KING'S BUSINESS SCHOOL

Data Analytics for Finance & Macro Research Centre

KING'S College LONDON

Testing for Regime Changes in Portfolios with a Large Number of Assets: A Robust Approach to Factor Heteroskedasticity

D. Massacci

Working paper No. 2020/1 | July 2020 | ISSN 2516-5933

Testing for Regime Changes in Portfolios with a Large Number of Assets: A Robust Approach to Factor Heteroskedasticity*

Daniele Massacci

King's College London

July 23, 2020

Abstract

We develop a new test for threshold-type regime changes in the risk exposures in portfolios with a large number of financial assets whose returns exhibit an approximate factor structure. Unlike existing procedures to detect discrete shifts in factor models, our test is robust to regime-specific second moment of the common factors. We rely on an auxiliary threshold regression: we take a weighted cross-sectional average of the cross-sectional units; we estimate the factors from the original model under the null hypothesis of no regime changes; we construct a Lagrange multiplier statistic to test for threshold effect in the auxiliary regression. Numerical results show the good finite sample properties of our procedure. The empirical analysis uncovers the dynamics of portfolio weights and diversification benefits in factor mimicking portfolios across different regimes.

JEL classification: C12, C38, G11.

Keywords: Large Factor Model, Portfolio Choice, Threshold Model, Linearity Testing, Principal Component Analysis.

*This paper benefits from comments from participants at the 11th Annual SoFiE Conference in Lugano, the 2018 London Empirical Asset Pricing Workshop, and the 4thVienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance 2019, and at the seminar series at King's College London Business School, Cass Business School, Aston Business School, Royal Holloway University of London, University of Manchester, and University of Surrey; and from conversations with Matteo Barigozzi, Daniele Bianchi, Valentina Corradi, Pasquale Della Corte, Marco Lippi, Hashem Pesaran, Ivan Petrella, Mirco Rubin, Dario Ruzzi, Andrea Tamoni, and Lorenzo Trapani. Errors and omissions are the author's responsibility. Address correspondence to Daniele Massacci, King's Business School, Bush House, 30 Aldwych, London, WC2B 4BG, United Kingdom. E-mail: daniele.massacci@kcl.ac.uk .

1 Introduction

Are risk exposures in large dimensional factor models for financial returns constant over time? We answer this question by proposing a novel test for recurring regime changes in large dimensional factor models estimated by asymptotic principal components.[1] The test is robust to regime-specific second moment of the factors: it is thus suitable to detect recurring regime changes in risk exposures in portfolios with a large number of financial assets.

Large dimensional factor models describe the sources of common variations in vast datasets of financial variables: as stressed in Giglio and Xiu (2019), they allow to uncover the true drivers of asset returns. Seminal contributions such as Connor and Korajczyk (1986, 1988, 1993), Bai and Ng (2002), Stock and Watson (2002), and Bai (2003), assume that risk exposures are constant over time. However, as argued in Ang and Timmermann (2012), the maintained assumption of linearity is likely to be violated in financial markets due to structural breaks or recurring regime shifts: the former are induced by specific low frequency episodes such as regulatory changes and generate nonrecurring regimes; the latter relate to higher frequency recurring fluctuations.

A well established literature addresses structural instability in large dimensional factor models: Bates *et al.* (2013) study the robustness properties of the asymptotic principal components estimator in the presence of structural instability; Breitung and Eickmeier (2011), Chen *et al.* (2014), Corradi and Swanson (2014), Han and Inoue (2015), Yamamoto and Tanaka (2015), and Barigozzi and Trapani (2020), propose inferential procedures to detect breaks; Cheng *et al.* (2016), Baltagi *et al.* (2017, 2020), Su and Wang (2017), Barigozzi *et al.* (2018), Ma and Su (2018), and Massacci (2019), address the estimation problem.

Ang and Timmermann (2012) argue that recurring regime shifts arise in situations where "history repeats". An established literature considers small dimensional Markov switching factor models: see Kim (1994) for a seminal methodological contribution; Diebold and Rudebusch (1996), and Kim and Nelson (1998), for applications to business cycle analysis; and Baele *et al.* (2010), and more recently Guidolin and Pedio (2019) and references therein, for studies related to financial markets. As discussed in Chan *et al.* (2017), a strategy to account for recurring regimes involves threshold models. To the very best of our knowledge, Massacci (2017) is the first study allowing for regime changes in large dimensional factor models through the threshold

[1] We consider static factor models, where the dependent variables are static functions of the factors. For generalized dynamic factor models see Forni *et al.* (2017) and Barigozzi *et al.* (2019), and references therein.

principle. Further related contributions are Liu and Chen (2019), and Liu and Chen (2020). Regime changes in risk exposures are important for portfolios choice: Lehmann and Modest (2005) show that risk exposures can be mapped into weights of factor-mimicking portfolios; regime changes in risk exposures therefore lead to regime-specific investment opportunities. Despite the high number of contributions on large dimensional factors models with structural breaks, to the very best of our knowledge factor models with recurring regime changes are at a very early stage and vast scope for further research is available: we make a step in this direction.

We study testing for regime changes: this problem suffers from the course of dimensionality. Following Chen *et al.* (2014), and Han and Inoue (2015), Massacci (2017) tests for discrete shifts in the risk exposures by testing for threshold effect in the covariance matrix of the estimated factors. This strategy suffers from a main drawback: the covariance matrix of the true factors has to be independent of the regimes. This maintained assumption is unlikely to hold in financial markets: Fama and French (1993), and Lustig *et al.* (2011), show that the common factors driving the cross section of returns in equity and currency markets, respectively, are returns themselves; for example, it is violated if the factor volatility process follows the threshold GARCH model of Glosten *et al.* (1993), and recently studied in Cai and Stander (2020). Should the covariance matrix of the factors depend upon the regimes, the test would suffer from size distortions, which may have severe consequences for portfolio choice: since risk exposures can be mapped into weights of factor-mimicking portfolios, erroneous detection of regime changes in risk exposures may lead to sub-optimal asset allocation decisions.

We consider a portfolio of assets and assume that returns follow a factor structure: under the null hypothesis, returns allow for a linear factor representation; risk exposures experience threshold-type recurring regime changes under the alternative. We propose a Lagrange multiplier test for linearity in the risk exposures that is robust to regime-specific second moment of the factors. Our procedure follows three steps. First, we estimate number of factors and factors under the null hypothesis. Second, in the unconstrained model we replace the unknown factors with their estimates and take a weighted cross-sectional average of the returns: this results in an auxiliary threshold regression where the slope coefficients are weighted averages of the corresponding risk exposures. Finally, we test for a threshold effect in the auxiliary regression. Taking a weighted cross-sectional average of the returns transforms the infinite-dimensional problem of comparing two matrices of risk exposures into a finite-dimensional one. We show that our procedure results in a valid test in terms of size and power: in particular, the test does not suffer from size distortions when the second moment of the factors is regime-specific.

Finally, we apply our test to large portfolios of financial assets and show that risk exposures depend on recurring regimes. We map the regime-specific exposures to the weights of factor-mimicking portfolios and derive implications for risk management. We thus contribute to the literature on portfolio allocation under regime changes, which so far has involved a limited number of assets: on this respect, see Ang and Bekaert (2002), and Guidolin and Timmermann (2008). More generally, our paper relates to the growing literature on high dimensional factor analysis as applied to financial markets: see Giglio and Xiu (2019), and references therein.

The rest of the paper is organized as follows. Section 2 defines the set-up. Section 3 details the testing procedure. Section 4 links the test to portfolio choice and diversification. Section 5 performs a set of Monte Carlo experiments. Section 6 provides the empirical analysis. Section 7 concludes. The Appendix collects the technical proofs.

Notation. $\mathbb{I}(\cdot)$ is the indicator function. Given a square matrix $\mathbf{A}$, $\mathrm{tr}(\mathbf{A})$ denotes the trace of $\mathbf{A}$. The norm of a matrix $\mathbf{A}$ is $\|\mathbf{A}\| = [\mathrm{tr}(\mathbf{A}'\mathbf{A})]^{1/2}$. Given a scalar A, $|A|$, $\boldsymbol{\iota}_A$, $\mathbf{I}_A$ and $\mathbf{0}_A$ are the absolute value of A, the $A \times 1$ vector of ones, the $A \times A$ identity matrix and the $A \times A$ zero matrix, respectively. $\xrightarrow{p}$ and $\xrightarrow{d}$ denote convergence in probability and in distribution, respectively. $\Rightarrow$ denotes weak convergence with respect to the uniform metric. $\otimes$ denotes the Kronecker product. $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution and $\Phi^{-1}(\cdot)$ is the associated quantile function. $\lfloor \cdot \rfloor$ is the integer part of the argument.

2 Set-up

We describe the multi-factor model in Section 2.1, make assumptions on the model in Section 2.2, and discuss in details time-variation in the second moment of the factors in Section 2.3.

2.1 Multi-factor model with regime changes

2.1.1 Model

We consider the model

$$\mathbf{R}_t = \mathbb{I}(z_t \le \theta)\, \mathbf{B}_1 \mathbf{f}_t + \mathbb{I}(z_t > \theta)\, \mathbf{B}_2 \mathbf{f}_t + \mathbf{e}_t, \quad t = 1, \ldots, T, \tag{1}$$

where T denotes the time series dimension; $\mathbf{R}_t = (R_{1t}, \ldots, R_{Nt})'$ is an $N \times 1$ vector of asset (excess) returns; $\mathbf{f}_t = (f_{1t}, \ldots, f_{Pt})'$ is the $P \times 1$ vector of latent factors; z_t is an observable variable; θ is the potentially unknown threshold value; $\mathbf{e}_t = (e_{1t}, \ldots, e_{Nt})'$ is an $N \times 1$ vector of idiosyncratic errors; $\mathbf{B}_j = (\boldsymbol{\beta}_{j1}, \ldots, \boldsymbol{\beta}_{jN})'$ is the $N \times P$ matrix of risk exposures, with $i-th$ row defined as $\boldsymbol{\beta}'_{ji} = (\beta_{ji1}, \ldots, \beta_{jiP})$, for $j = 1, 2$, and $i = 1, \ldots, N$.

The specification in (1) belongs to the general class of large dimensional threshold factor models: see Massacci (2017), Liu and Chen (2019), and Liu and Chen (2020). According to the threshold principle of Pearson (1900), the regime prevailing at time t depends on the position of z_t with respect to θ. For $j = 1, 2$, the risk exposures $\mathbf{B}_j$ and the factors $\mathbf{f}_t$ are identified up to a rotation as they both are latent. The model allows for changes in the risk exposures: our test has power against the alternative hypothesis that the number of factors or the risk exposures (or both) are regime-specific. In the empirical analysis in Section 6 we investigate whether the number of common factors driving the comovement of equity returns changes between the regimes identified by the model: the common factors are the sources of systematic risk; a change in the number of factors implies that the sources of systematic risk depend upon the regime. Finally, the model in (1) allows for two regimes: this is consistent with Ang and Timmermann (2012), who focus on two-regime low dimensional models for financial markets. Our test has power against the alternative hypothesis of multiple regimes: however, to the very best of our knowledge, estimation and inference on multiple regimes in large dimensional factor models has not been studied and represents an interesting line of future research.

The observable threshold variable z_t in (1) drives the recurring regimes. Notice that if $z_t = t/T$ and $\theta \in (0, 1)$, then (1) becomes a multi-factor model with a single structural break: our procedure can be applied also to detect structural instability, although we plan to investigate this possibility in future work.

2.1.2 Interpreting regime changes

Ang and Timmermann (2012) study recurring regime changes in financial markets in low dimensional settings. We extend their arguments to a high dimensional framework: in particular, we consider the effect of regime changes in risk exposures on comovement among asset returns.

Given (1), let $\mathbf{c}_t = \mathbb{I}(z_t \leq \theta) \mathbf{B}_1 \mathbf{f}_t + \mathbb{I}(z_t > \theta) \mathbf{B}_2 \mathbf{f}_t$ be the $P \times 1$ vector of common components driving the comovement among the $N \times 1$ elements of $\mathbf{R}_t$. For ease of exposition, let $\{\mathbf{f}_t\}_{t=1}^{T}$ be an independently and identically distributed sequence: it follows that $\mathrm{E}(\mathbf{c}_t \mathbf{c}'_t | z_t \leq \theta) =$

$\mathbf{B}_1\mathrm{E}(\mathbf{f}_t\mathbf{f}_t')\mathbf{B}_1'$ and $\mathrm{E}(\mathbf{c}_t\mathbf{c}_t'|z_t > \theta) = \mathbf{B}_2\mathrm{E}(\mathbf{f}_t\mathbf{f}_t')\mathbf{B}_2'$. Therefore, regime changes in risk exposures lead to a shift in the covariance structure of the common components. The model in (1) matches the asymmetric cross-sectional dependence in asset returns: this extends Ang and Timmermann (2012), who focus on the low dimensional case through the analysis of pairwise correlation.

2.2 Assumptions

We now introduce the assumptions on the model in (1): these are valid both under the null and under the alternative hypothesis, which are stated and discussed in details in Section 3.1. Let $\mathbb{I}_{1t}(\theta) = \mathbb{I}(z_t \leq \theta)$ and $\mathbb{I}_{2t}(\theta) = \mathbb{I}(z_t > \theta)$. For $j = 1, 2$, denote P^0, $\mathbf{B}_j^0 = (\boldsymbol{\beta}_{j1}^0, \ldots, \boldsymbol{\beta}_{jN}^0)'$, θ^0 and $\mathbf{f}_t^0$ the true values of P, $\mathbf{B}_j$, θ and $\mathbf{f}_t$, respectively. Define $\mathbf{f}_{jt}^0(\theta^0) = \mathbb{I}_{jt}(\theta^0)\mathbf{f}_t^0$, for $j = 1, 2$.

Assumption A1 - Factors. $\mathrm{E}\|\mathbf{f}_t^0\|^4 < \infty$; for $j = 1, 2$, $T^{-1}\sum_{t=1}^T \mathbf{f}_{jt}^0(\theta^0)\mathbf{f}_{jt}^0(\theta^0)' \xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{f}jj}^0(\theta^0, \theta^0)$ as $T \to \infty$, for some $P^0 \times P^0$ positive definite matrix $\boldsymbol{\Sigma}_{\mathbf{f}jj}^0(\theta^0, \theta^0)$.

Assumption A2 - Risk Exposures. For $j = 1, 2$, and $i = 1, \ldots, N$, $\|\boldsymbol{\beta}_{ji}^0\| \leq \bar{\beta} < \infty$, and $\left\|\mathbf{B}_j^{0\prime}\mathbf{B}_j^0/N - \mathbf{D}_{\mathbf{B}jj}^0\right\| \to 0$ as $N \to \infty$, for some $P^0 \times P^0$ positive definite matrix $\mathbf{D}_{\mathbf{B}jj}^0$.

Assumption A3 - Time and Cross Section Dependence and Heteroskedasticity. There exists a positive $M < \infty$ such that for $j = 1, 2$, and for all (N, T),

- **(a)** $\mathrm{E}(e_{it}) = 0$ and $\mathrm{E}|e_{it}|^8 \leq M$ for all i;
- **(b)** $\mathrm{E}[\mathbb{I}_{jt}(\theta^0)\mathbb{I}_{jv}(\theta^0)e_{it}e_{iv}] = \tau_{jitv}(\theta^0)$ with $|\tau_{jitv}(\theta^0)| \leq |\tau_{jtv}|$ for some τ_{jtv} and for all i, and $T^{-1}\sum_{t=1}^T\sum_{v=1}^T|\tau_{jtv}| \leq M$;
- **(c)** $\mathrm{E}\left[T^{-1}\sum_{t=1}^T \mathbb{I}_{jt}(\theta^0)e_{it}e_{lt}\right] = \sigma_{jil}(\theta^0)$, $|\sigma_{jll}(\theta^0)| \leq M$ for all l, and $N^{-1}\sum_{i=1}^N\sum_{l=1}^N|\sigma_{jil}(\theta^0)| \leq M$;
- **(d)** $\mathrm{E}\left|T^{-1/2}\sum_{t=1}^T \mathbb{I}_{jt}(\theta^0)e_{it}e_{lt} - \mathrm{E}[\mathbb{I}_{jt}(\theta^0)e_{it}e_{lt}]\right|^4 \leq M$ for every (i, l).

Assumption A4 - Weak Dependence between $\mathbf{f}_t^0$, z_t and e_{it}. There exists some positive constant $M < \infty$ such that for all (N, T),

$$\mathrm{E}\left\{N^{-1}\sum_{i=1}^N\left\|T^{-1/2}\left[\sum_{t=1}^T \mathbb{I}_{jt}(\theta^0)\mathbf{f}_t^0 e_{it}\right]\right\|^2\right\} \leq M, \quad j = 1, 2.$$

Assumptions A1 to A4 are the natural extensions of Assumptions A to D imposed on linear factor models in Bai and Ng (2002) and accommodate the threshold effect: they are analogous to Assumptions C1 to C4, respectively, in Massacci (2017), to which we refer to for further comments; notice that we set $\theta = \theta^0$ as we are not interested in estimating θ^0. According to Assumption A1, factors can have regime-specific second moment: Assumption A1 plays a central role in this paper and we discuss it in details in Section 2.3 below. Assumption A2 implies that risk exposures are nonstochastic and in each state factors have a nonnegligible effect on the variance of $\mathbf{R}_t$. Assumption A3 follows Chamberlain and Rothschild (1983) and allows for some degree of cross-sectional correlation in the idiosyncratic components within each regime: in particular, it is less restrictive than Assumption 2 in Breitung and Eickmeier (2011), which requires that e_{it} and e_{lt} are independent for all $i \neq l$ to build the suggested pooled Lagrange multiplier test. Assumption A4 puts an upper bound on the degree of dependence among latent factors, threshold variable and idiosyncratic components: if $\mathbf{f}_t^0$, z_t and e_{it} are independent of each other, Assumption A4 is implied by Assumptions A1 and A3(a).

