	0.941	−0.244	−0.166
HML			0.778	−0.222	0.141	0.664	0.030	−0.124	0.654	0.216	0.064	0.728	−0.328
HML_m				−0.664	0.010	0.486	0.067	−0.443	0.481	0.035	0.091	0.489	−0.630
UMD					0.071	−0.015	−0.034	0.487	0.003	0.121	−0.038	0.002	0.713
RMW						0.042	−0.376	0.653	0.127	0.862	−0.309	0.270	0.420
CMA							−0.003	−0.044	0.902	0.192	0.021	0.740	−0.059
r_{ME}								−0.336	−0.076	−0.464	0.927	−0.227	−0.149
r_{ROE}									0.051	0.607	−0.339	0.103	0.631
$r_{I/A}$										0.246	−0.038	0.733	−0.061
RMW_C											−0.380	0.390	0.428
SMB_{SY}												−0.161	−0.120
MGMT													−0.031

Panel B: Bull state

	SMB	HML	HML_m	UMD	RMW	CMA	r_{ME}	r_{ROE}	$r_{I/A}$	RMW_C	SMB_{SY}	MGMT	PERF
MKT	0.188	−0.311	−0.255	−0.030	−0.079	−0.322	0.234	0.064	−0.295	−0.327	0.173	−0.456	−0.117
SMB		−0.149	−0.089	−0.021	−0.296	−0.109	0.970	−0.184	−0.101	−0.445	0.933	−0.360	−0.101
HML			0.873	−0.157	−0.256	0.686	−0.133	−0.352	0.625	0.050	−0.090	0.701	−0.518
HML_m				−0.450	−0.278	0.642	−0.102	−0.499	0.545	−0.018	−0.065	0.602	−0.662
UMD					0.119	−0.063	0.021	0.391	0.030	0.123	−0.002	−0.034	0.603
RMW						−0.266	−0.319	0.562	−0.197	0.758	−0.270	−0.020	0.466
CMA							−0.132	−0.349	0.914	0.049	−0.042	0.749	−0.286
r_{ME}								−0.105	−0.111	−0.460	0.918	−0.377	−0.101
r_{ROE}									−0.185	0.428	−0.129	−0.176	0.574
$r_{I/A}$										0.087	−0.028	0.710	−0.219
RMW_C											−0.400	0.300	0.360
SMB_{SY}												−0.266	−0.075
MGMT													−0.237

Panel C: Bear state

	SMB	HML	HML_m	UMD	RMW	CMA	r_{ME}	r_{ROE}	$r_{I/A}$	RMW_C	SMB_{SY}	MGMT	PERF
MKT	0.327	−0.130	0.117	−0.298	−0.301	−0.405	0.293	−0.424	−0.372	−0.405	0.321	−0.499	−0.377
SMB		0.104	0.175	−0.095	−0.442	0.024	0.975	−0.561	−0.116	−0.489	0.949	−0.152	−0.215
HML			0.727	−0.249	0.366	0.646	0.149	0.016	0.680	0.314	0.189	0.748	−0.205
HML_m				−0.753	0.140	0.368	0.159	−0.406	0.434	0.054	0.188	0.416	−0.623
UMD					0.051	0.051	−0.053	0.521	0.011	0.127	−0.051	0.038	0.777
RMW						0.245	−0.421	0.697	0.355	0.910	−0.346	0.458	0.399
CMA							0.088	0.188	0.890	0.292	0.062	0.734	0.111
r_{ME}								−0.479	−0.065	−0.482	0.936	−0.108	−0.184
r_{ROE}									0.240	0.700	−0.485	0.297	0.666
$r_{I/A}$										0.365	−0.066	0.755	0.062
RMW_C											−0.383	0.455	0.464
SMB_{SY}												−0.071	−0.157
MGMT													0.119

Table IA8
Pairwise tests of equality of (bias-adjusted) squared Sharpe ratios

The table reports pairwise tests of equality of the squared Sharpe ratios of seven different factor models in the single state (Panel A), bull state (Panel B), and bear state (Panel C). We report the difference between the (bias-adjusted) sample squared Sharpe ratios of the factor models in row i and column j, denoted by $\widehat{Sh}_i^2 - \widehat{Sh}_j^2$, and the corresponding asymptotic p-value (in parentheses) for the test of H_0: $Sh_i^2 = Sh_j^2$. The bias adjustment in Panel A is based on Jobson and Korkie (1980), whereas the bias adjustments in Panels B and C are based on a bootstrap approach. The number of bootstrap replications is set to 10,000. The sample period is from January 1972 to December 2021. See also the notes to Table 2 in the main paper for a description of the models.