2.3 Regime-dependent factor covariance matrix

Aligned to Assumption A in Bai and Ng (2002), Assumption A1 allows for regime-dependent second moment of the factors. This contrasts with the literature that tests for discrete shifts in high dimensional factor models by testing for changes in the covariance matrix of the estimated factors. Let $T^{-1} \sum_{t=1}^{T} \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \xrightarrow{p} \mathbf{\Sigma}_{\mathbf{f}}^0$, where $\mathbf{\Sigma}_{\mathbf{f}}^0$ is a $P^0 \times P^0$ positive definite matrix: the existence of $\mathbf{\Sigma}_{\mathbf{f}}^0$ is guaranteed by Assumption A1. Define $\pi^0 = \mathrm{E}\left[\mathbb{I}_{1t}\left(\theta^0\right)\right]$. Monitoring the covariance matrix of the estimated factors requires $T^{-1} \sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right) \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \xrightarrow{p} \pi^0 \mathbf{\Sigma}_{\mathbf{f}}^0$, as stated in Assumption LT2 in Massacci (2017). If $z_t = t/T$, it is sufficient to assume that $\mathrm{E}\left(\mathbf{f}_t^0 \mathbf{f}_t^{0\prime}\right) = \mathbf{\Sigma}_{\mathbf{f}}^0$ (i.e., the second moment of the factors is time-invariant), as imposed in Assumption 2 and Assumption 1 in Chen *et al.* (2014), and Han and Inoue (2015), respectively. Allowing for heteroskedastic factors is important when modelling comovement among financial asset returns: on this, see for example Section 1.2.5 in Baele *et al.* (2010). In particular, Forbes and Rigobon (2002) argue that regime-dependent second moments play a key role in testing for changes in comovement among financial returns.

3 Testing for regime changes

Section 3.1 carefully explains the null and the alternative hypotheses. Section 3.2 deals with principal components estimation. Section 3.3 introduces the testing strategy. Sections 3.4 and

3.5 derive the limiting distributions of the test statistics under the null hypotheses. Section 3.6 looks at the power properties. Section 3.7 outlines the procedure to obtain the critical values. Section 3.8 analyzes the robustness of the test statistics to misspecification in the number of factors. Section 3.9 compares the size properties of our statistics with those of existing tests when the factor covariance matrix is regime-dependent.

3.1 Null and alternative hypotheses

We now state the null and the alternative hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. Define

$$\boldsymbol{\delta}_i^0 = \boldsymbol{\beta}_{2i}^0 - \mathbf{L}'\boldsymbol{\beta}_{1i}^0, \quad i = 1, \ldots, N,$$

where $\mathbf{L}$ is a $P^0 \times P^0$ full rank matrix: for $i = 1, \ldots, N$, $\boldsymbol{\delta}_i^0$ measures the deviation between the risk exposures in $\mathbf{B}_2^0$ and those in the linear rotation $\mathbf{B}_1^0\mathbf{L}$ of $\mathbf{B}_1^0$.

$$\mathcal{H}_0 : \boldsymbol{\delta}_i^0 = \mathbf{0} \text{ for some } \mathbf{L}, \text{ for } i = \left\lfloor N^{0.5} \right\rfloor + 1, \ldots, N. \tag{2}$$

$$\mathcal{H}_1 : \text{For } 0.5 < \alpha^0 \leq 1,\ \boldsymbol{\delta}_i^0 \neq \mathbf{0} \text{ for any } \mathbf{L}, \text{ for } i = 1, \ldots, \left\lfloor N^{\alpha^0} \right\rfloor. \tag{3}$$

The ordering of the cross-sectional units in (2) and (3) is for expositional convenience only and it not required for the validity of our test. To interpret $\mathcal{H}_0$ and $\mathcal{H}_1$, the parameter α^0 and the rotation matrix $\mathbf{L}$ deserve further discussion. Let us start from α^0, which depends on the number of cross-sectional units subject to a regime change. Under $\mathcal{H}_0$, *no more* than a fraction $O\left(N^{0.5}\right)$ of the cross-sectional units undergoes a regime shift; under $\mathcal{H}_1$, *at least* a fraction $O\left(N^{\alpha^0}\right)$ of the N cross-sectional units experiences a threshold effect in the risk exposures, for $0.5 < \alpha^0 \leq 1$. The parameter α^0 determines the convergence rate of the principal components estimator as applied to a linear model. Bates *et al.* (2013) show that if at most a fraction $O\left(N^{0.5}\right)$ of the series undergo a structural break in the risk exposures then the principal components estimator as applied to the misspecified linear model achieves the same Bai and Ng (2002) convergence rate $C_{NT} = \min\left\{\sqrt{N}, \sqrt{T}\right\}$. Massacci (2017) proves that the convergence rate of the principal components estimator as applied to the misspecified linear model when the data are generated according to (1) is $C_{NT}^{\alpha^0} = \min\left\{\sqrt{N}, \sqrt{T}, N^{1-\alpha^0}\right\}$: $C_{NT}^{\alpha^0} = C_{NT}$ under $\mathcal{H}_0$; the convergence rate is slower than C_{NT} under $\mathcal{H}_1$ for $0.5 < \alpha^0 < 1$; the estimator is inconsistent for $\alpha^0 = 1$.

The rotation matrix $\mathbf{L}$ is due to the rotational indeterminacy issue typical of large dimensional factor models. Under $\mathcal{H}_0$, a model with a regime change in the risk exposures is *observationally equivalent* to a linear model with a regime change in the covariance matrix of the factors, the latter scenario being covered by Assumption A in Bai and Ng (2002). Under $\mathcal{H}_1$, a model with a regime change in the risk exposures no longer is observationally equivalent to a linear model. Notice that $\mathcal{H}_1$ includes a regime change in the risk exposures, in the number of factors, or both: our test thus has power against a more general scenario than the one we consider for ease of exposition in Assumption A2.

We now link $\mathcal{H}_0$ and $\mathcal{H}_1$ in (2) and (3), respectively, to the literature on testing for structural instability in large dimensional factor models. Breitung and Eickmeier (2011) test for the null hypothesis $\boldsymbol{\delta}_i^0 = \mathbf{0}$ for given i by directly testing for equality in the estimated risk exposures: their test cannot be applied to $\mathcal{H}_0$ in (2) as it would reject $\mathcal{H}_0$ with probability approaching one as $N \to \infty$ because it involves an increasing number of estimated risk exposures. Breitung and Eickmeier (2011) also propose a pooled LM test for $\mathcal{H}_0$ in (2): the test requires independence in the idiosyncratic errors e_{it} and e_{jt} for all $i \neq j$; we allow for the more empirically realistic approximate factor structure. Chen *et al.* (2014), and Han and Inoue (2015), indirectly test for $\mathcal{H}_0$ in (2) by testing for structural change in the second moment of the factors estimated from the equivalent linear representation: Chen *et al.* (2014) follow a regression-based approach and test for structural stability in a model in which one estimated factor is regressed upon the remaining estimated factors; Han and Inoue (2015) work directly with the covariance matrix of the estimated factors. Massacci (2017) adapts the test of Chen *et al.* (2014) to the threshold factor model. As pointed out in Section 2.3, looking at estimated factors rather than estimated risk exposures requires the covariance matrix of the true factors to be regime independent. Regime changes in the covariance matrix of the true factors are observationally equivalent to a rotation in the risk exposures. Recall $\boldsymbol{\Sigma}_{\mathbf{f}}^0$ and π^0 as defined in Section 2.3: formally, the test in Massacci (2017) tests for $\mathcal{H}_0$ in (2) with $\mathbf{L} = \mathbf{I}_{P^0}$ under $T^{-1}\sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right) \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \xrightarrow{p} \pi^0 \boldsymbol{\Sigma}_{\mathbf{f}}^0$, which is more stringent than Assumption A1.

The second moment of the common factors in asset returns is unlikely to be independent of the regime: the Fama and French (1993) three factor model for equity returns includes three factors, which are returns from investment strategies; Lustig *et al.* (2011) document a similar factor structure in exchange rates. In Section 3.9 below we show that tests based on the second moment of the estimated factors suffer from size distortion when the covariance matrix of the true factors depends on the regime. We propose a test aligned to the empirically more plausible

Assumption A1 in this paper and Assumption A in Bai and Ng (2002).

3.2 Principal components estimation

Based on the considerations about $\mathcal{H}_0$ and $\mathcal{H}_1$ made in Section 3.1, we estimate the misspecified linear model $\mathbf{R}_t = \mathbf{B}_1\mathbf{f}_t + \mathbf{e}_t$ when regime changes are neglected. Following Bai and Ng (2002), Stock and Watson (2002), and Bai (2003), we estimate factors and loadings by principal components: this is a least squares estimator; as discussed in Bai and Li (2012), it is a quasi-maximum likelihood estimator when both factors and loadings are treated as parameters. Define $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_T)$, and let $\mathbf{F}^0 = \left(\mathbf{f}_1^0, \ldots, \mathbf{f}_T^0\right)$ be the true value of $\mathbf{F}$. For a given number of factors $P \geq 1$, the objective function in terms of $\mathbf{B}_1^P = \left(\boldsymbol{\beta}_{11}^P, \ldots, \boldsymbol{\beta}_{1N}^P\right)'$ and $\mathbf{F}^P = \left(\mathbf{f}_1^P, \ldots, \mathbf{f}_T^P\right)$ is the sum of squared residuals (divided by NT)

$$S\left(\mathbf{B}_1^P, \mathbf{F}^P\right) = (NT)^{-1} \sum_{t=1}^{T} \left(\mathbf{R}_t - \mathbf{B}_1^P \mathbf{f}_t^P\right)' \left(\mathbf{R}_t - \mathbf{B}_1^P \mathbf{f}_t^P\right), \tag{4}$$

where the superscript P on $\mathbf{B}_1^P$ and $\mathbf{F}^P$ denotes dependence on the number of factors. The estimators for $\mathbf{B}_1^0$ and $\mathbf{F}^0$ are obtained by minimizing $S\left(\mathbf{B}_1^P, \mathbf{F}^P\right)$ with respect to $\mathbf{B}_1^P$ and $\mathbf{F}^P$: as discussed in Bai and Ng (2002), for identification purposes this requires concentrating out either $\mathbf{B}_1^P$ or $\mathbf{F}^P$. As in Massacci (2017), we follow the second route. The estimators $\tilde{\mathbf{B}}_1^P = \left(\tilde{\boldsymbol{\beta}}_{11}^P, \ldots, \tilde{\boldsymbol{\beta}}_{1N}^P\right)'$ and $\tilde{\mathbf{F}}^P = \left(\tilde{\mathbf{f}}_1^P, \ldots, \tilde{\mathbf{f}}_T^P\right)$ for $\mathbf{B}_1^0$ and $\mathbf{F}^0$, respectively, solve

$$\tilde{\mathbf{B}}_1^P, \tilde{\mathbf{F}}^P = \arg \min_{\mathbf{B}_1^P, \mathbf{F}^P} S\left(\mathbf{B}_1^P, \mathbf{F}^P\right),$$

subject to the normalization $\mathbf{B}_1^{P\prime}\mathbf{B}_1^P / N = \mathbf{I}_P$. The estimator $\tilde{\mathbf{B}}_1^P$ is equal to $\sqrt{N}$ times the $N \times P$ matrix of eigenvectors of the sample covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{R}} = (NT)^{-1} \sum_{t=1}^{T} \mathbf{R}_t \mathbf{R}_t'$ corresponding to its largest P eigenvalues. The normalization $\mathbf{B}_1^{P\prime}\mathbf{B}_1^P / N = \mathbf{I}_P$ implies $\tilde{\mathbf{f}}_t^P = \left(\tilde{\mathbf{B}}_1^{P\prime}\tilde{\mathbf{B}}_1^P\right)^{-1} \tilde{\mathbf{B}}_1^{P\prime}\mathbf{R}_t = N^{-1}\tilde{\mathbf{B}}_1^{P\prime}\mathbf{R}_t$, for $t = 1, \ldots, T$.

The estimators $\tilde{\mathbf{B}}_1^P$ and $\tilde{\mathbf{F}}^P$ depend on the number of factors $P \geq 1$. The null and the alternative hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, affect existing model selection criteria. Given the loss function $S\left(\mathbf{B}_1^P, \mathbf{F}^P\right)$, let $\tilde{P}$ be the number of factors estimated by any of the information criteria of Bai and Ng (2002). Recall $\boldsymbol{\delta}_i^0 = \boldsymbol{\beta}_{2i}^0 - \mathbf{L}'\boldsymbol{\beta}_{1i}^0 = \left(\delta_{i1}^0, \ldots, \delta_{iP^0}^0\right)'$, for $i = 1, \ldots, N$. Let $P_{\boldsymbol{\delta}}^0$ be the number of factors $p = 1, \ldots, P^0$ such that $\sum_{i=1}^{N} \left|\delta_{ip}^0\right| = O\left(N^{\alpha^0}\right)$, for $0.5 < \alpha^0 \leq 1$: $P_{\boldsymbol{\delta}}^0$ is the number of factors whose loadings experience an identifiable regime change.

Proposition 3.1 *Let Assumptions A1 - A4 hold. Then:* (*i*) $\lim_{N,T\to\infty} \Pr\left(\tilde{P} = P^0\right) = 1$ *under* $\mathcal{H}_0$*; and* (*ii*) $\lim_{N,T\to\infty} \Pr\left(\tilde{P} = P^0 + P^0_{\boldsymbol{\delta}}\right) = 1$ *under* $\mathcal{H}_1$*.*

Under $\mathcal{H}_0$, the true number of factors P^0 is consistently estimated as $N, T \to \infty$; under $\mathcal{H}_1$, Bai and Ng (2002) information criteria overestimate the number of factors by $P^0_{\boldsymbol{\delta}}$. We conjecture that analogous results hold for other existing criteria applicable to the static factor model: see Trapani (2018) and references therein. Proposition 3.1 implies that the estimated number of factors in the linear model $\tilde{P}$ depends only on the dynamics of the risk exposures and not on the structure of the factor covariance matrix: this provides further justification for developing a test for regime changes that is robust to factor heteroskedasticity.

3.3 Testing strategy

From Section 3.2, $\tilde{P} \geq 1$ is the estimated number of factors in the linear model $\mathbf{R}_t = \mathbf{B}_1\mathbf{f}_t + \mathbf{e}_t$: $\tilde{P}$ is the true number of factors under $\mathcal{H}_0$, namely $\tilde{P} = P^0$; $\tilde{P} = \left(P^0 + P^0_{\boldsymbol{\delta}}\right)$ due to neglected regime shifts under $\mathcal{H}_1$. If $\tilde{P} = 1$, a regime shift is ruled out. If $\tilde{P} > 1$, we proceed as follows.

3.3.1 Weighting scheme

Assumption B1 - Weights. There exists a sequence of weights $\{w_i\}_{i=1}^N$ such that $w_i = O\left(N^{-1}\right)$ for $i = 1, \ldots, N$, and $\sum_{i=1}^N w_i = 1$.

Assumption B2 - Power. $\sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i\boldsymbol{\delta}_i^0 = \mathbf{0}$ if and only if $\sum_{i=\lfloor N^{0.5}\rfloor+1}^N \left\|\boldsymbol{\delta}_i^0\right\| = 0$.

Assumption B1 draws from Assumption 5 in Pesaran (2006). For a sequence of weights $\{w_i\}_{i=1}^N$ consistent with Assumption B1, Assumption B2 ensures that $\sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i\boldsymbol{\delta}_i^0 = \mathbf{0}$ if and only if under $\mathcal{H}_0$. This rules out the scenario where the weighted cross-sectional average of the sequence $\left\{\boldsymbol{\delta}_i^0\right\}_{i=\lfloor N^{0.5}\rfloor+1}^N$ (i.e., $\sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i\boldsymbol{\delta}_i^0$) vanishes even if an identifiable regime shift occurs, namely under $\mathcal{H}_1$: the proposed test would not have power in this case. For example, this happens if the sequence $\left\{\delta_{ip}^0\right\}_{i=\lfloor N^{0.5}\rfloor+1}^N$ is symmetric around zero and $w_i = 1/N$, for $p = 1, \ldots, P^0$, and $i = 1, \ldots, N$: in this case, $\sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i\delta_{ip}^0 = 0$, for $p = 1, \ldots, P^0$. Violation of Assumption B2 is extremely unlikely: the condition $\sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i\boldsymbol{\delta}_i^0 = \mathbf{0}$ is satisfied if and only if $\sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i\delta_{ip}^0 = 0$, for $p = 1, \ldots, P^0$; the weighted averages for *all* factors thus have to vanish under $\mathcal{H}_1$ for Assumption B2 to be violated. In the empirically relevant scenario in which the sequence of weights is chosen from a continuum of values, the probability of violating Assumption B2 is equal to zero. In practice, Assumption B2 thus is a regularity condition that is trivially satisfied. The weights tilt the distribution of $\left\{\boldsymbol{\delta}_i^0\right\}_{i=1}^N$ and are not unique. The weights are chosen *a priori*: an obvious option is $w_i = 1/N$, for $i = 1, \ldots, N$ (i.e., the equal-weight scheme); if applicable (e.g., in the case of individual stocks), the weight assigned to an asset may depend on its market capitalization. As shown in Theorems 3.1 and 3.2 below, the asymptotic distribution of the test statistics under the null hypothesis is independent of the weights. According to Theorem 3.3, the test is consistent for any weighting scheme that satisfies Assumption B2.