Model	FF6	$FF5_C$	$FF6_C$	HXZ4	BS6	SY4
Panel A: Single state						
FF5	−0.029	−0.047	−0.070	−0.045	−0.093	−0.080
	(0.000)	(0.001)	(0.001)	(0.037)	(0.001)	(0.003)
FF6		−0.018	−0.041	−0.016	−0.063	−0.051
		(0.381)	(0.002)	(0.413)	(0.005)	(0.021)
$FF5_C$			−0.023	0.002	−0.046	−0.033
			(0.000)	(0.953)	(0.151)	(0.260)
$FF6_C$				0.025	−0.022	−0.010
				(0.305)	(0.400)	(0.696)
HXZ4					−0.047	−0.035
					(0.016)	(0.231)
BS6						0.012
						(0.686)
Panel B: Bull state						
FF5	−0.117	−0.076	−0.178	−0.100	−0.222	−0.175
	(0.000)	(0.006)	(0.000)	(0.030)	(0.000)	(0.001)
FF6		0.041	−0.061	0.016	−0.105	−0.058
		(0.347)	(0.021)	(0.731)	(0.013)	(0.228)
$FF5_C$			−0.102	−0.024	−0.146	−0.099
			(0.000)	(0.636)	(0.014)	(0.075)
$FF6_C$				0.077	−0.044	0.003
				(0.144)	(0.352)	(0.955)
HXZ4					−0.122	−0.074
					(0.001)	(0.164)
BS6						0.047
						(0.444)
Panel C: Bear state						
FF5	0.004	−0.026	−0.024	0.030	0.044	0.053
	(0.373)	(0.252)	(0.410)	(0.471)	(0.347)	(0.360)
FF6		−0.030	−0.028	0.026	0.040	0.049
		(0.297)	(0.235)	(0.573)	(0.412)	(0.449)
$FF5_C$			0.001	0.056	0.070	0.078
			(0.289)	(0.287)	(0.205)	(0.207)
$FF6_C$				0.054	0.068	0.077
				(0.345)	(0.237)	(0.271)
HXZ4					0.014	0.022
					(0.393)	(0.611)
BS6						0.008
						(0.849)

Table IA9
Multiple-model comparison tests of (bias-adjusted) squared Sharpe ratios

The table reports multiple-model comparison tests of the (bias-adjusted) squared Sharpe ratios of seven different factor models in the single state (Panel A), bull state (Panel B), and bear state (Panel C). q is the number of alternative models in each multiple nonnested model comparison; LR is the value of the likelihood ratio statistic with corresponding asymptotic p-value reported in the adjacent column; q_N is the number of factor models that nest the benchmark model; and $\widehat{Sh}^2_{EM} - \widehat{Sh}^2$ is the difference between the (bias-adjusted) sample squared Sharpe ratios of the expanded model (EM) and the benchmark model, with associated asymptotic p-value (in parentheses) reported in the adjacent column. The bias adjustment in Panel A is based on Jobson and Korkie (1980), whereas the bias adjustments in Panels B and C are based on a bootstrap approach. The number of bootstrap replications is set to 10,000. The sample period is from January 1972 to December 2021. See also the notes to Table 2 in the main paper for a description of the models.

Benchmark	q	LR	p-value	q_N	$\widehat{Sh}^2_{EM} - \widehat{Sh}^2$	p-value
Panel A: Single state						
FF5	4	16.327	(0.000)	1	0.029	(0.000)
FF6	4	21.292	(0.000)			
FF5_C	4	2.264	(0.170)	1	0.023	(0.000)
FF6_C	4	0.725	(0.449)			
HXZ4	4	6.113	(0.029)			
BS6	4	0.000	(0.768)			
SY4	4	0.163	(0.561)			
Panel B: Bull state						
FF5	4	22.489	(0.000)	1	0.117	(0.000)
FF6	4	11.171	(0.004)			
FF5_C	4	6.678	(0.020)	1	0.102	(0.000)
FF6_C	4	0.866	(0.425)			
HXZ4	4	11.324	(0.002)			
BS6	4	0.000	(0.790)			
SY4	4	0.586	(0.412)			
Panel C: Bear state						
FF5	4	0.677	(0.516)	1	−0.004	(0.373)
FF6	4	1.408	(0.334)			
FF5_C	4	0.000	(0.770)	1	−0.001	(0.289)
FF6_C	4	0.000	(0.731)			
HXZ4	4	0.893	(0.478)			
BS6	4	1.860	(0.279)			
SY4	4	1.210	(0.256)			

Table IA10
Out-of-sample tests of differences in squared Sharpe ratios

The table reports the differences in the squared Sharpe ratios of seven different factor models evaluated using an out-of-sample scheme. We report the difference between the out-of-sample squared Sharpe ratios of the factor models in row i and column j, denoted by $\widehat{Sh}_i^2 - \widehat{Sh}_j^2$, and the corresponding p-value (in parentheses) computed using the dependent wild bootstrap proposed by Shao (2010). The initial in-sample estimation window is chosen to be half of the full sample ($T/2$). The number of bootstrap replications is set to 10,000 and the value of the bandwidth parameter is set to 10. The column labeled "RC" reports the reality check statistics and the corresponding p-values (in parentheses). See also the notes to Table 2 in the main paper for a description of the models.

Model	FF6	$FF5_C$	$FF6_C$	HXZ4	BS6	SY4	RC
FF5	0.003	−0.020	−0.008	−0.039	−0.022	−0.062	−0.062
	(0.868)	(0.309)	(0.690)	(0.230)	(0.323)	(0.178)	(0.147)
FF6		−0.023	−0.010	−0.042	−0.025	−0.065	−0.065
		(0.070)	(0.244)	(0.031)	(0.072)	(0.016)	(0.056)
$FF5_C$			0.012	−0.020	−0.002	−0.043	−0.043
			(0.588)	(0.585)	(0.954)	(0.269)	(0.778)
$FF6_C$				−0.032	−0.014	−0.055	−0.055
				(0.434)	(0.738)	(0.295)	(0.424)
HXZ4					0.017	−0.023	−0.023
					(0.263)	(0.273)	(0.915)
BS6						−0.040	−0.040
						(0.067)	(0.802)
SY4							0.023
							(0.985)

Table IA11
Marginal contributions to the squared Sharpe ratios

The table reports estimates of alphas ($\hat{\alpha}$) and residual standard errors (RSE) from spanning regressions of each factor from the set of factors that constitute a given model on the model's other factors. The table also reports marginal contribution of each factor k to a model's squared Sharpe ratio, that is, $\hat{\alpha_k}^2/\text{RSE}_k^2$. Panel A summarizes the results for the single state, Panel B summarizes the results for the bull state, and Panel C summarizes the results for the bear state. Numbers in parentheses are the p-values for the test of $H_0 : \alpha_k = 0$. The sample period is from January 1972 to December 2021. See also the notes to Tables 1 and 2 in the main paper for a description of the factors and the models, respectively.