3.3.2 Test statistics

Based on the weights in Assumption B1, define the cross-sectional weighted averages

$$\bar{R}_{\mathbf{w}t} = \sum_{i=1}^N w_i R_{it}, \qquad \bar{\boldsymbol{\beta}}_{\mathbf{w}j} = \sum_{i=1}^N w_i \boldsymbol{\beta}_{ji}, \qquad \bar{e}_{\mathbf{w}t} = \sum_{i=1}^N w_i e_{it}, \qquad j = 1, 2.$$

Taking the cross-sectional weighted average of left and right-hand side of (1) leads to

$$\sum_{i=1}^N w_i R_{it} = \mathbb{I}_{1t}(\theta) \left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}\right)' \mathbf{f}_t + \mathbb{I}_{2t}(\theta) \left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{2i}\right)' \mathbf{f}_t + \sum_{i=1}^N w_i e_{it},$$

or equivalently

$$\bar{R}_{\mathbf{w}t} = \mathbb{I}_{1t}(\theta) \bar{\boldsymbol{\beta}}_{\mathbf{w}1}' \mathbf{f}_t + \mathbb{I}_{2t}(\theta) \bar{\boldsymbol{\beta}}_{\mathbf{w}2}' \mathbf{f}_t + \bar{e}_{\mathbf{w}t} : \tag{5}$$

under Assumption B1, the null hypothesis $\mathcal{H}_0$ implies

$$\sum_{i=\lfloor N^{0.5} \rfloor+1}^N w_i \boldsymbol{\beta}_{2i}^0 = \mathbf{L}' \sum_{i=\lfloor N^{0.5} \rfloor+1}^N w_i \boldsymbol{\beta}_{1i}^0.$$

Let $\tilde{\mathbf{B}}_1 = \tilde{\mathbf{B}}_1^{\tilde{P}} = \left(\tilde{\boldsymbol{\beta}}_{11}, \ldots, \tilde{\boldsymbol{\beta}}_{1N}\right)'$ and $\tilde{\mathbf{f}}_t = \tilde{\mathbf{f}}_t^{\tilde{P}}$ be the $N \times \tilde{P}$ matrix of estimated risk exposures and the $\tilde{P} \times 1$ vector of estimated factors, respectively, from $\mathbf{R}_t = \mathbf{B}_1 \mathbf{f}_t + \mathbf{e}_t$, for $t = 1, \ldots, T$: $\tilde{\mathbf{B}}_1^P$, $\tilde{\mathbf{f}}_t^P$ and $\tilde{P}$ are obtained as in Section 3.2. As in Hansen (1996), we build a Lagrange multiplier statistic. Define $\tilde{\mathbf{f}}_t(\theta) = \left[\mathbb{I}_{1t}(\theta) \tilde{\mathbf{f}}_t', \mathbb{I}_{2t}(\theta) \tilde{\mathbf{f}}_t'\right]'$. From the auxiliary model in (5) and for given θ, consider the least squares estimator $\widehat{\bar{\boldsymbol{\beta}}}_{\mathbf{w}}(\theta)$ for $\bar{\boldsymbol{\beta}}_{\mathbf{w}}^0 = \left(\bar{\boldsymbol{\beta}}_{\mathbf{w}1}^{0\prime}, \bar{\boldsymbol{\beta}}_{\mathbf{w}2}^{0\prime}\right)' = \left[\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)', \left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{2i}^0\right)'\right]'$ defined as

$$\widehat{\bar{\boldsymbol{\beta}}}_{\mathbf{w}}(\theta) = \left[\widehat{\bar{\boldsymbol{\beta}}}_{\mathbf{w}1}(\theta)', \widehat{\bar{\boldsymbol{\beta}}}_{\mathbf{w}2}(\theta)'\right]' = \left[\sum_{t=1}^T \tilde{\mathbf{f}}_t(\theta) \tilde{\mathbf{f}}_t(\theta)'\right]^{-1} \left[\sum_{t=1}^T \tilde{\mathbf{f}}_t(\theta) \bar{R}_{\mathbf{w}t}\right].$$

For any (θ_1, θ_2), define the matrix $\hat{\mathbf{M}}(\theta_1, \theta_2) = T^{-1} \sum_{t=1}^{T} \tilde{\mathbf{f}}_t(\theta_1) \tilde{\mathbf{f}}_t(\theta_2)'$. The regression scores $\mathbf{k}_t(\theta) = \tilde{\mathbf{f}}_t(\theta) \bar{e}_{\mathbf{w}t}$ are estimated under the null hypothesis as $\tilde{\mathbf{k}}_t(\theta) = \tilde{\mathbf{f}}_t(\theta) \tilde{\tilde{e}}_{\mathbf{w}t}$, where $\tilde{\tilde{e}}_{\mathbf{w}t} = \bar{R}_{\mathbf{w}t} - \tilde{\tilde{\boldsymbol{\beta}}}'_{\mathbf{w}1} \tilde{\mathbf{f}}_t$ and $\tilde{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}1} = \sum_{i=1}^{N} w_i \tilde{\boldsymbol{\beta}}_{1i}$. From Newey and West (1987), define: $\hat{\mathbf{K}}_d(\theta_1, \theta_2) = T^{-1} \sum_{t=d+1}^{T} \tilde{\mathbf{k}}_t(\theta_1) \tilde{\mathbf{k}}_{t-d}(\theta_2)$, for $d = 0, \ldots, D_T$, with $D_T = o\left(T^{1/4}\right)$; $\hat{\boldsymbol{\Omega}}(\theta_1, \theta_2) = \hat{\mathbf{K}}_0(\theta_1, \theta_2) + \sum_{d=1}^{D_T} w(d, D_T) \left[\hat{\mathbf{K}}_d(\theta_1, \theta_2) + \hat{\mathbf{K}}_d(\theta_1, \theta_2)'\right]$, where $w(d, D_T) = [1 - d/(D_T + 1)]$ is the Bartlett kernel. Define the $\tilde{P} \times 2\tilde{P}$ matrix $\mathbf{G} = \left(\mathbf{I}_{\tilde{P}}, -\mathbf{I}_{\tilde{P}}\right)'$. For given θ, the heteroskedasticity and autocorrelation robust Lagrange multiplier test statistic is

$$\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta) = T \hat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}(\theta)' \mathbf{G} \left[\mathbf{G}' \hat{\mathbf{M}}(\theta, \theta)^{-1} \hat{\boldsymbol{\Omega}}(\theta, \theta) \hat{\mathbf{M}}(\theta, \theta)^{-1} \mathbf{G}\right]^{-1} \mathbf{G}' \hat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}(\theta). \tag{6}$$

For known θ^0 (i.e., for $\theta = \theta^0$) and under the null hypothesis, $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}\left(\theta^0\right)$ has a χ^2 limiting distribution with P^0 degrees of freedom as $N, T \to \infty$ (see Theorem 3.1 in Section 3.4 below). However, θ^0 is generally unknown and not identified under the null hypothesis. Following Hansen (1996), we propose the statistic

$$\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}} = \sup_{\theta} \widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta). \tag{7}$$

When the idiosyncratic errors are time homoskedastic and do not exhibit serial correlation, the $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ statistic in (6) simplifies to

$$\widehat{LM}_{\mathbf{w}}(\theta) = T \hat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}(\theta)' \mathbf{G} \left[\left(T^{-1} \sum_{t=1}^{T} \tilde{\tilde{e}}'_{\mathbf{w}t} \tilde{\tilde{e}}_{\mathbf{w}t}\right) \mathbf{G}' \hat{\mathbf{M}}(\theta, \theta)^{-1} \mathbf{G}\right]^{-1} \mathbf{G}' \hat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}(\theta): \tag{8}$$

when θ^0 is unknown, the relevant statistic thus is

$$\sup \widehat{LM}_{\mathbf{w}} = \sup_{\theta} \widehat{LM}_{\mathbf{w}}(\theta).$$

3.4 Limiting distribution of $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}\left(\theta^0\right)$ under the null hypothesis

Recall $\mathbf{D}_{\mathbf{B}11}^0$ in Assumption A2 and that $T^{-1} \sum_{t=1}^{T} \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{f}}^0$ from Section 2.3. Notice that

$$T^{-1/2} \sum_{t=1}^{T} \mathbb{I}_{jt}(\theta) \mathbf{f}_t^0 \bar{e}_{\mathbf{w}t} = \sum_{i=1}^{N} w_i \left[T^{-1/2} \sum_{t=1}^{T} \mathbb{I}_{jt}(\theta) \mathbf{f}_t^0 e_{it}\right], \quad j = 1, 2,$$

and define

$$\begin{aligned}\boldsymbol{\Omega}^0_{jk}\left(\theta_1,\theta_2\right) &= \lim_{N,T\rightarrow\infty} T^{-1}\sum_{t=1}^{T}\sum_{v=1}^{T}\mathrm{E}\left[\mathbb{I}_{jt}\left(\theta_1\right)\mathbb{I}_{kv}\left(\theta_2\right)\mathbf{f}^0_t\mathbf{f}^{0\prime}_v\bar{e}_{\mathbf{w}t}\bar{e}_{\mathbf{w}v}\right] \\ &= \lim_{N,T\rightarrow\infty}\sum_{i=1}^{N}\sum_{l=1}^{N}w_iw_l\left\{T^{-1}\sum_{t=1}^{T}\sum_{v=1}^{T}\mathrm{E}\left[\mathbb{I}_{jt}\left(\theta_1\right)\mathbb{I}_{kv}\left(\theta_2\right)\mathbf{f}^0_t\mathbf{f}^{0\prime}_v e_{it}e_{lv}\right]\right\},\end{aligned} \qquad j,k=1,2.$$

Assumption C1 - Eigenvalues under the Null Hypothesis. The eigenvalues of the $P^0\times P^0$ matrix $\left(\mathbf{D}^0_{\mathbf{B}11}\cdot\boldsymbol{\Sigma}^0_{\mathbf{f}}\right)$ are distinct.

Assumption C2 - Convergence Rates. $\sqrt{T}/N\rightarrow 0$ as $N,T\rightarrow\infty$.

Assumption C3 - Central Limit Theorem. For $j=1,2$, $T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{jt}\left(\theta^0\right)\mathbf{f}^0_t\bar{e}_{\mathbf{w}t}\xrightarrow{d}\mathcal{N}\left[\mathbf{0},\boldsymbol{\Omega}^0_{jj}\left(\theta^0,\theta^0\right)\right]$ as $N,T\rightarrow\infty$, where $\boldsymbol{\Omega}^0_{jj}\left(\theta^0,\theta^0\right)$ is a positive definite matrix.

Assumption C1 is analogous to Assumption G in Bai (2003): it guarantees a unique probability limit for $\left(\mathbf{B}^{0\prime}_j\tilde{\mathbf{B}}_j/N\right)$, which enters the asymptotic distributions of the test statistics $\widehat{LM}^{\text{HAC}}_{\mathbf{w}}\left(\theta^0\right)$ and $\sup\widehat{LM}^{\text{HAC}}_{\mathbf{w}}$ defined according to (6) and (7), respectively. Assumption C2 restricts the convergence rates to ensure that the inclusion of estimated factors does not affect the asymptotic distribution of the test statistic. Given Assumption B1, Assumption C3 is a central limit theorem similar to Assumption F4 in Bai (2003) as applied to the weighted average of $T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{jt}\left(\theta^0\right)\mathbf{f}^0_t e_{it}$, with weights given by w_i, for $i=1,\ldots,N$: the covariance matrix $\boldsymbol{\Omega}^0_{jj}\left(\theta^0,\theta^0\right)$ is the weighted average of the asymptotic covariance matrix of $T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{jt}\left(\theta^0\right)\mathbf{f}^0_t e_{it}$, with weights given by w_iw_l, for $i,l=1,\ldots,N$, and the assumption that $\boldsymbol{\Omega}^0_{jj}\left(\theta^0,\theta^0\right)$ is positive definite is thus not restrictive.

Define

$$\boldsymbol{\Omega}^0\left(\theta_1,\theta_2\right)=\begin{bmatrix}\boldsymbol{\Omega}^0_{11}\left(\theta_1,\theta_2\right) & \boldsymbol{\Omega}^0_{12}\left(\theta_1,\theta_2\right)\\ \boldsymbol{\Omega}^0_{21}\left(\theta_1,\theta_2\right) & \boldsymbol{\Omega}^0_{22}\left(\theta_1,\theta_2\right)\end{bmatrix}:$$

by construction, $\boldsymbol{\Omega}^0_{12}\left(\theta_1,\theta_2\right)=\mathbf{0}_{P^0}$ and $\boldsymbol{\Omega}^0_{21}\left(\theta_1,\theta_2\right)=\mathbf{0}_{P^0}$ if $\theta_1<\theta_2$ and $\theta_1>\theta_2$, respectively, and $\boldsymbol{\Omega}^0_{12}\left(\theta_1,\theta_2\right)=\boldsymbol{\Omega}^0_{21}\left(\theta_1,\theta_2\right)=\mathbf{0}_{P^0}$ if $\theta_1=\theta_2$. Let $\tilde{\mathbf{V}}_1$ be the $\tilde{P}\times\tilde{P}$ diagonal matrix of the first $\tilde{P}$ largest eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\mathbf{R}}=(NT)^{-1}\sum_{t=1}^{T}\mathbf{R}_t\mathbf{R}_t'$ in decreasing order; define the $P^0\times\tilde{P}$ rotation matrix $\tilde{\mathbf{H}}_1$ as

$$\tilde{\mathbf{H}}_1=\frac{\mathbf{F}^0\mathbf{F}^{0\prime}}{T}\frac{\mathbf{B}^{0\prime}_1\tilde{\mathbf{B}}_1}{N}\tilde{\mathbf{V}}_1^{-1}.$$

Theorem 3.1 *Let Assumptions A1-A4, B1, C1-C3 hold. Then* $\widehat{LM}^{\text{HAC}}_{\mathbf{w}}\left(\theta^0\right)\xrightarrow{d}\chi^2\left(P^0\right)$ *under* $\mathcal{H}_0$*, provided that* $\hat{\boldsymbol{\Omega}}\left(\theta^0,\theta^0\right)\xrightarrow{p}\left(\mathbf{I}_2\otimes\mathbf{H}^0_1\right)^{-1}\boldsymbol{\Omega}^0\left(\theta^0,\theta^0\right)\left(\mathbf{I}_2\otimes\mathbf{H}^{0\prime}_1\right)^{-1}$*, where* $\mathbf{H}^0_1=p\lim_{N,T\rightarrow\infty}\tilde{\mathbf{H}}_1$.

The asymptotic distribution of $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}\left(\theta^0\right)$ stated in Theorem 3.1 is valid for any weighting scheme that satisfies Assumption B1. Theorem 3.1 holds under Assumption A1, which allows for regime-dependent second moment of the true factors $\mathbf{f}_t^0$. Chen *et al.* (2014), and Han and Inoue (2015), test for structural break in the risk exposures by comparing the second moment of the estimated factors before and after the break: this procedure requires time-invariant second moment of the true factors, as imposed in Assumption 2 in Chen *et al.* (2014), and in Assumption 1 in Han and Inoue (2015), and it implies $\mathbf{L} = \mathbf{I}_{P^0}$ in $\mathcal{H}_0$ in (2); should the second moment of $\mathbf{f}_t^0$ be time-varying, the test would erroneously reject the null hypothesis $\mathcal{H}_0$ when in fact $\mathcal{H}_0$ is true, as discussed in Section 4.4 in Chen *et al.* (2014), and in note 4 in Han and Inoue (2015). Massacci (2017) follows up on Chen *et al.* (2014), and Han and Inoue (2015), and tests for a regime change in the risk exposures according to the model in (1) by comparing the second moment of the estimated factors in the two regimes: the test requires regime-invariant second moment of the true factors, as imposed in Assumption LT2. In line with the analytical results in Section 3.9, the results from the Monte Carlo experiment in Section 5.1 show that tests for linearity based on the second moment of the estimated factors are likely to suffer from severe size distortions if the second moment of the true factors is regime-dependent.