Model		MKT	SMB	HML	HML_m	UMD	RMW	CMA	r_ME	r_ROE	$r_\text{I/A}$	RMW_C	SMB_SY	MGMT	PERF
						Panel A: Single state									
FF5	$\hat{\alpha_k}$	0.90	0.20	−0.12			0.40	0.27							
		(0.000)	(0.085)	(0.225)			(0.000)	(0.000)							
	RSE_k	4.03	2.73	2.22			2.08	1.37							
	$\hat{\alpha_k}^2/\text{RSE}_k^2$	0.050	0.005	0.003			0.037	0.040							
FF6	$\hat{\alpha_k}$	1.00	0.20	0.00		0.73	0.37	0.24							
		(0.000)	(0.093)	(0.969)		(0.000)	(0.000)	(0.000)							
	RSE_k	3.97	2.73	2.12		4.09	2.07	1.35							
	$\hat{\alpha_k}^2/\text{RSE}_k^2$	0.063	0.005	0.000		0.032	0.031	0.031							
FF5_C	$\hat{\alpha_k}$	1.05	0.34	−0.14				0.25				0.45			
		(0.000)	(0.003)	(0.140)				(0.000)				(0.000)			
	RSE_k	3.94	2.64	2.23				1.38				1.57			
	$\hat{\alpha_k}^2/\text{RSE}_k^2$	0.071	0.017	0.004				0.033				0.084			
FF6_C	$\hat{\alpha_k}$	1.12	0.33	−0.03		0.65		0.22				0.42			
		(0.000)	(0.004)	(0.747)		(0.000)		(0.000)				(0.000)			
	RSE_k	3.89	2.64	2.13		4.08		1.37				1.56			
	$\hat{\alpha_k}^2/\text{RSE}_k^2$	0.083	0.016	0.000		0.026		0.026				0.073			
HXZ4	$\hat{\alpha_k}$	0.94							0.33	0.65	0.44				
		(0.000)							(0.006)	(0.000)	(0.000)				
	RSE_k	4.10							2.83	2.47	1.80				
	$\hat{\alpha_k}^2/\text{RSE}_k^2$	0.053							0.014	0.069	0.059				
BS6	$\hat{\alpha_k}$	1.01	0.29		0.43	0.52				0.51	0.08				
		(0.000)	(0.014)		(0.000)	(0.000)				(0.000)	(0.219)				
	RSE_k	4.08	2.66		2.00	2.76				2.01	1.38				
	$\hat{\alpha_k}^2/\text{RSE}_k^2$	0.061	0.012		0.046	0.036				0.064	0.003				
SY4	$\hat{\alpha_k}$	1.13											0.33	0.83	0.99
		(0.000)											(0.007)	(0.000)	(0.000)
	RSE_k	3.73											2.77	2.51	3.60
	$\hat{\alpha_k}^2/\text{RSE}_k^2$	0.092											0.014	0.109	0.076

(Continued)

Table IA11 – *Continued*

Model		MKT	SMB	HML	HML_m	UMD	RMW	CMA	r_{ME}	r_{ROE}	$r_{I/A}$	RMW_C	SMB_{SY}	MGMT	PERF
						Panel B: Bull state									
FF5	$\hat{\alpha_k}$	1.07	0.21	0.18			0.40	0.14							
		(0.000)	(0.069)	(0.029)			(0.000)	(0.017)							
	RSE_k	3.27	2.15	1.57			1.25	1.08							
	$\hat{\alpha_k}^2/RSE_k^2$	0.108	0.010	0.014			0.104	0.016							
FF6	$\hat{\alpha_k}$	1.14	0.22	0.26		0.91	0.37	0.11							
		(0.000)	(0.084)	(0.003)		(0.000)	(0.000)	(0.070)							
	RSE_k	3.27	2.15	1.55		2.60	1.25	1.08							
	$\hat{\alpha_k}^2/RSE_k^2$	0.121	0.010	0.029		0.121	0.089	0.011							
$FF5_C$	$\hat{\alpha_k}$	1.25	0.35	0.14				0.10				0.44			
		(0.000)	(0.002)	(0.112)				(0.122)				(0.000)			
	RSE_k	3.17	2.07	1.59				1.10				1.03			
	$\hat{\alpha_k}^2/RSE_k^2$	0.156	0.029	0.008				0.008				0.181			
$FF6_C$	$\hat{\alpha_k}$	1.28	0.34	0.22		0.84		0.08				0.40			
		(0.000)	(0.006)	(0.018)		(0.000)		(0.249)				(0.000)			
	RSE_k	3.17	2.07	1.57		2.59		1.10				1.03			
	$\hat{\alpha_k}^2/RSE_k^2$	0.164	0.027	0.020		0.106		0.005				0.150			
HXZ4	$\hat{\alpha_k}$	0.90							0.13	0.69	0.42				
		(0.000)							(0.323)	(0.000)	(0.000)				
	RSE_k	3.33							2.26	1.62	1.45				
	$\hat{\alpha_k}^2/RSE_k^2$	0.073							0.003	0.184	0.085				
BS6	$\hat{\alpha_k}$	1.06	0.22		0.46	0.64				0.56	0.06				
		(0.000)	(0.102)		(0.000)	(0.000)				(0.000)	(0.391)				
	RSE_k	3.33	2.21		1.50	2.16				1.36	1.19				
	$\hat{\alpha_k}^2/RSE_k^2$	0.101	0.010		0.095	0.088				0.169	0.003				
SY4	$\hat{\alpha_k}$	1.38											0.41	0.87	1.04
		(0.000)											(0.001)	(0.000)	(0.000)
	RSE_k	3.07											2.17	1.85	2.42
	$\hat{\alpha_k}^2/RSE_k^2$	0.203											0.036	0.224	0.184

(Continued)

Table IA11 – *Continued*

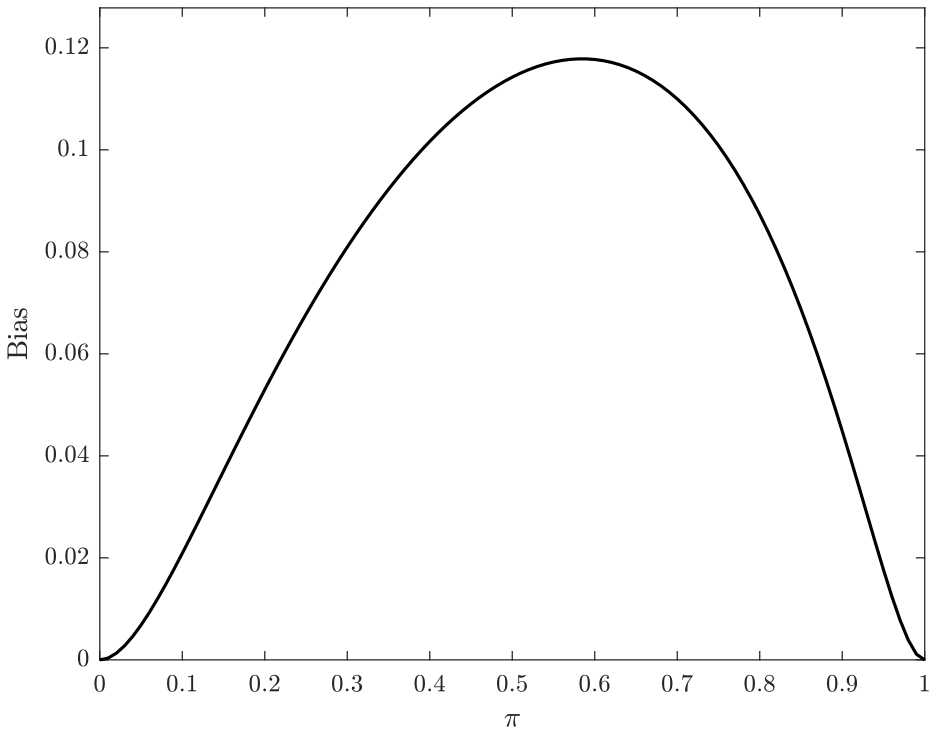
Model		MKT	SMB	HML	HML_m	UMD	RMW	CMA	r_ME	r_ROE	$r_\text{I/A}$	RMW_C	SMB_SY	MGMT	PERF
							Panel C: Bear state								
FF5	$\hat{\alpha}_k$	0.60	0.52	−0.48			0.36	0.65							
		(0.298)	(0.178)	(0.149)			(0.277)	(0.000)							
	RSE_k	5.78	3.87	3.37			3.28	1.99							
	$\hat{\alpha}_k{}^2/\text{RSE}_k^2$	0.011	0.018	0.020			0.012	0.106							
FF6	$\hat{\alpha}_k$	0.45	0.54	−0.51		−0.59	0.38	0.66							
		(0.420)	(0.158)	(0.099)		(0.402)	(0.238)	(0.000)							
	RSE_k	5.61	3.86	3.15		6.98	3.26	1.96							
	$\hat{\alpha}_k{}^2/\text{RSE}_k^2$	0.007	0.020	0.027		0.007	0.014	0.112							
FF5_C	$\hat{\alpha}_k$	0.75	0.68	−0.58				0.63				0.51			
		(0.192)	(0.075)	(0.101)				(0.001)				(0.046)			
	RSE_k	5.68	3.83	3.49				2.00				2.55			
	$\hat{\alpha}_k{}^2/\text{RSE}_k^2$	0.017	0.032	0.027				0.101				0.039			
FF6_C	$\hat{\alpha}_k$	0.59	0.71	−0.62		−0.70		0.65				0.53			
		(0.297)	(0.062)	(0.058)		(0.323)		(0.000)				(0.034)			
	RSE_k	5.55	3.81	3.25		6.95		1.97				2.53			
	$\hat{\alpha}_k{}^2/\text{RSE}_k^2$	0.011	0.035	0.036		0.010		0.109				0.045			
HXZ4	$\hat{\alpha}_k$	0.24							0.63	0.13	0.72				
		(0.672)							(0.112)	(0.727)	(0.002)				
	RSE_k	5.94							4.18	3.86	2.60				
	$\hat{\alpha}_k{}^2/\text{RSE}_k^2$	0.002							0.023	0.001	0.076				
BS6	$\hat{\alpha}_k$	0.16	0.54		0.16	0.02				0.24	0.28				
		(0.775)	(0.117)		(0.572)	(0.962)				(0.417)	(0.110)				
	RSE_k	5.82	3.61		2.93	3.97				3.06	1.85				
	$\hat{\alpha}_k{}^2/\text{RSE}_k^2$	0.001	0.023		0.003	0.000				0.006	0.023				
SY4	$\hat{\alpha}_k$	0.34											0.64	0.81	0.68
		(0.496)											(0.092)	(0.023)	(0.230)
	RSE_k	5.34											4.09	3.89	6.03
	$\hat{\alpha}_k{}^2/\text{RSE}_k^2$	0.004											0.025	0.043	0.013