3.5 Limiting distribution of $\sup \widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}$ under the null hypothesis

Define $\hat{\mathbf{k}}(\theta) = T^{-1/2}\sum_{t=1}^{T}\tilde{\mathbf{k}}_t(\theta) = T^{-1/2}\sum_{t=1}^{T}\tilde{\mathbf{f}}_t(\theta)\widetilde{\widetilde{e}}_{\mathbf{w}t}$. Let $\mathbf{k}^0(\theta)$ be a zero mean Gaussian process with covariance kernel

$$\boldsymbol{\Omega}^0_{\mathbf{H}_1^0}(\theta_1,\theta_2) = \mathrm{E}\left[\mathbf{k}^0(\theta_1)\,\mathbf{k}^0(\theta_2)'\right] = \left[\boldsymbol{\iota}_2\boldsymbol{\iota}_2'\otimes\left(\mathbf{H}_1^0\right)^{-1}\right]\boldsymbol{\Omega}^0(\theta_1,\theta_2)\left[\boldsymbol{\iota}_2\boldsymbol{\iota}_2'\otimes\left(\mathbf{H}_1^{0\prime}\right)^{-1}\right],$$

where $\mathbf{H}_1^0$ is as in Theorem 3.1. Further, define

$$\mathbf{M}^0(\theta_1,\theta_2) = \lim_{T\to\infty}\sum_{t=1}^{T}\mathrm{E}\left\{\left[\begin{array}{c}\mathbb{I}_{1t}(\theta_1)\,\mathbf{f}_t^0\\ \mathbb{I}_{2t}(\theta_1)\,\mathbf{f}_t^0\end{array}\right]\left[\begin{array}{c}\mathbb{I}_{1t}(\theta_2)\,\mathbf{f}_t^0\\ \mathbb{I}_{2t}(\theta_2)\,\mathbf{f}_t^0\end{array}\right]'\right\} = \left[\begin{array}{cc}\boldsymbol{\Sigma}^0_{\mathbf{f}11}(\theta_1,\theta_2) & \boldsymbol{\Sigma}^0_{\mathbf{f}12}(\theta_1,\theta_2)\\ \boldsymbol{\Sigma}^0_{\mathbf{f}21}(\theta_1,\theta_2) & \boldsymbol{\Sigma}^0_{\mathbf{f}22}(\theta_1,\theta_2)\end{array}\right]$$

and

$$\mathbf{M}^0_{\mathbf{H}_1^0}(\theta_1,\theta_2) = \left[\boldsymbol{\iota}_2\boldsymbol{\iota}_2'\otimes\left(\mathbf{H}_1^0\right)^{-1}\right]\mathbf{M}^0(\theta_1,\theta_2)\left[\boldsymbol{\iota}_2\boldsymbol{\iota}_2'\otimes\left(\mathbf{H}_1^{0\prime}\right)^{-1}\right].$$

Assumption D1 - Mixing Condition and Moment Bound.

(a) $\left\{\bar{R}_{\mathbf{w}t},\mathbf{f}_t^0,z_t\right\}_{t=1}^{T}$ is strictly stationary and $\beta-$mixing, with $\beta-$mixing coefficients satisfying $\beta_m = O\left(m^{-\nu}\right)$ for some $\nu > \xi/(\xi-1)$ and $k \geq \xi > 1$;

(b) $\mathrm{E}\left\{\left|\max_{j=1,2}\left[\sup_\theta \left\|\mathbb{I}_{jt}(\theta)\,\mathbf{f}_t^0\right\|\right]\right|^{4k}\right\} < \infty$;

(c) $\mathrm{E}\left|e_{it}\right|^{4k} < \infty$.

Assumption D2 - Bracketing. For all θ, and for some $M < \infty$ and $\gamma > 0$, there exists some $\bar{\theta}$ such that $\left\{\mathrm{E}\left|\max_{j=1,2}\left\|\left[\mathbb{I}_{jt}(\theta) - \mathbb{I}_{jt}(\bar{\theta})\right]\mathbf{f}_t^0 e_{it}\right\|\right|^{2\xi}\right\}^{1/(2\xi)} \leq M\left|\theta - \bar{\theta}\right|^{\gamma}$.

Assumption D3 - Uniform Convergence under the Null Hypothesis. $\hat{\mathbf{M}}(\theta_1, \theta_2)$ and $\hat{\mathbf{\Omega}}(\theta_1, \theta_2)$ converge in probability to $\mathbf{M}^0_{\mathbf{H}_1^0}(\theta_1, \theta_2)$ and $\mathbf{\Omega}^0_{\mathbf{H}_1^0}(\theta_1, \theta_2)$, respectively, uniformly over (θ_1, θ_2), where $\mathbf{M}^0_{\mathbf{H}_1^0}(\theta_1, \theta_2)$ and $\mathbf{\Omega}^0_{\mathbf{H}_1^0}(\theta_1, \theta_2)$ are positive definite matrices.

Assumptions D1-D3 are analogous to Assumptions 1-3 in Hansen (1996): in particular, Assumption D1(a) applies to the auxiliary threshold regression in (5) and it imposes a stationarity condition, which allows for regime-specific covariance matrix of the factors. The uniform convergence of $\hat{\mathbf{\Omega}}(\theta_1, \theta_2)$ to $\mathbf{\Omega}^0_{\mathbf{H}_1^0}(\theta_1, \theta_2)$ is not stringent: factors are consistently estimated from a linear model under $\mathcal{H}_0$ in (2); $\hat{\mathbf{\Omega}}(\theta_1, \theta_2)$ is a HAC estimator for $\mathbf{\Omega}^0_{\mathbf{H}_1^0}(\theta_1, \theta_2)$. In particular, $\mathbf{\Omega}^0(\theta_1, \theta_2)$ is the weighted average of the covariance kernel of $T^{-1/2}\sum_{t=1}^T\left[\mathbb{I}_{1t}(\theta), \mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0 e_{it}$, with weights given by $w_i w_l$, for $i, l = 1, \ldots, N$: given the features of $\mathbf{\Omega}^0(\theta_1, \theta_2)$ discussed in Section 3.4, it is not stringent to assume that $\mathbf{\Omega}^0_{\mathbf{H}_1^0}(\theta_1, \theta_2)$ is positive definite. Assumptions D1 and D3 jointly imply Assumption A1.

Define

$$
\begin{aligned}
& LM^{\mathrm{HAC},0}_{\mathbf{w}}(\theta) \\
= & \left[\mathbf{M}^0_{\mathbf{H}_1^0}(\theta,\theta)^{-1}\mathbf{k}^0(\theta)\right]'\mathbf{G}\left[\mathbf{G}'\mathbf{M}^0_{\mathbf{H}_1^0}(\theta,\theta)^{-1}\mathbf{\Omega}^0_{\mathbf{H}_1^0}(\theta,\theta)\mathbf{M}^0_{\mathbf{H}_1^0}(\theta,\theta)^{-1}\mathbf{G}\right]^{-1}\mathbf{G}'\left[\mathbf{M}^0_{\mathbf{H}_1^0}(\theta,\theta)^{-1}\mathbf{k}^0(\theta)\right],
\end{aligned}
$$

where $\mathbf{G}'\left[\mathbf{M}^0_{\mathbf{H}_1^0}(\theta,\theta)^{-1}\mathbf{k}^0(\theta)\right]$ is a zero mean Gaussian process with covariance Kernel equal to $\mathbf{G}'\mathbf{M}^0_{\mathbf{H}_1^0}(\theta_1,\theta_1)^{-1}\mathbf{\Omega}^0_{\mathbf{H}_1^0}(\theta_1,\theta_2)\mathbf{M}^0_{\mathbf{H}_1^0}(\theta_2,\theta_2)^{-1}\mathbf{G}$.

Theorem 3.2 *Let Assumptions A2-A4, B1, C1, C2 and D1-D3 hold. Then* $\hat{\mathbf{k}}^{\mathbb{H}_0}(\theta) \Rightarrow \mathbf{k}^0(\theta)$, $\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}(\theta) \Rightarrow LM^{\mathrm{HAC},0}_{\mathbf{w}}(\theta)$ *and* $\sup \widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}} \xrightarrow{d} \sup LM^{\mathrm{HAC},0}_{\mathbf{w}}$ *under* $\mathcal{H}_0$.

The asymptotic distribution of $\sup \widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}$ stated in Theorem 3.2 holds for any weighting scheme described in Assumption B1. The distribution of $\sup LM^{\mathrm{HAC},0}_{\mathbf{w}}$ is generally unknown; however, the critical values can be obtained as detailed in Section 3.7.

3.6 Power properties

Given Proposition 3.1, under $\mathcal{H}_1$ in (3) the model in (1) can be equivalently written as[2]

$$\mathbf{R}_t = \mathbf{B}_{\boldsymbol{\delta}}^0 \mathbf{f}_{\boldsymbol{\delta} t}^0 \left(\theta^0\right) + \mathbf{e}_t, \quad t = 1, \ldots, T, \tag{9}$$

where $\mathbf{f}_{\boldsymbol{\delta} t}^0 \left(\theta^0\right) = \left[\mathbf{f}_t^{0\prime}, \mathbf{f}_{\boldsymbol{\delta} 2t}^0 \left(\theta^0\right)'\right]'$ and $\mathbf{B}_{\boldsymbol{\delta}}^0 = \left(\mathbf{B}_1^0, \boldsymbol{\Delta}_{\boldsymbol{\delta} 2}^0\right)$, $\mathbf{f}_{\boldsymbol{\delta} 2t}^0 \left(\theta^0\right) = \mathbb{I}_{2t} \left(\theta^0\right) \mathbf{f}_{\boldsymbol{\delta} t}^0$ and $\mathbf{f}_{\boldsymbol{\delta} t}^0$ denotes the factors in $\mathbf{f}_t^0$ with an identifiable regime change in the loadings, and $\boldsymbol{\Delta}_{\boldsymbol{\delta} 2}^0 = \left(\boldsymbol{\delta}_{\boldsymbol{\delta} 21}^0, \ldots, \boldsymbol{\delta}_{\boldsymbol{\delta} 2N}^0\right)'$ is the $N \times P_{\boldsymbol{\delta}}^0$ matrix of loadings of $\mathbf{f}_{\boldsymbol{\delta} t}^0$ when $\mathbb{I}_{2t} \left(\theta^0\right) = 1$. Assumptions A1 and A2 imply that $T^{-1} \sum_{t=1}^T \mathbf{f}_{\boldsymbol{\delta} t}^0 \left(\theta^0\right) \mathbf{f}_{\boldsymbol{\delta} t}^0 \left(\theta^0\right)' \xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{f}\boldsymbol{\delta}}^0 \left(\theta^0, \theta^0\right)$ as $T \to \infty$ and $\left\|\mathbf{B}_{\boldsymbol{\delta}}^{0\prime} \mathbf{B}_{\boldsymbol{\delta}}^0 / N - \mathbf{D}_{\mathbf{B}\boldsymbol{\delta}}^0\right\| \to 0$ as $N \to \infty$, respectively, where $\boldsymbol{\Sigma}_{\mathbf{f}\boldsymbol{\delta}}^0 \left(\theta^0, \theta^0\right)$ and $\mathbf{D}_{\mathbf{B}\boldsymbol{\delta}}^0$ are $\left(P^0 + P_{\boldsymbol{\delta}}^0\right) \times \left(P^0 + P_{\boldsymbol{\delta}}^0\right)$ positive definite matrices.

Let us define: $\mathbf{F}_{\boldsymbol{\delta}}^0 \left(\theta^0\right) = \left[\mathbf{f}_{\boldsymbol{\delta} 1}^0 \left(\theta^0\right), \ldots, \mathbf{f}_{\boldsymbol{\delta} T}^0 \left(\theta^0\right)\right]$; the principal components estimator $\tilde{\mathbf{B}}_{\boldsymbol{\delta}}$ for $\mathbf{B}_{\boldsymbol{\delta}}^0$ as $\sqrt{N}$ times the $N \times \tilde{P}$ matrix of eigenvectors of the sample covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{R}} = (NT)^{-1} \sum_{t=1}^T \mathbf{R}_t \mathbf{R}_t'$ corresponding to its largest $\tilde{P}$ eigenvalues; $\tilde{\mathbf{V}}_{\boldsymbol{\delta}}$ as the $\tilde{P} \times \tilde{P}$ diagonal matrix containing the first $\tilde{P}$ eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\mathbf{R}}$ in decreasing order. Define the $\left(P^0 + P_{\boldsymbol{\delta}}^0\right) \times \tilde{P}$ rotation matrix

$$\tilde{\mathbf{H}}_{\boldsymbol{\delta}} \left(\theta^0\right) = \frac{\mathbf{F}_{\boldsymbol{\delta}}^0 \left(\theta^0\right) \mathbf{F}_{\boldsymbol{\delta}}^0 \left(\theta^0\right)'}{T} \frac{\mathbf{B}_{\boldsymbol{\delta}}^{0\prime} \tilde{\mathbf{B}}_{\boldsymbol{\delta}}}{N} \tilde{\mathbf{V}}_{\boldsymbol{\delta}}^{-1},$$

and let $\mathbf{H}_{\boldsymbol{\delta}}^0 \left(\theta^0\right)$ be the $\left(P^0 + P_{\boldsymbol{\delta}}^0\right) \times \left(P^0 + P_{\boldsymbol{\delta}}^0\right)$ matrix such that $\mathbf{H}_{\boldsymbol{\delta}}^0 \left(\theta^0\right) = p \lim_{N,T \to \infty} \tilde{\mathbf{H}}_{\boldsymbol{\delta}} \left(\theta^0\right)$. Define

$$\boldsymbol{\Omega}_{\mathbf{H}_{\boldsymbol{\delta}}^0}^0 \left(\theta_1, \theta_2\right) = \left\{\boldsymbol{\iota}_2 \boldsymbol{\iota}_2' \otimes \left[\mathbf{H}_{\boldsymbol{\delta}}^0 \left(\theta^0\right)\right]^{-1}\right\} \boldsymbol{\Omega}_{\boldsymbol{\delta}}^0 \left(\theta_1, \theta_2\right) \left\{\boldsymbol{\iota}_2 \boldsymbol{\iota}_2' \otimes \left[\mathbf{H}_{\boldsymbol{\delta}}^0 \left(\theta^0\right)'\right]^{-1}\right\}$$

with

$$\boldsymbol{\Omega}_{\boldsymbol{\delta}}^0 \left(\theta_1, \theta_2\right) = \begin{bmatrix} \boldsymbol{\Omega}_{\boldsymbol{\delta} 11}^0 \left(\theta_1, \theta_2\right) & \boldsymbol{\Omega}_{\boldsymbol{\delta} 12}^0 \left(\theta_1, \theta_2\right) \\ \boldsymbol{\Omega}_{\boldsymbol{\delta} 21}^0 \left(\theta_1, \theta_2\right) & \boldsymbol{\Omega}_{\boldsymbol{\delta} 22}^0 \left(\theta_1, \theta_2\right) \end{bmatrix},$$

where

$$\begin{aligned} \boldsymbol{\Omega}_{\boldsymbol{\delta} jk}^0 \left(\theta_1, \theta_2\right) &= \lim_{N,T \to \infty} T^{-1} \sum_{t=1}^T \sum_{v=1}^T \mathrm{E}\left[\mathbb{I}_{jt} \left(\theta_1\right) \mathbb{I}_{kv} \left(\theta_2\right) \mathbf{f}_{\boldsymbol{\delta} t}^0 \left(\theta^0\right) \mathbf{f}_{\boldsymbol{\delta} v}^0 \left(\theta^0\right)' \bar{e}_{\mathbf{w}t} \bar{e}_{\mathbf{w}v}\right] \\ &= \lim_{N,T \to \infty} \sum_{i=1}^N \sum_{l=1}^N w_i w_l \left\{T^{-1} \sum_{t=1}^T \sum_{v=1}^T \mathrm{E}\left[\mathbb{I}_{jt} \left(\theta_1\right) \mathbb{I}_{kv} \left(\theta_2\right) \mathbf{f}_{\boldsymbol{\delta} t}^0 \left(\theta^0\right) \mathbf{f}_{\boldsymbol{\delta} v}^0 \left(\theta^0\right)' e_{it} e_{lv}\right]\right\}, \end{aligned} \quad j,k = 1,2:$$

[2] Formally

$$\begin{aligned} \mathbf{R}_t &= \mathbb{I}_{1t} \left(\theta^0\right) \mathbf{B}_1^0 \mathbf{f}_t^0 + \mathbb{I}_{2t} \left(\theta^0\right) \mathbf{B}_2^0 \mathbf{f}_t^0 + \mathbf{e}_t \\ &= \mathbf{B}_1^0 \mathbf{f}_t^0 + \left(\mathbf{B}_2^0 - \mathbf{B}_1^0\right) \left[\mathbb{I}_{2t} \left(\theta^0\right) \mathbf{f}_t^0\right] + \mathbf{e}_t \\ &= \mathbf{B}_1^0 \mathbf{f}_t^0 + \boldsymbol{\Delta}_{\boldsymbol{\delta} 2}^0 \left[\mathbb{I}_{2t} \left(\theta^0\right) \mathbf{f}_{\boldsymbol{\delta} t}^0\right] + \mathbf{e}_t \\ &= \mathbf{B}_1^0 \mathbf{f}_t^0 + \boldsymbol{\Delta}_{\boldsymbol{\delta} 2}^0 \mathbf{f}_{\boldsymbol{\delta} 2t}^0 \left(\theta^0\right) + \mathbf{e}_t \\ &= \mathbf{B}_{\boldsymbol{\delta}}^0 \mathbf{f}_{\boldsymbol{\delta} t}^0 \left(\theta^0\right) + \mathbf{e}_t. \end{aligned}$$

the matrix $\mathbf{\Omega}^0_{\boldsymbol{\delta}}(\theta_1,\theta_2)$ is such that $\mathbf{\Omega}^0_{\boldsymbol{\delta}12}(\theta_1,\theta_2)=\mathbf{0}_{P^0+P^0_{\boldsymbol{\delta}}}$ if $\theta_1<\theta_2$, $\mathbf{\Omega}^0_{\boldsymbol{\delta}21}(\theta_1,\theta_2)=\mathbf{0}_{P^0+P^0_{\boldsymbol{\delta}}}$ if $\theta_1>\theta_2$, and $\mathbf{\Omega}^0_{\boldsymbol{\delta}12}(\theta_1,\theta_2)=\mathbf{\Omega}^0_{\boldsymbol{\delta}21}(\theta_1,\theta_2)=\mathbf{0}_{P^0+P^0_{\boldsymbol{\delta}}}$ if $\theta_1=\theta_2$. Let

$$\mathbf{M}^0_{\boldsymbol{\delta}}(\theta_1,\theta_2)=\lim_{T\to\infty}\sum_{t=1}^{T}\mathrm{E}\left\{\left[\begin{array}{c}\mathbb{I}_{1t}(\theta_1)\,\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0)\\ \mathbb{I}_{2t}(\theta_1)\,\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0)\end{array}\right]\left[\begin{array}{c}\mathbb{I}_{1t}(\theta_2)\,\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0)\\ \mathbb{I}_{2t}(\theta_2)\,\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0)\end{array}\right]'\right\}$$

and

$$\mathbf{M}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}(\theta_1,\theta_2)=\left\{\boldsymbol{\iota}_2\boldsymbol{\iota}_2'\otimes\left[\mathbf{H}^0_{\boldsymbol{\delta}}(\theta^0)\right]^{-1}\right\}\mathbf{M}^0_{\boldsymbol{\delta}}(\theta_1,\theta_2)\left\{\boldsymbol{\iota}_2\boldsymbol{\iota}_2'\otimes\left[\mathbf{H}^0_{\boldsymbol{\delta}}(\theta^0)'\right]^{-1}\right\}.$$

Assumption E1 - Eigenvalues under the Alternative Hypothesis. The eigenvalues of the $\left(P^0+P^0_{\boldsymbol{\delta}}\right)\times\left(P^0+P^0_{\boldsymbol{\delta}}\right)$ matrix $\left[\mathbf{D}^0_{\mathbf{B}\boldsymbol{\delta}}\cdot\mathbf{\Sigma}^0_{\mathbf{f}\boldsymbol{\delta}}(\theta^0,\theta^0)\right]$ are distinct.

Assumption E2 - Uniform Convergence under the Alternative Hypothesis. $\hat{\mathbf{M}}(\theta_1,\theta_2)$ and $\hat{\mathbf{\Omega}}(\theta_1,\theta_2)$ converge in probability to $\mathbf{M}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}(\theta_1,\theta_2)$ and $\mathbf{\Omega}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}(\theta_1,\theta_2)$, respectively, uniformly over (θ_1,θ_2), where $\mathbf{M}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}(\theta_1,\theta_2)$ and $\mathbf{\Omega}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}(\theta_1,\theta_2)$ are positive definite matrices.

Assumption E1 is similar to Assumption C1 and the former reduces to the latter when $P^0_{\boldsymbol{\delta}}=0$. Assumption E2 generalizes Assumption D3 when $P^0_{\boldsymbol{\delta}}>0$. Notice that $\mathbb{I}_{1t}(\theta)\,\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0)=\left[\mathbb{I}_{1t}(\theta)\,\mathbf{f}^{0\prime}_t,\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{2t}(\theta^0)\,\mathbf{f}^{0\prime}_{\boldsymbol{\delta}t}\right]'$: thus $\mathbb{I}_{1t}(\theta)\,\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0)=\left[\mathbb{I}_{1t}(\theta)\,\mathbf{f}^{0\prime}_t,\mathbf{0}'\right]'$ for $\theta\le\theta^0$, and $\mathbf{\Omega}^0_{\boldsymbol{\delta}}(\theta_1,\theta_2)$ and $\mathbf{M}^0_{\boldsymbol{\delta}}(\theta_1,\theta_2)$ are rank deficient for some (θ_1,θ_2). However,

$$\begin{aligned}\mathbb{I}_{1t}(\theta)\left[\mathbf{H}^0_{\boldsymbol{\delta}}(\theta^0)\right]^{-1}\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0) &= \left[\mathbf{H}^0_*(\theta^0),\mathbf{H}^0_{*2}(\theta^0)\right]\left[\mathbb{I}_{1t}(\theta)\,\mathbf{f}^{0\prime}_t,\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{2t}(\theta^0)\,\mathbf{f}^{0\prime}_{\boldsymbol{\delta}t}\right]'\\ &= \mathbb{I}_{1t}(\theta)\,\mathbf{H}^0_*(\theta^0)\,\mathbf{f}^0_t+\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{2t}(\theta^0)\,\mathbf{H}^0_{*2}(\theta^0)\,\mathbf{f}^0_{\boldsymbol{\delta}t},\end{aligned}$$

with $\left[\mathbf{H}^0_{\boldsymbol{\delta}}(\theta^0)\right]^{-1}=\left[\mathbf{H}^0_*(\theta^0),\mathbf{H}^0_{*2}(\theta^0)\right]$, where $\mathbf{H}^0_*(\theta^0)$ and $\mathbf{H}^0_{*2}(\theta^0)$ are $\left(P^0+P^0_{\boldsymbol{\delta}}\right)\times P^0$ and $\left(P^0+P^0_{\boldsymbol{\delta}}\right)\times P^0_{\boldsymbol{\delta}}$ matrices, respectively: $\mathbb{I}_{1t}(\theta)\left[\mathbf{H}^0_{\boldsymbol{\delta}}(\theta^0)\right]^{-1}\mathbf{f}^0_{\boldsymbol{\delta}t}(\theta^0)=\mathbb{I}_{1t}(\theta)\,\mathbf{H}^0_*(\theta^0)\,\mathbf{f}^0_t$ for $\theta\le\theta^0$, and positive definiteness of $\mathbf{M}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}(\theta_1,\theta_2)$ and $\mathbf{\Omega}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}(\theta_1,\theta_2)$ in Assumption E2 is not restrictive.

Theorem 3.3 *Let Assumptions A1-A4, B1, B2, E1, E2 hold. Then under $\mathcal{H}_1$ and as $N,T\to\infty$: (a) there exists some $\left(P^0+P^0_{\boldsymbol{\delta}}\right)\times1$ nonrandom vector $\mathbf{c}\neq\mathbf{0}$ such that $\left[\hat{\overline{\boldsymbol{\beta}}}_{\mathbf{w}1}(\theta^0)-\hat{\overline{\boldsymbol{\beta}}}_{\mathbf{w}2}(\theta^0)\right]\xrightarrow{p}\mathbf{c}$; (b) $\mathrm{P}\left[\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}(\theta^0)>c\right]\to1$ and $\mathrm{P}\left(\sup\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}>c\right)\to1$ for any constant $c\in\mathfrak{R}$.*

Theorem 3.3 resembles Theorem 4 in Han and Inoue (2015). Theorem 3.3(a) shows that the estimators for $\left(\sum_{i=1}^{N}w_i\boldsymbol{\beta}^0_{1i}\right)'$ and $\left(\sum_{i=1}^{N}w_i\boldsymbol{\beta}^0_{2i}\right)'$ converge to different probability limits: this is why our test detects regime changes in the loadings. Theorem 3.3(b) implies that the probability limit of $\left[\hat{\overline{\boldsymbol{\beta}}}_{\mathbf{w}1}(\theta^0)-\hat{\overline{\boldsymbol{\beta}}}_{\mathbf{w}2}(\theta^0)\right]$ transfers to the $\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}(\theta^0)$ and $\sup\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}$ statistics and makes the tests consistent. Theorem 3.3 does not require *a priori* knowledge of the true number of factors P^0. According to Remark B in Breitung and Eickmeier (2011), a test for structural stability that compares estimated risk exposures may lack power when the number of factors is determined from the entire sample: this is because a factor model with a structural break admits an equivalent representation with a larger number of factors and constant risk exposures. By Proposition 3.1, the estimator $\tilde{P}$ for P^0 obtained from the entire sample overestimates P^0: however, Theorem 3.3 shows that our test detects regime changes.

3.7 Fixed regressor bootstrap

The critical values of the asymptotic distribution of $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ cannot be tabulated. We thus implement the fixed regressor bootstrap of Hansen (1996). For $b = 1, \ldots, B$: (i) generate $u_{bt}^* \sim \text{IID}\mathcal{N}(0,1)$, for $t = 1, \ldots, T$; (ii) let $\hat{\mathbf{k}}_b^*(\theta) = T^{-1/2} \sum_{t=1}^T \tilde{\mathbf{f}}_t(\theta) \tilde{\tilde{e}}_{\mathbf{w}t} u_{bt}^*$; (iii) define $\sup \widehat{LM}_{\mathbf{w},b}^{\text{HAC},*} = \sup_\theta \widehat{LM}_{\mathbf{w},b}^{\text{HAC},*}(\theta)$, where

$$
\begin{aligned}
& LM_{\mathbf{w},b}^{\text{HAC},*}(\theta) \\
= & \left[\hat{\mathbf{M}}(\theta,\theta)^{-1} \hat{\mathbf{k}}_b^*(\theta)\right]' \mathbf{G} \left[\mathbf{G}'\hat{\mathbf{M}}(\theta,\theta)^{-1} \hat{\mathbf{\Omega}}_{\mathbf{w}}(\theta,\theta) \hat{\mathbf{M}}(\theta,\theta)^{-1} \mathbf{G}\right]^{-1} \mathbf{G}' \left[\hat{\mathbf{M}}(\theta,\theta)^{-1} \hat{\mathbf{k}}_b^*(\theta)\right].
\end{aligned}
$$

The empirical distribution of $\left\{\sup \widehat{LM}_{\mathbf{w},b}^{\text{HAC},*}\right\}_{b=1}^B$ approximates the limiting distribution of $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ under $\mathcal{H}_0$. Formally, let P^* denote the bootstrap probability measure conditional upon the original data: the following theorem states the validity of the fixed regressor bootstrap.

Theorem 3.4 (a) *Let Assumptions A2-A4, B1, C1, C2, D1-D3 hold and $N, T \to \infty$. Then for any constant $c \in \mathfrak{R}^+$, under $\mathcal{H}_0$:*

$$
\text{P}\left[\sup_{q \in \mathfrak{R}^+} \left|\text{P}^*\left(\sup \widehat{LM}_{\mathbf{w},b}^{\text{HAC},*} \leq q\right) - \text{P}\left(\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}} \leq q\right)\right| > c\right] \to 0.
$$

(b) *Let Assumptions A1-A4, B1, B2, E1, E2 hold and $N, T \to \infty$. Then for any constant $c \in \mathfrak{R}^+$, under $\mathcal{H}_1$:*

$$
\text{P}\left(\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}} - \sup \widehat{LM}_{\mathbf{w},b}^{\text{HAC},*} > c\right) \to 1.
$$

3.8 Robustness to unknown number of factors

Theorems 3.1 and 3.2 require consistent estimation of the number of factors under the null hypothesis. Theorem 3.3 is valid provided that the dimension $P^0 + P_\delta^0$ of the augmented factor space is consistently estimated. Proposition 3.1 relies on Bai and Ng (2002) information criteria, which may overestimate the number of factors in the presence of cross-sectional dependence in the idiosyncratic errors: on this, see Trapani (2018). We show that our procedure remains valid when the number of factors is unknown.

For ease of exposition, we generalize Theorem 3.1: Theorems 3.2 and 3.3 can be dealt with in a similar way. Recall $\tilde{\mathbf{f}}_t^P$ from Section 3.2. Define $\tilde{\mathbf{f}}_t^P\left(\theta^0\right)=\left[\mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t^{P\prime},\mathbb{I}_{2t}\left(\theta^0\right)\tilde{\mathbf{f}}_t^{P\prime}\right]^{\prime}$ and

$$\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}}^P\left(\theta^0\right)=\left[\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}1}^P\left(\theta^0\right)^{\prime},\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}2}^P\left(\theta^0\right)^{\prime}\right]^{\prime}=\left[\sum_{t=1}^T\tilde{\mathbf{f}}_t^P\left(\theta^0\right)\tilde{\mathbf{f}}_t^P\left(\theta^0\right)^{\prime}\right]^{-1}\left[\sum_{t=1}^T\tilde{\mathbf{f}}_t^P\left(\theta^0\right)\bar{R}_{\mathbf{w}t}\right].$$

Let $\hat{\mathbf{M}}^P\left(\theta^0,\theta^0\right)=T^{-1}\sum_{t=1}^T\tilde{\mathbf{f}}_t^P\left(\theta^0\right)\tilde{\mathbf{f}}_t^P\left(\theta^0\right)^{\prime}$ and $\tilde{\mathbf{k}}_t^P\left(\theta^0\right)=\tilde{\mathbf{f}}_t^P\left(\theta^0\right)\widetilde{\widetilde{e}}_{\mathbf{w}t}^P$, where $\widetilde{\widetilde{e}}_{\mathbf{w}t}^P=\bar{R}_{\mathbf{w}t}-\widetilde{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}1}^{P\prime}\tilde{\mathbf{f}}_t^P$ and $\widetilde{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}1}^P=\sum_{i=1}^N w_i\tilde{\boldsymbol{\beta}}_{1i}^P$. Define $\hat{\mathbf{K}}_d^P\left(\theta^0,\theta^0\right)=T^{-1}\sum_{t=d+1}^T\tilde{\mathbf{k}}_t^P\left(\theta^0\right)\tilde{\mathbf{k}}_{t-d}^P\left(\theta^0\right)$, for $d=0,\ldots,D_T$, and $\hat{\boldsymbol{\Omega}}_{\mathbf{w}}^P\left(\theta^0,\theta^0\right)=\hat{\mathbf{K}}_0^P\left(\theta^0,\theta^0\right)+\sum_{d=1}^{D_T}w\left(d,D_T\right)\left[\hat{\mathbf{K}}_d^P\left(\theta^0,\theta^0\right)+\hat{\mathbf{K}}_d^P\left(\theta^0,\theta^0\right)^{\prime}\right]$, with D_T and $w\left(d,D_T\right)$ as in Section 3.3. Define $\mathbf{G}^P=\left(\mathbf{I}_P,-\mathbf{I}_P\right)^{\prime}$. The relevant test statistic is

$$\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC},P}\left(\theta^0\right)=T\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}}^P\left(\theta^0\right)^{\prime}\mathbf{G}^P\left[\mathbf{G}^{P\prime}\hat{\mathbf{M}}^P\left(\theta^0,\theta^0\right)^{-1}\hat{\boldsymbol{\Omega}}_{\mathbf{w}}^P\left(\theta^0,\theta^0\right)\hat{\mathbf{M}}^P\left(\theta^0,\theta^0\right)^{-1}\mathbf{G}^P\right]^{-1}\mathbf{G}^{P\prime}\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}}^P\left(\theta^0\right)$$

and it is such that $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC},\tilde{P}}\left(\theta^0\right)=\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}\left(\theta^0\right)$. Define the $P^0\times P$ rotation matrix

$$\tilde{\mathbf{H}}_1^P=\frac{\mathbf{F}^0\mathbf{F}^{0\prime}}{T}\frac{\mathbf{B}_1^{0\prime}\tilde{\mathbf{B}}_1^P}{N}\left(\tilde{\mathbf{V}}_1^P\right)^{-1},$$

where $\tilde{\mathbf{V}}_1^P$ is the $P\times P$ diagonal matrix of the first P largest eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\mathbf{R}}=\left(NT\right)^{-1}\sum_{t=1}^T\mathbf{R}_t\mathbf{R}_t^{\prime}$ in decreasing order: $\tilde{\mathbf{H}}_1^P$ is such that $\tilde{\mathbf{H}}_1^{\tilde{P}}=\tilde{\mathbf{H}}_1$.

Theorem 3.5 *Let Assumptions A1-A4, B1, C1-C3 hold. Then* $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC},\bar{P}}\left(\theta^0\right)\xrightarrow{d}\chi^2\left(\bar{P}\right)$ *under* $\mathcal{H}_0$ *for any a priori chosen number of factors* $P=\bar{P}$ *such that* $\bar{P}\geq P^0$*, provided that* $\hat{\boldsymbol{\Omega}}_{\mathbf{w}}^{\bar{P}}\left(\theta^0,\theta^0\right)\xrightarrow{p}\left(\mathbf{I}_2\otimes\mathbf{H}_1^{0,\bar{P}}\right)^{-1}\boldsymbol{\Omega}_{\mathbf{w}}^0\left(\theta^0,\theta^0\right)\left(\mathbf{I}_2\otimes\mathbf{H}_1^{0,\bar{P}\prime}\right)^{-1}$*, where* $\mathbf{H}_1^{0,\bar{P}}=p\lim_{N,T\to\infty}\tilde{\mathbf{H}}_1^{\bar{P}}$.