Table IA12
Sharpe ratio decomposition of tangency portfolios based on factor models for the single state

The table reports the Sharpe ratio decomposition of the tangency portfolios based on seven different factor models in the single state. $\mu_{\mathfrak{p}}$ is the expected return on the portfolio $\mathfrak{p}$, $\sigma_{\mathfrak{p}}$ is the portfolio volatility, Sh is the Sharpe ratio of the portfolio, $Sh(f_k)$ is the Sharpe ratio of factor k, $\rho_{k\mathfrak{p}}$ is the correlation of the factor k with the portfolio $\mathfrak{p}$, $\mathfrak{w}_k$ is the weight (%) of the factor k in the portfolio $\mathfrak{p}$, and $\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$ is the weighted marginal contribution of the factor k to the portfolio volatility. The sample period is from January 1972 to December 2021. See also the notes to Tables 1 and 2 in the main paper for a description of the factors and the models, respectively.

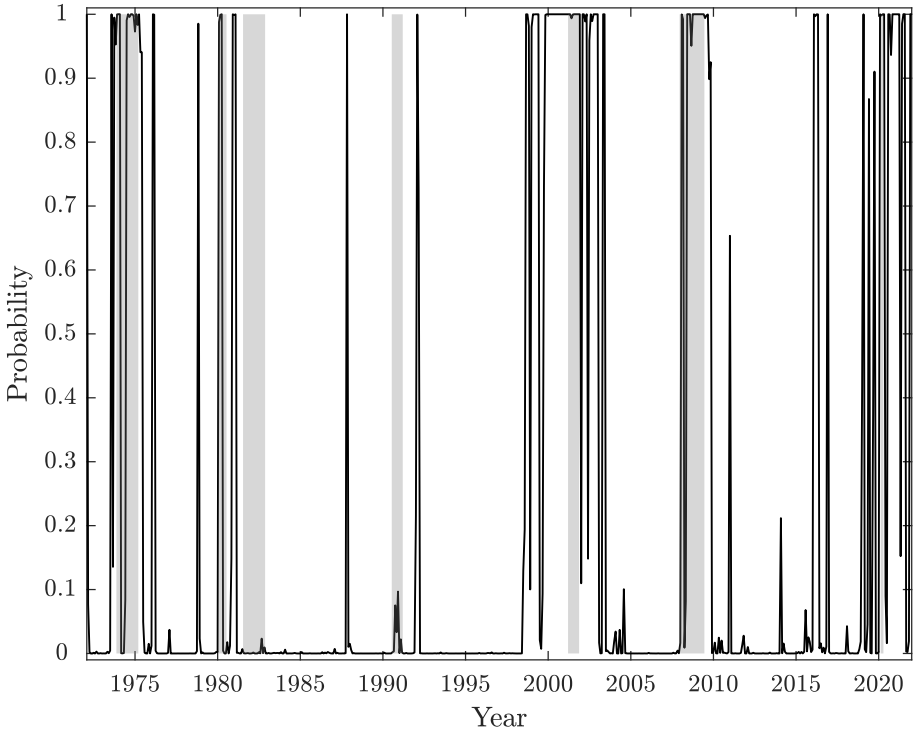
Model	$\mu_{\mathfrak{p}}$	$\sigma_{\mathfrak{p}}$	Sh		MKT	SMB	HML	HML_m	UMD	RMW	CMA	r_ME	r_ROE	$r_\text{I/A}$	RMW_C	SMB_SY	MGMT	PERF
FF5	0.346	1.079	0.320	$Sh(f_k)$	0.138	0.057	0.092			0.130	0.148							
				$\rho_{k\mathfrak{p}}$	0.431	0.178	0.286			0.406	0.463							
				$\mathfrak{w}_k$	19	9	−8			31	49							
				$\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$	0.367	0.049	−0.068			0.290	0.441							
FF6	0.385	1.051	0.366	$Sh(f_k)$	0.138	0.057	0.092		0.142	0.130	0.148							
				$\rho_{k\mathfrak{p}}$	0.377	0.156	0.250		0.389	0.355	0.405							
				$\mathfrak{w}_k$	18	8	0		12	24	37							
				$\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$	0.311	0.036	0.002		0.211	0.200	0.291							
FF5_C	0.370	0.955	0.388	$Sh(f_k)$	0.138	0.057	0.092				0.148				0.198			