Theorem 3.5 shows that the $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC},\bar{R}}\left(\theta^0\right)$ statistic can still give a correctly sized test even when the true number of factors is neither known nor consistently estimated: when $\bar{P}>P^0$, the asymptotic distribution $\chi^2\left(\bar{P}\right)$ has more degrees of freedom than $\chi^2\left(P^0\right)$ in Theorem 3.1, and the test may suffer from a power loss.

3.9 Discussion of size properties

Chen *et al.* (2014), Han and Inoue (2015), and Massacci (2017), test for discrete changes in the factor loadings by testing for a change in the covariance matrix of the estimated factors. We now explain why this strategy produces size distortions if the covariance matrix of the true factors is regime-dependent. For ease of exposition, set $\theta = \theta^0$ and $\tilde{P} = P^0$. Define $\hat{\mathbf{\Sigma}}_{\mathbf{f}jj}\left(\theta^0, \theta^0\right) = T^{-1} \sum_{t=1}^{T} \mathbb{I}_{jt}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t'$, for $j = 1, 2$. Following Han and Inoue (2015), consider

$$
\begin{aligned}
& \frac{\sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t'}{\sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right)} - \frac{\sum_{t=1}^{T} \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t'}{\sum_{t=1}^{T} \mathbb{I}_{2t}\left(\theta^0\right)} \\
= & \frac{T}{\sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right)} \hat{\mathbf{\Sigma}}_{\mathbf{f}11}\left(\theta^0, \theta^0\right) - \frac{T}{\sum_{t=1}^{T} \mathbb{I}_{2t}\left(\theta^0\right)} \hat{\mathbf{\Sigma}}_{\mathbf{f}22}\left(\theta^0, \theta^0\right).
\end{aligned}
$$

Recall the $P^0 \times P^0$ matrix $\mathbf{H}_1^0$ in Theorem 3.1. Given Assumption A1, and under $\mathcal{H}_0$ in (2), $T^{-1} \sum_{t=1}^{T} \mathbb{I}_{jt}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t' \xrightarrow{p} \left(\mathbf{H}_1^0\right)^{-1} \mathbf{\Sigma}_{\mathbf{f}jj}^0\left(\theta^0, \theta^0\right) \left(\mathbf{H}_1^{0\prime}\right)^{-1}$, for $j = 1, 2$. It follows that

$$
\begin{aligned}
& \frac{T}{\sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right)} \hat{\mathbf{\Sigma}}_{\mathbf{f}11}\left(\theta^0, \theta^0\right) - \frac{T}{\sum_{t=1}^{T} \mathbb{I}_{2t}\left(\theta^0\right)} \hat{\mathbf{\Sigma}}_{\mathbf{f}22}\left(\theta^0, \theta^0\right) \\
\xrightarrow{p} & \frac{1}{\pi_0} \left(\mathbf{H}_1^0\right)^{-1} \mathbf{\Sigma}_{\mathbf{f}11}^0\left(\theta^0, \theta^0\right) \left(\mathbf{H}_1^{0\prime}\right)^{-1} - \frac{1}{1 - \pi_0} \left(\mathbf{H}_1^0\right)^{-1} \mathbf{\Sigma}_{\mathbf{f}22}^0\left(\theta^0, \theta^0\right) \left(\mathbf{H}_1^{0\prime}\right)^{-1} = \mathbf{C},
\end{aligned}
$$

where $\pi_0 = \mathrm{E}\left[\mathbb{I}_{1t}\left(\theta^0\right)\right]$ and $\mathbf{C}$ is a $P^0 \times P^0$ matrix. Recall $T^{-1} \sum_{t=1}^{T} \mathbf{f}_t \mathbf{f}_t' \xrightarrow{p} \mathbf{\Sigma}_{\mathbf{f}}^0$ as defined in Section 2.3. Unless $\mathbf{\Sigma}_{\mathbf{f}11}^0\left(\theta^0, \theta^0\right) = \pi_0 \mathbf{\Sigma}_{\mathbf{f}}^0$ and $\mathbf{\Sigma}_{\mathbf{f}22}^0\left(\theta^0, \theta^0\right) = \left(1 - \pi_0\right) \mathbf{\Sigma}_{\mathbf{f}}^0$ (namely, the covariance matrix of the true factors is constant across regimes), $\mathbf{C} \neq \mathbf{0}_{P^0}$ because $\mathbf{\Sigma}_{\mathbf{f}11}^0\left(\theta^0, \theta^0\right)$ and $\mathbf{\Sigma}_{\mathbf{f}22}^0\left(\theta^0, \theta^0\right)$ are positive definite by Assumption A1: under $\mathcal{H}_0$, a regime change in the covariance matrix of the *estimated* factors would falsely detect a regime change in the loadings when the covariance structure of the *true* factors is regime-specific. This is aligned to Remark 3 in Baltagi *et al.* (2017), which emphasizes that the tests for structural stability of Chen *et al.* (2014), and Han and Inoue (2015), are not robust to heteroskedasticity of the factors.

Chen *et al.* (2014), Han and Inoue (2015), and Massacci (2017), rule out size distortions by imposing that the covariance matrix of the true factors remains constant across states: this is stated in Assumption 2, Assumption 1 and Assumption LT2, respectively. In line with Assumption A in Bai and Ng (2002), our test allows for regime-dependent covariance structure in the true factors: the dimensionality problem is solved by taking cross-sectional averages according to the weighting scheme in Assumption B1.

4 Link to portfolio choice and diversification

In (1) the common components $\mathbb{I}_{1t}(\theta)\mathbf{B}_1\mathbf{f}_t + \mathbb{I}_{2t}(\theta)\mathbf{B}_2\mathbf{f}_t$ and the errors $\mathbf{e}_t$ are sources of systematic and idiosyncratic risk, respectively. Assumption B1 has implications for portfolio construction: the weights of order $1/N$ and Assumption A3(c) guarantee that $\text{Var}(\bar{e}_{\mathbf{w}t}) \to 0$ as $N \to \infty$; together with Assumption A3(a), this ensures that idiosyncratic risk is diversified away in the limit (see Proof of Lemma A.1 in the Appendix). Recall $\bar{R}_{\mathbf{w}t}$ as defined in (5): as $N \to \infty$, $\{\bar{R}_{\mathbf{w}t}\}_{t=1}^{T}$ is the sequence of returns from a portfolio that only bears systematic risk; under $\mathcal{H}_0$ in (2) the regime shifts in the loadings are negligible and do not affect the risk profile of this portfolio.

Under $\mathcal{H}_1$ in (3) the risk exposures are regime-specific and we estimate number of factors, risk exposures and factors as in Massacci (2017). Let $\hat{\mathbf{B}}_j$ be the $N \times \hat{P}$ matrix of estimated regime-specific risk exposures, where $\hat{P}$ is the estimated number of factors, for $j = 1, 2$. Due to rotational indeterminacy, without further restrictions $\hat{\mathbf{B}}_j$ does not allow to identify the sign of the true risk exposures. In financial markets the first factor is likely to be a level factor, as documented in Fama and French (1993), and Lustig *et al.* (2011), for equity and foreign exchange markets, respectively. To match the sign of the true risk exposures, we multiply $\hat{\mathbf{B}}_j$ by $+1$ or -1 depending on whether the correlation between the first estimated factor and the return on the market within the corresponding regime is positive or negative, respectively.

We construct portfolio weights from the estimated risk exposures by extending the approach of Lehmann and Modest (2005) to allow for the presence of regimes: the resulting portfolios bear no idiosyncratic risk as $N \to \infty$. Since $\hat{\mathbf{B}}_j'\hat{\mathbf{B}}_j/N = \mathbf{I}_{\hat{P}}$, from Lehmann and Modest (2005) the $N \times \hat{P}$ matrix of portfolio weights is $\hat{\mathbf{B}}_j\left(\hat{\mathbf{B}}_j'\hat{\mathbf{B}}_j\right)^{-1} = \hat{\mathbf{B}}_j/N$. The weights $\hat{\mathbf{B}}_j/N$ do not necessarily add up to unity: as in Lehmann and Modest (2005), we normalize them to ensure such a condition is met.

We then study the dynamics of portfolio diversification across the two regimes. We follow Pukthuanthong and Roll (2009) and measure diversification through the $\text{R}-\text{squared}$ from the multifactor model: the higher the $\text{R}-\text{squared}$, the lower the benefits from diversification as factors capture a higher degree of comovement across assets.

5 Monte Carlo analysis

We conduct six experiments to evaluate the finite sample properties of our test. In line with Section 3.9, Experiment 1 in Section 5.1 provides a size comparison with the test proposed in Massacci (2017) when the covariance matrix of the factors depends upon the regimes. We then turn to a deep analysis of the test we propose in this paper. Experiment 2 in Section 5.2 assesses size and power. Experiment 3 in Section 5.3 focuses on α^0, which determines the number of cross-sectional units subject to threshold effect. Experiment 4 in Section 5.4 studies how the test performs when the number of factors is unknown. Experiment 5 in Section 5.5 investigates how idiosyncratic components with heavy-tailed distribution affect the finite sample properties of the test: this is important in light of Assumption A3(a), according to which the idiosyncratic components have eight finite moments. Experiment 6 in Section 5.6 studies how the dimension of the factor space shapes size and power of the test. Finally, Section 5.7 provides a discussion of the whole set of results. In all experiments we consider 95% statistical coverage, which corresponds to 5% size. Let $s = 1, \ldots, S$ refer to the replication; for the generic test statistic $\hat{\mathcal{T}}^s$, for $s = 1, \ldots, S$, we report the frequency of violations

$$\hat{v} = \frac{1}{S} \sum_{s=1}^{S} \mathbb{I}\left(\hat{\mathcal{T}}^s > q_{0.95}\right), \tag{10}$$

where $q_{0.95}$ is the 0.95 quantile of the distribution of $\hat{\mathcal{T}}^s$ under the null hypothesis. We fix $S = 2000$ in all experiments.

5.1 Experiment 1: size comparison

5.1.1 Data generating process

The data generating process (DGP) for the observable dependent variables is the linear two-factor model

$$R_{it}^s = \beta_{1i1}^0 f_{1t}^{0s} + \beta_{1i2}^0 f_{2t}^{0s} + e_{it}^s, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T. \tag{11}$$

We fix the risk exposures $\beta_{1ip}^0 \sim \mathcal{N}(1, 1)$ in repeated samples, for $p = 1, 2$. We generate the factors as

$$f_{pt}^{0s} = \frac{\mathbb{I}\left(z_t^s \leq \theta^0\right) \lambda_{f1}^0 f_{ft}^s + \mathbb{I}\left(z_t^s > \theta^0\right) \lambda_{f2}^0 f_{ft}^s + \epsilon_{fpt}^s}{\left[\pi^0 \left(\lambda_{f1}^0\right)^2 + \left(1 - \pi^0\right) \left(\lambda_{f2}^0\right)^2 + 1\right]^{1/2}}, \quad p = 1, 2, \quad t = 1, \ldots, T,$$

with $f_{ft}^{s} \sim \text{IID}\mathcal{N}(0,1)$ and $\epsilon_{fpt}^{s} \sim \text{IID}\mathcal{N}(0,1)$, and $\theta^0 = 2$ fixed in repeated samples. We let $z_t^s = \mu_z + \epsilon_{zt}^s$, with $\epsilon_{zt}^s \sim \text{IID}\mathcal{N}(0,1)$ and μ_z fixed in repeated samples. Given $\pi^0 = \text{P}\left(z_t^s \leq \theta^0\right) = \text{P}\left(z_t^s - \mu_z \leq \theta^0 - \mu_z\right) = \Phi\left(\theta^0 - \mu_z\right)$, we have $\mu_z = \theta^0 - \Phi^{-1}\left(\pi^0\right)$: given θ^0, μ_z controls for π^0. In this way,

$$\text{Var}\left(f_{pt}^{0s} \left| z_t^s \leq \theta^0\right.\right) = \frac{\lambda_{f1}^2 + 1}{\pi^0 \lambda_{f1}^2 + \left(1 - \pi^0\right)\lambda_{f2}^2 + 1}, \qquad \text{Var}\left(f_{pt}^{0s} \left| z_t^s > \theta^0\right.\right) = \frac{\lambda_{f2}^2 + 1}{\pi^0 \lambda_{f1}^2 + \left(1 - \pi^0\right)\lambda_{f2}^2 + 1},$$

$$\text{Corr}\left(f_{1t}^{0s}, f_{2t}^{0s} \left| z_t^s \leq \theta^0\right.\right) = \frac{\lambda_{f1}^2}{\lambda_{f1}^2 + 1}, \qquad \text{Corr}\left(f_{1t}^{0s}, f_{2t}^{0s} \left| z_t^s > \theta^0\right.\right) = \frac{\lambda_{f2}^2}{\lambda_{f2}^2 + 1}:$$

we fix $\lambda_{f1} = 1$ in repeated samples and define $\delta_{\mathbf{f}} = \lambda_{f2} - \lambda_{f1}$ to control for regime changes in the covariance matrix of the factors: when $\delta_{\mathbf{f}} = 0$, the factors covariance matrix is independent of the regimes. Finally, we generate the idiosyncratic components as $e_{it}^s = \sigma_{ii}^{1/2} \epsilon_{e_i t}^s$, with $\epsilon_{e_i t}^s \sim \text{IID}\mathcal{N}(0,1)$ and $\sigma_{ii} \sim \chi^2(1)$ fixed in repeated samples.

5.1.2 Results

Consistently with the DGP described in Section 5.1.1, we compute $\widehat{LM}_{\mathbf{w}}(\theta)$ in (8) with equal weights $w_i = 1/N$, for $i = 1, \ldots, N$. We set $\theta = \theta^0$: under the null hypothesis, $\widehat{LM}_{\mathbf{w}}\left(\theta^0\right)$ has a χ^2 limiting distribution with R^0 degrees of freedom as $N, T \to \infty$. We also compute the Lagrange multiplier statistic of Massacci (2017), which we denote $\widehat{LM}_{\mathbf{f}}\left(\theta^0\right)$: under the null hypothesis, and provided that the covariance matrix of the factors is independent of the state, this statistic has a χ^2 limiting distribution with $\left(P^0 - 1\right)$ degrees of freedom as $N, T \to \infty$. To allow for a meaningful comparison, $\widehat{LM}_{\mathbf{f}}\left(\theta^0\right)$ is robust neither to time heteroskedasticity nor to serial correlation.

Table 1 about here

Table 1 collects the results. We set $N = 50, 100$, and $T = 100, 200, 400$. We consider $\pi^0 = 0.15, 0.30, 0.50, 0.70, 0.85$ to control for a wide range of regime probabilities. We compute the size for $\delta_{\mathbf{f}} = 0.00, 0.25, 1.00, 1.75$. The $\widehat{LM}_{\mathbf{f}}\left(\theta^0\right)$ statistic has the correct size when $\delta_{\mathbf{f}} = 0.00$, namely when the covariance matrix of the factors is independent of the regimes; the size distortion becomes more pronounced as $\delta_{\mathbf{f}}$ increases for any N, T and π^0. On the other hand, the $\widehat{LM}_{\mathbf{w}}\left(\theta^0\right)$ statistic always delivers the correct size regardless of the degree of regime-dependence in the covariance matrix of the factors.

5.2 Experiment 2: size and power

5.2.1 Data generating process

We compute the size of the test proposed in this paper using the DGP for x_{it}^s in (11), with $\beta_{1ip}^0 \sim \mathcal{N}(1,1)$ fixed in repeated samples, and f_{pt}^{0s} as in Experiment 5.1, for $p = 1, 2$. The threshold variable z_t^s is as in Experiment 1. We let the idiosyncratic components e_{it}^s be time and cross-sectionally dependent and heteroskedastic with DGP

$$e_{it}^s = \rho_e e_{i,t-1}^s + \sigma_{ii}^{1/2}\left(1-\rho_e^2\right)^{1/2}\varpi_{e_it}^s\epsilon_{e_it}^s, \quad e_{i,-50}^s = 0, \quad i = 1, \ldots, N, \quad t = -49, \ldots, 0, \ldots, T,$$

with $\rho_e = 0.50$ and $\sigma_{ii} \sim \chi^2(1)$ fixed in repeated samples. The conditional volatility $\varpi_{e_it}^s$ follows the GARCH$(1,1)$ process

$$\left(\varpi_{e_it}^s\right)^2 = \lambda_{e1} + \lambda_{e2}\left(\varpi_{e_i,t-1}^s\right)^2 + \lambda_{e3}\left(\varpi_{e_i,t-1}^s\epsilon_{e_i,t-1}^s\right)^2, \quad \left(\varpi_{e_i,-50}^s\right)^2 = \mathrm{E}\left[\left(\varpi_{e_it}^s\right)^2\right] = 1.$$

Let $\boldsymbol{\epsilon}_{et}^s = \left(\epsilon_{e_1t}^s, \ldots, \epsilon_{e_Nt}^s\right)'$. We allow for cross-sectional dependence through the first order spatial autoregressive process $\boldsymbol{\epsilon}_{et}^s = \bar{\mathbf{G}}\boldsymbol{\varrho}_{et}^s$, where

$$\bar{\mathbf{G}} = \mathbf{G}\left[\frac{N}{\mathrm{tr}\left(\boldsymbol{\Sigma}_{\mathbf{e},\mathrm{diag}}^{1/2}\mathbf{G}\mathbf{G}'\boldsymbol{\Sigma}_{\mathbf{e},\mathrm{diag}}^{1/2}\right)}\right]^{1/2}, \qquad \boldsymbol{\Sigma}_{\mathbf{e},\mathrm{diag}}^{1/2} = \mathrm{diag}\left[\left(\sigma_{11}^{1/2}, \ldots, \sigma_{NN}^{1/2}\right)'\right],$$

with $\boldsymbol{\varrho}_{et}^s \sim \mathrm{IID}\mathcal{N}(\mathbf{0}, \mathbf{I}_N)$, $\mathbf{G} = (\mathbf{I}_N - g\mathbf{V})^{-1}$, where g regulates the degree of cross-sectional dependence: $\mathbf{V} = (v_{il})$ is a rook-type matrix, namely all elements in $\mathbf{V}$ are zero except $v_{i+1,i} = v_{l-1,l} = 0.5$, for $i = 1, \ldots, N-2$ and $l = 3, \ldots, N$, with $v_{12} = v_{N,N-1} = 1$. It follows that $\mathrm{Var}\left(e_{it}^s\right) = \left[\sigma_{ii} \Big/ \left(N^{-1}\sum_{i=1}^N \sigma_{ll}\right)\right]$, $\lim_{N\to\infty}\mathrm{Var}\left(e_{it}^s\right) = \sigma_{ii}$ and $N^{-1}\sum_{i=1}^N \mathrm{Var}\left(e_{it}^s\right) = 1$. We consider three scenarios: (i) time homoskedastic and cross-sectionally independent idiosyncratic components ($\mathrm{CSI}_{\mathbf{e}}$); (ii) time homoskedastic and cross-sectionally dependent idiosyncratic components ($\mathrm{CSD}_{\mathbf{e}}$); (iii) time heteroskedastic and cross-sectionally dependent idiosyncratic components ($\mathrm{CSDH}_{\mathbf{e}}$). Under $\mathrm{CSI}_{\mathbf{e}}$ we set $\lambda_{e1} = 1$, $\lambda_{e2} = 0$, $\lambda_{e3} = 0$ and $g = 0$. We build $\mathrm{CSD}_{\mathbf{e}}$ by imposing $\lambda_{e1} = 1$, $\lambda_{e2} = 0$, $\lambda_{e3} = 0$ and $g = 0.4$. We parameterize $\mathrm{CSDH}_{\mathbf{e}}$ by setting $\lambda_{e1} = 0.1$, $\lambda_{e2} = 0.8$, $\lambda_{e3} = 0.1$ and $g = 0.4$. We reduce the effect of the initial values $e_{i,-50}^s = 0$ and $\varpi_{e_i,-50}^s = 1$ by discarding the first 50 observations in the DGPs for $e_{it}^s = 0$ and $\varpi_{e_it}^s = 1$.

We obtain the power of the test by simulating the data from the one-factor DGP

$$R_{it}^{s}=\mathbb{I}\left(z_{t}^{s}\leq\theta^{0}\right)\beta_{1i1}^{0}f_{1t}^{0s}+\mathbb{I}\left(z_{t}^{s}>\theta^{0}\right)\beta_{2i1}^{0}f_{1t}^{0s}+e_{it}^{s},\quad i=1,\ldots,N,\quad t=1,\ldots,T,$$

where z_t^s and θ^0 are as in Experiment 1. We control for the shift in the risk exposures through $\delta_i^0=\beta_{2i1}^0-\beta_{1i1}^0$, with $\beta_{1i1}^0\sim\mathcal{N}(1,1)$, and we set $\delta_i^0=\delta^0$, for $i=1,\ldots,N$: we obtain the power for $\delta^0>0$. We generate f_{1t}^{0s} as in Experiment 5.1. The idiosyncratic components e_{it}^s are generated under the three scenarios $\mathrm{CSI}_{\mathbf{e}}$, $\mathrm{CSD}_{\mathbf{e}}$ and $\mathrm{CSDH}_{\mathbf{e}}$ described above and the first 50 observations are again discarded.

5.2.2 Results

Table 2 displays size and power for $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}(\theta)$ in (6) with $\theta=\theta^0$, and for $\sup\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}$ in (7) (see Panels A and B, respectively). In the latter case, in each replication we compute the maximum of $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}(\theta)$ by selecting 19 equally spaced quantiles of the empirical distribution of z_t^s, namely $\{5\%,10\%,15\%,\ldots,85\%,90\%,95\%\}$, and the true value $\theta^0=2$; we opt for $B=1000$ fixed regressor bootstrap replications. We fix $N=50,100$ in both cases. For $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}\left(\theta^0\right)$, we set $T=100,200,400$ and $\pi^0=0.15,0.30,0.50$; we consider $T=100,200,400,1000$ and $\pi^0=0.50$ for $\sup\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}$. Regime shifts in the DGP for f_{pt}^{0s} are allowed through $\delta_{\mathbf{f}}=0.00,1.75$, for $p=1,2$. In computing the power, the threshold effect on the risk exposures is controlled for by setting $\delta^0=0.25,1.00$.

Table 2 about here

The empirical size of $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}\left(\theta^0\right)$ matches the theoretical value for $T=400$ regardless of N, $\delta_{\mathbf{f}}$, π^0 and the DGP of the idiosyncratic components. The power increases in N, T and δ^0; other conditions being equal, it is maximized at $\pi^0=0.50$ and under the $\mathrm{CSI}_{\mathbf{e}}$ scenario. The power decreases in $\delta_{\mathbf{f}}$ as it becomes harder to identify regime shifts in the loadings from changes in factor dynamics. The size of $\sup\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}$ improves as T increases; the power depends on δ^0 and $\delta_{\mathbf{f}}$ in a similar way as that of $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}\left(\theta^0\right)$. Our test thus has good size and power under different scenarios.

5.3 Experiment 3: α^0 and power

5.3.1 Data generating process

We assess the effect of α^0 on the power of the test under time heteroskedasticity in the factors. We take as a starting point Experiment 2 under the CSDH$_{\mathbf{e}}$ scenario for the idiosyncratic components as described in Section 5.2.1; to compute size and power, we depart from the linear two-factor model and the threshold one-factor model, respectively, in two ways. We generate the factors f_{pt}^{0s} as

$$f_{pt}^{0s} = \varpi_{fpt}^{s} \epsilon_{\epsilon pt}^{s}, \quad p = 1, 2, \quad t = -49, \ldots, 0, \ldots, T,$$

with $\epsilon_{\epsilon pt}^{s} \sim \text{IID}\mathcal{N}(0, 1)$; the conditional volatility ϖ_{fpt}^{s} follows the GARCH$(1,1)$ process

$$\left(\varpi_{fpt}^{s}\right)^2 = \lambda_{f1} + \lambda_{f2}\left(\varpi_{fp,t-1}^{s}\right)^2 + \lambda_{f3}\left(\varpi_{fp,t-1}^{s}\epsilon_{\epsilon p,t-1}^{s}\right)^2, \quad p = 1, 2, \quad t = -48, \ldots, 0, \ldots, T,$$

with $\lambda_{f1} = 0.1$, $\lambda_{f2} = 0.8$ and $\lambda_{f3} = 0.1$, and with starting value $\left(\varpi_{fp,-49}^{s}\right)^2 = \mathrm{E}\left[\left(\varpi_{fpt}^{s}\right)^2\right] = 1$. Given $\delta_i^0 = \lambda_{2i1}^0 - \lambda_{1i1}^0$, we set $\delta_i^0 > 0$ for $i = 1, \ldots, \left\lfloor N^{\alpha^0} \right\rfloor$, and $\delta_i^0 = 0$ for $i = \left\lfloor N^{\alpha^0} \right\rfloor + 1, \ldots, N$. To reduce the effect induced by the initial values, we discard the first 50 observations in the DGPs for e_{it}^{s} and f_{pt}^{0s}.

5.3.2 Results

Table 3 displays size and power for $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ in (6) with $\theta = \theta^0$, and for $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ in (7) (see Panels A and B, respectively). In the latter case, the maximum of $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ is computed as in Section 5.2.2. In both cases, we fix $N = 50, 100$ and $\pi^0 = 0.50$. We consider $T = 100, 200, 400$ for $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta^0)$, and $T = 100, 200, 400, 1000$ for $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$; in the latter case, we stick to $B = 1000$. In computing the power, we control for the threshold effect through $\delta_i^0 = 0.25, 1.00$, for $i = 1, \ldots, \left\lfloor N^{\alpha^0} \right\rfloor$, with $\alpha^0 = 0.60, 0.80, 1.00$.

Table 3 about here

For given values of N, T and δ_i^0, the power of both $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta^0)$ and $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ monotonically increases in α^0. This result is intuitive: α^0 regulates the number of cross-sectional units $\left\lfloor N^{\alpha^0} \right\rfloor$ subject to a regime shift; the higher α^0, the higher $\left\lfloor N^{\alpha^0} \right\rfloor$ and, other conditions being equal, the stronger the overall threshold effect.

5.4 Experiment 4: unknown number of factors

Table 4 about here

The DGPs to compute size and power are those described in Section 5.3.1: for the power, we fix $\alpha^0 = 1$. In line with Theorem 3.5, Table 4 displays results for $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}\left(\theta^0\right)$. In computing the statistic, we augment the number of estimated factors that would be obtained according to Proposition 3.1 by two and four units: this gives us $\bar{P} = P^0 + 2 = 4$ and $\bar{P} = P^0 + 4 = 6$ to compute the size, and $\bar{P} = 2P^0 + 2 = 4$ and $\bar{P} = 2P^0 + 4 = 6$ for the power. The results are highly encouraging. Regardless of the number of redundant factors, the $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}\left(\theta^0\right)$ statistic is correctly sized for $T = 200, 400$, namely for large enough time series dimension. As for the power, the test preserves its good finite sample properties.

5.5 Experiment 5: heavy-tailed distribution of idiosyncratic components

Table 5 about here

We study how fat-tailness in the distribution of the idiosyncratic components impact size and power. We take the DGP in Section 5.3.1 as a starting point and simulate the idiosyncratic components e_{it}^s under the CSDH$_{\mathbf{e}}$ scenario with $\epsilon_{e_it}^s \sim \text{IID}t\,(\text{d.o.f.})$, with number of degrees of freedom d.o.f. $= 10, 15, 20$: for d.o.f. $= 10$ the idiosyncratic components still satisfy Assumption A3(a). As in Experiment 5.4, we report size and power for $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}\left(\theta^0\right)$; in the case of the power, we focus on $\alpha^0 = 1$. The results in Table 5 show that the test is correctly sized and has excellent power properties for all values of the degrees of freedom under consideration: in particular, in line with Assumption A3(a), the test performs well in finite samples for d.o.f. $= 10$.

5.6 Experiment 6: dimension of factor space

Table 6 about here

We assess how the test performs depending on the dimension of the factor space: this is important as the number of restrictions grows linearly with the number of factors P^0 both under the null and under the alternative hypothesis. We employ the DGP described in Section 5.3.1 for $P^0 = 2, 4, 6, 8, 10$. We compute the power for $\alpha^0 = 1$ and $\delta^0 = 0.25$. Given the relatively high number of factors involved, we conduct the experiment over $N = 100, 200$, and $T = 200, 400$. The results displayed in Table 6 show that the test is correctly sized for $(N, T) = (200, 400)$, namely for sufficiently high values of N and T; in the remaining cases the size slightly declines in P^0. Regardless of N and T, the test always has excellent power properties.

5.7 Discussion

Our test displays remarkable finite sample properties. Unlike existing procedures, it is robust to regime changes in the covariance structure of the factors: this reinforces the contribution of our paper in relation to the existing literature. The test displays additional highly desirable features: it has good size and power; it is robust to the inclusion of redundant factors; it is not affected by idiosyncratic components with heavy-tailed distribution; it performs well regardless of the true number of factors. Our test thus is a valid tool for empirical work.

6 Empirical analysis

We use our test to study portfolio choice and diversification across regimes in a high dimensional setting: to the very best of our knowledge, this issue has not been previously addressed in the literature. Section 6.1 describes data and model specification. Section 6.2 discusses the estimation results. Section 6.3 addresses the implications for portfolio choice. Section 6.4 presents estimation results from a large dimensional factor model with structural instability: this allows to highlight the benefits of our approach based on recurring regime changes.

6.1 Data and model specification

The dependent variables are daily excess returns R_{it} from U.S. stock portfolios publicly available from Kenneth French website.[3] We study two groups of variables. We first look at 49 industry portfolios: this choice is aligned to Smith and Timmermann (2020), who employ a set of 30 industry portfolios. We then consider a combined set of $N = 255$ portfolios: 25 portfolios sorted by size and book-to-market ratio; 25 portfolios sorted by size and operating profitability; 25 portfolios sorted by size and investment; 25 portfolios sorted by book-to-market and operating profitability; 25 portfolios sorted by book-to-market and investment; 25 portfolios sorted by operating profitability and investment; 25 portfolios sorted by size and momentum; 25 portfolios sorted by size and short-term reversal; 25 portfolios sorted by size and long-term reversal; 30 industry portfolios. These two groups of variables allow us to assess how highly different cross-

[3]The data are available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

sectional dimensions and sorting schemes affect the empirical results. Our sample period is defined as January 2, 1985, through February 28, 2020, which gives $T = 8862$ time series observations.

We assess the empirical relevance of the threshold variable by monitoring two candidates, namely the one-day lagged value of the economic policy uncertainty index for the U.S. of Baker *et al.* (2016) (EPU_{t-1}) and of the momentum factor (MOM_{t-1}).[4] With the former, we look at the portfolio implications of uncertainty as a predictor of equity market regimes: this complements Brogaard *et al.* (2020), who document the effects of economic policy uncertainty on equity returns. The momentum factor measures the recent performance of the underlying index and it is replicable in markets for which the uncertainty index may not be available, such as emerging markets: Asness *et al.* (2013) show evidence of momentum return premia on several markets and asset classes; compared to one-period returns, momentum returns are more persistent and therefore provide a more accurate predictive signal about the regimes. The empirical correlation between EPU_{t-1} and MOM_{t-1} is -0.026: the two threshold variables are thus very mildly negatively correlated; in this way, we can compare the results from two almost orthogonal predictive signals. By construction, bad states are identified by high values of EPU_{t-1} and low values of MOM_{t-1}, namely $\text{EPU}_{t-1} \leq \hat{\theta}$ and $\text{MOM}_{t-1} > \hat{\theta}$, respectively.

The test is carried out using the heteroskedasticity and autocorrelation robust statistic $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ in (6): due to the daily frequency of the data, the bound on the number of sample autocovariances is $D_T = 5$; the number of fixed regressor bootstrap replications is $B = 1000$. We opt for the weights $w_i = 1/N$, for $i = 1, \ldots, N$. As detailed in Section 3.3, we fit a linear factor model to the portfolio returns. We obtain the estimate $\tilde{P}$ for the number of factors P^0 using the $IC_{p2}(P)$ criterion of Bai and Ng (2002).

Under the alternative hypothesis we estimate the model in (1) and select the number of factors as proposed in Massacci (2017): the estimator for the threshold parameter is consistent for any *a priori* given number of factors greater than or equal to the true number; the $IC_{p2}(P, P)$ criterion consistently estimates the true number of factors provided that this does not depend upon the regimes.[5] If this last assumption does not hold, the $IC_{p2}(P, P)$ criterion gives an upper bound on the true number of factors within each state, as the factor space cannot be

[4] The economic policy uncertainty index of Baker *et al.* (2016) is available at http://www.policyuncertainty.com/. The momentum factor is available from Kenneth French website.

[5] Monte Carlo simulations carried out in Massacci (2017) show that the $IC_{p2}(P, P)$ criterion performs better in finite samples than the other proposed criteria.

spanned with an insufficient number of factors: we thus apply the $IC_{p2}(P)$ criterion of Bai and Ng (2002) within each state and set the estimate obtained from $IC_{p2}(P, P)$ as an upper bound.

6.2 Results

6.2.1 Industry portfolios

In the case of the 49 industry portfolios, the realizations $\sup \widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}} = 37.598$ and $\sup \widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}} = 25.432$ for EPU_{t-1} and MOM_{t-1}, respectively, correspond to p-values equal to 0.000 and 0.004, respectively: these outcomes provide evidence against the null hypothesis of linearity at any conventional significance level for both threshold variables under consideration.