				$\rho_{k\mathfrak{p}}$	0.357	0.147	0.237				0.383				0.510			
				$\mathfrak{w}_k$	17	12	−7				33				46			
				$\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$	0.270	0.054	−0.051				0.243				0.439			
FF6_C	0.394	0.940	0.419	$Sh(f_k)$	0.138	0.057	0.092		0.142		0.148				0.198			
				$\rho_{k\mathfrak{p}}$	0.330	0.136	0.219		0.340		0.353				0.471			
				$\mathfrak{w}_k$	17	11	−2		9		27				39			
				$\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$	0.248	0.044	−0.010		0.131		0.183				0.346			
HXZ4	0.432	1.127	0.383	$Sh(f_k)$	0.138							0.079	0.198	0.177				
				$\rho_{k\mathfrak{p}}$	0.361							0.205	0.516	0.461				
				$\mathfrak{w}_k$	16							12	31	40				
				$\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$	0.270							0.078	0.429	0.351				
BS6	0.448	1.007	0.445	$Sh(f_k)$	0.138	0.057		0.080	0.142				0.198	0.177				
				$\rho_{k\mathfrak{p}}$	0.310	0.128		0.180	0.320				0.444	0.396				
				$\mathfrak{w}_k$	14	9		24	15				28	9				
				$\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$	0.194	0.036		0.160	0.216				0.333	0.068				
SY4	0.547	1.283	0.427	$Sh(f_k)$	0.138											0.115	0.177	0.167
				$\rho_{k\mathfrak{p}}$	0.324											0.270	0.415	0.392
				$\mathfrak{w}_k$	25											13	40	23
				$\mathfrak{w}_k\rho_{k\mathfrak{p}}\sigma_k$	0.361											0.100	0.480	0.343

Figure IA1
Bias as a function of regime likelihood



The figure plots the last term in Eq. (12) of the main paper, $\gamma' \tilde{V}_{r \cdot f}^{-1} \gamma$, as a function of π_1 when mispricing in each regime is precluded.

Figure IA2
Probability of the bear state



This figure plots the probability of being in the bear state from January 1972 through December 2021. Shaded areas represent recessionary periods as defined by the National Bureau of Economic Research.