In the empirical specification involving EPU_{t-1}, the estimated threshold value is $\hat{\theta} = 93.859$: the events $\mathrm{EPU}_{t-1} \leq \hat{\theta}$ and $\mathrm{EPU}_{t-1} > \hat{\theta}$ occur with sample frequencies approximately equal to 0.60 and 0.40, respectively. The variable $\mathbb{I}\left(\mathrm{EPU}_{t-1} > \hat{\theta}\right)$ has positive correlation equal to 0.15 with the NBER U.S. recession indicator[6]: this finding resembles the results of Henkel *et al.* (2011), who show that regimes in stock return dynamics are related to the business cycle. When we consider MOM_{t-1}, the estimated threshold $\hat{\theta} = -0.96\%$ identifies the regimes $\mathrm{MOM}_{t-1} \leq \hat{\theta}$ and $\mathrm{MOM}_{t-1} > \hat{\theta}$ with sample frequencies of 0.07 and 0.93, respectively. The correlation between $\mathbb{I}\left(\mathrm{MOM}_{t-1} \leq \hat{\theta}\right)$ and the NBER U.S. recession indicator is 0.05: albeit positive, this number is lower than the corresponding figure obtained for EPU_{t-1}. For both threshold variables the estimated number of factors within each regime is equal to 2: we therefore do not find any evidence of a change in the number of factors between the regimes.

Figure 1 about here

In order to provide further interpretation of the regimes identified by EPU_{t-1} and MOM_{t-1}, we study the sequences of monthly averages of the dummy variables $\mathbb{I}\left(\mathrm{EPU}_{t-1} > \hat{\theta}\right)$ and $\mathbb{I}\left(\mathrm{MOM}_{t-1} \leq \hat{\theta}\right)$: these are interpretable as the monthly probabilities of being in the bad state as identified by the corresponding threshold variable and are displayed in Figure 1. The correlation coefficient between the two sequences is approximately equal to 0.23. Economic policy uncertainty captures the following noticeable episodes of market distress: the early 1990s; the turmoil generated by the September 11, 2001, terrorist attack; the aftermath of Lehmann Brothers collapse in September 2008. As compared to policy uncertainty, momentum misses

[6] The daily NBER recession indicator is publicly available at https://fred.stlouisfed.org/series/USRECDM .

the early 1990s, but it generates a spike in October 1987.

Finally, we assess whether the assumptions imposed on the model and stated in Section 2.2 are empirically plausible. In particular, under Assumption A3(a) the idiosyncratic components e_{it} in (1) have eight finite moments: this may be restrictive when dealing with heavy-tailed equity market returns. Let $\hat{e}_{it}$ be the estimator for e_{it}, for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. We assess the empirical validity of Assumption A3(a) in two ways: we study the sequence $(\tau N)^{-1} \sum_{t=1}^{\tau} \sum_{i=1}^{N} |\hat{e}_{it}|^8$, for $\tau = 1, \ldots, T$; we analyze the right tail of the distribution of the $N \times T$ realizations of the positive random variable $|\hat{e}_{it}|^8$. The original sample of daily excess returns R_{it} exhibits outliers, which affect the realizations of $|\hat{e}_{it}|^8$ when both EPU_{t-1} and MOM_{t-1} are used as threshold variables. In both cases, we winsorize the top 0.5% of the empirical distribution of $|\hat{e}_{it}|^8$: the sequence $(\tau N)^{-1} \sum_{t=1}^{\tau} \sum_{i=1}^{N} |\hat{e}_{it}|^8$, for $\tau = 1, \ldots, T$, does not exhibit any sizeable structural break; the right-hand side of the empirical distribution of $|\hat{e}_{it}|^8$ does not display any noticeable mass concentration apart from the winsorization point. Since we winsorised only the top 0.5% of the distribution of $|\hat{e}_{it}|^8$, we interpret these two findings as evidence that Assumption A3(a) is not restrictive for empirical purposes.

6.2.2 Combined portfolios

Evidence of regimes is also found when we consider the larger set of 255 combined portfolios: the test statistics $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}} = 155.700$ and $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}} = 113.920$ for EPU_{t-1} and MOM_{t-1}, respectively, both have p-values equal to 0.000. In the former case, the threshold value $\hat{\theta} = 122.858$ splits the sample into the events $\text{EPU}_{t-1} \leq \hat{\theta}$ and $\text{EPU}_{t-1} > \hat{\theta}$ with sample frequencies approximately equal to 0.75 and 0.25, respectively: the event $\text{EPU}_{t-1} > \hat{\theta}$ happens less frequently than it does with industry portfolios; the correlation between $\mathbb{I}\left(\text{EPU}_{t-1} > \hat{\theta}\right)$ and the NBER U.S. recession indicator is 0.16. As for MOM_{t-1}, the estimated regimes are perfectly synchronized with those from industry portfolios since the estimates for the threshold value are the same.

Figure 2 about here

Figure 2 plots the monthly averages of $\mathbb{I}\left(\text{EPU}_{t-1} > \hat{\theta}\right)$ and $\mathbb{I}\left(\text{MOM}_{t-1} \leq \hat{\theta}\right)$: they have correlation approximately equal to 0.18; the former series signals episodes of distress that are analogous to those identified from industry portfolios. As for the estimated number of factors, these remain stable across regimes and are equal to 6 and 7 when EPU_{t-1} and MOM_{t-1} are selected as threshold variable, respectively: as compared to industry portfolios, the higher number of estimated factors is likely to be due to the additional sorting schemes involved. Finally, the empirical properties of $|\hat{e}_{it}|^8$ are analogous to those discussed in Section 6.2.1 for both threshold variables, which confirms that Assumption A3(a) is not restrictive.

6.3 Implications for portfolio choice

Regime-specific portfolio weights come from a switching strategy that mimics the estimated factors (see Section 4). The average weight is $1/N$ by construction and it is equal to 0.02 and 0.004 in the case of industry and combined portfolios, respectively.

Table 7 about here

Table 8 about here

Tables 7 and 8 provide summary statistics for the resulting portfolio weights, and the correlations between the two sequences of weights for each factor. In the case of industry portfolios, Table 7 shows that, regardless of the threshold variable, the weights associated to $\hat{f}_{1t}$ have the lowest standard deviation and all are of positive sign: $\hat{f}_{1t}$ is likely to be an estimated market factor. As for $\hat{f}_{2t}$, the corresponding weights allow for long and short positions, since their maximum and minimum values are positive and negative, respectively; these weights also tend to switch sign between regimes, as evidenced by the negative correlation between the two sequences. Table 8 shows that analogous arguments hold for the set of combined portfolios: the weights related to $\hat{f}_{1t}$ are always positive and display very low standard deviation; all other factors lead to both long and short positions.

Figure 3 about here

Figure 4 about here

As discussed in Section 4, we follow Pukthuanthong and Roll (2009) and study the diversification benefits through the $\mathrm{R}-$squared of the model. For each of the four combinations of dependent variables (i.e., returns from industry and combined portfolios) and threshold variables (i.e., EPU_{t-1} and MOM_{t-1}) as described in Section 6.1, we compute the $\mathrm{R}-$squared from the regime-specific regressions of each portfolio return on a given factor. We then compute the average values of the $\mathrm{R}-$squared associated to each factor within each regime. Finally, since the estimated factors are mutually orthogonal by construction, the sum of the individual average $\mathrm{R-squared}$ measures diversification. Figures 3 and 4 display the results for industry and combined portfolios, respectively. In the former case, when $z_t = \mathrm{EPU}_{t-1}$ the averages for the first and second factors are 0.48 and 0.04, respectively, for $\mathrm{EPU}_{t-1} \leq \hat{\theta}$, and 0.64 and 0.03 for $\mathrm{EPU}_{t-1} > \hat{\theta}$ (see Panel A in Figure 3): this leads to an average $\mathrm{R-squared}$ equal to 0.52 and 0.67 for $\mathrm{EPU}_{t-1} \leq \hat{\theta}$ and $\mathrm{EPU}_{t-1} > \hat{\theta}$, respectively; the benefits from diversification thus diminish when economic policy uncertainty is high, as factors become stronger during these periods. For $z_t = \mathrm{MOM}_{t-1}$ the average $\mathrm{R-squared}$ is 0.65 and 0.59 for $\mathrm{MOM}_{t-1} \leq \hat{\theta}$ and $\mathrm{MOM}_{t-1} > \hat{\theta}$, respectively (see Panel B in Figure 3) and diversification is less effective when momentum is low. Given the economic interpretation of the regimes $\mathrm{EPU}_{t-1} > \hat{\theta}$ and $\mathrm{MOM}_{t-1} \leq \hat{\theta}$, the results imply that diversification is weaker in bad times. These findings are confirmed when we look at the larger set of combined portfolios (see Figure 4): diversification becomes less effective when $\mathrm{EPU}_{t-1} > \hat{\theta}$ or $\mathrm{MOM}_{t-1} \leq \hat{\theta}$, as compared to $\mathrm{EPU}_{t-1} \leq \hat{\theta}$ and $\mathrm{MOM}_{t-1} > \hat{\theta}$, respectively.

6.4 Comparison with structural break model

If $z_t = t/T$ and $\theta \in (0, 1)$ then (1) becomes a multi-factor model with a single structural break: θ and $T_\theta = \lfloor \theta T \rfloor$ are the break fraction and the break date, respectively. We conjecture that our inferential procedure is applicable also in this case: Theorem 3.1 remains valid; Theorem 3.2 needs to be generalized since stationarity of the factors as imposed in Assumption D1(a) is unlikely to hold in this set up. Our test is informative about the global stability of the model; however, if the null hypothesis of stability is rejected, it cannot determine the exact number of breaks. Using a panel of equity returns, Smith and Timmermann (2020) find two and three breaks within the sample period we consider, depending on whether or not the breaks are cross-sectionally restricted, respectively. Under the maintained assumption of three breaks, we sequentially estimate their location by first estimating one break within the whole sample period and then one additional break within each resulting subsample: this is an application of the Bai and Perron (1998) procedure when the number of breaks is known.

Figure 5 about here

The estimated break dates are displayed in Figure 5. Using the set of industry portfolios we find breaks in September 1997, November 2004 and April 2012. The larger set of combined portfolios is likely to be more accurate at locating the breaks as it employs the information stemming from a higher cross-sectional dimension: the estimated break dates fall in June 1999, August 2003 and July 2010. These three breaks can be related to as many possible sources. The 1999 break may be due to the Gramm–Leach–Bliley Act of the same year, which repealed parts of the Glass-Steagall Act of 1933. The 2003 break occurred after the introduction of the Sarbanes-Oxley Act of 2002, which was enacted following major scandals such as Enron and WorldCom. The 2010 break may be caused by the Dodd-Frank Act of the same year, which followed the Great Recession. Structural breaks tend to capture major specific events such as regulatory changes; however, they miss higher frequency fluctuations that may be relevant to investors, such as those captured by EPU_{t-1} and MOM_{t-1}. Therefore, recurring regime changes and structural breaks should not be seen as mutually exclusive events, as they are likely to coexist.

7 Conclusions

We propose a Lagrange multiplier test for threshold-type regime changes in high dimensional factor models that is robust to regime-specific factor volatility. To reduce the dimensionality of the problem we work with an auxiliary threshold regression obtained by taking a weighted cross-sectional average of the returns: we obtain estimates for the factors from the large dimensional model constrained by the null hypothesis of linearity and we input them into the auxiliary regression; we then test for a threshold effect in the auxiliary regression. An application to a large set of equity portfolios illustrates the usefulness of our methodology for portfolio choice and risk measurement from an *ex-post* (or in-sample) perspective.

Our work can be extended along several dimensions: two are worth mentioning. To the very best of our knowledge, the literature on *ex-ante* (or out-of-sample) portfolio allocation under regime changes has studied low dimensional problems: see for example Guidolin and Timmermann (2007, 2008). In future work, we plan to implement our test within an out-of-sample asset allocation exercise in a high dimensional setting: this would extend the in-sample analysis conducted in this paper. Ours is a test for the null hypothesis of global linearity: however, should the null hypothesis be rejected, the test would not be informative about the actual number of recurrent regimes. The development of a recursive testing procedure to determine the number of regimes is at the top of our research agenda.

A Appendix

A.1 Proof of Proposition 3.1

Starting from (i), by Theorem 3.1 in Massacci (2017) the model in (1) is not identified from the linear model $\mathbf{R}_t = \mathbf{B}_1 \mathbf{f}_t + \mathbf{e}_t$ under $\mathcal{H}_0$. Assumptions A1 - A4 imply Assumptions A - D in Bai and Ng (2002), respectively, with respect to $\mathbf{B}_1^0$ and $\mathbf{f}_t^0$: the result in (i) follows from Theorem 2 in Bai and Ng (2002). As for (ii), Theorem 3.1 in Massacci (2017) implies that the model in (1) is identified from a linear factor model under $\mathcal{H}_1$. Consider

$$\begin{aligned}
\mathbf{R}_t &= \mathbb{I}_{1t}\left(\theta^0\right)\mathbf{B}_1^0\mathbf{f}_t^0 + \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{B}_2^0\mathbf{f}_t^0 + \mathbf{e}_t \\
&= \left[1 - \mathbb{I}_{2t}\left(\theta^0\right)\right]\mathbf{B}_1^0\mathbf{f}_t^0 + \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{B}_2^0\mathbf{f}_t^0 + \mathbf{e}_t \\
&= \mathbf{B}_1^0\mathbf{f}_t^0 + \mathbb{I}_{2t}\left(\theta^0\right)\left(\mathbf{B}_2^0 - \mathbf{B}_1^0\right)\mathbf{f}_t^0 + \mathbf{e}_t,
\end{aligned}$$

where $\left(\mathbf{B}_2^0 - \mathbf{B}_1^0\right)$ is a $N \times 2P^0$ matrix with column rank equal to $P_{\boldsymbol{\delta}}^0$ as $N \to \infty$: the proof of (ii) then follows along similar steps as those in Theorem 4.1 in Massacci (2017).

A.2 Proofs of Theorems 3.1, 3.2, 3.3 and 3.5

Lemma A.1 *Under Assumptions A3(a), A3(c) and B1,* $\bar{e}_{\mathbf{w}t} \xrightarrow{p} 0$.

Lemma A.2 *Under* $\mathcal{H}_0$ *and Assumptions A1-A4,* $\tilde{\mathbf{f}}_t \xrightarrow{p} \tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0$, *for* $t = 1, \ldots, T$, *as* $N, T \to \infty$.

Lemma A.3 *Under* $\mathcal{H}_0$ *and Assumptions A2 and B1,* $\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{2i}^0\right) \to \mathbf{L}' \left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)$ *as* $N \to \infty$.

Lemma A.4 *Under* $\mathcal{H}_0$ *and Assumptions A1-A4,*

(a) $C_{NT}^2 \left[\frac{1}{T}\sum_{t=1}^T \mathbb{I}_{jt}\left(\theta^0\right)\left(\tilde{\mathbf{f}}_t - \tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0\right)\bar{e}_{\mathbf{w}t}\right] = O_p(1)$, *for* $j = 1, 2$;

(b) $C_{NT}^2 \left[\frac{1}{T}\sum_{t=1}^T \mathbb{I}_{jt}\left(\theta^0\right)\tilde{\mathbf{f}}_t \left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{ji}^0\right)' \left(\mathbf{f}_t^0 - \tilde{\mathbf{H}}_1 \tilde{\mathbf{f}}_t\right)\right] = O_p(1)$, *for* $j = 1, 2$.

Proof of Theorem 3.1. Consider

$$
\begin{aligned}
\widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}}\left(\theta^0\right) &= \left[\widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)', \widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)'\right]' \\
&= \left[\sum_{t=1}^{T} \tilde{\mathbf{f}}_t\left(\theta^0\right) \tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1} \left[\sum_{t=1}^{T} \tilde{\mathbf{f}}_t\left(\theta^0\right) \bar{R}_{\mathbf{w}t}\right] \\
&= \left[\sum_{t=1}^{T} \tilde{\mathbf{f}}_t\left(\theta^0\right) \tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1} \left\{\sum_{t=1}^{T} \tilde{\mathbf{f}}_t\left(\theta^0\right) \left\{\sum_{i=1}^{N} w_i \left[\mathbb{I}_{1t}\left(\theta^0\right) \boldsymbol{\beta}_{1i}^{0\prime} \mathbf{f}_t^0 + \mathbb{I}_{2t}\left(\theta^0\right) \boldsymbol{\beta}_{2i}^{0\prime} \mathbf{f}_t^0 + e_{it}\right]\right\}\right\} \\
&= \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix}'\right\}^{-1} \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix} \left\{\sum_{i=1}^{N} w_i \left[\mathbb{I}_{1t}\left(\theta^0\right) \boldsymbol{\beta}_{1i}^{0\prime} \mathbf{f}_t^0\right]\right\}\right\} \\
&\quad + \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix}'\right\}^{-1} \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix} \left\{\sum_{i=1}^{N} w_i \left[\mathbb{I}_{2t}\left(\theta^0\right) \boldsymbol{\beta}_{2i}^{0\prime} \mathbf{f}_t^0\right]\right\}\right\} \\
&\quad + \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix}'\right\}^{-1} \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix} \left(\sum_{i=1}^{N} w_i e_{it}\right)\right\}, \\
&= \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t' & \mathbf{0}_{\tilde{P}} \\ \mathbf{0}_{\tilde{P}} & \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t' \end{bmatrix}\right\}^{-1} \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \mathbf{f}_t^{0\prime} \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0\right) \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \mathbf{f}_t^{0\prime} \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{2i}^0\right) \end{bmatrix}\right\} \\
&\quad + \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t' & \mathbf{0}_{\tilde{P}} \\ \mathbf{0}_{\tilde{P}} & \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t' \end{bmatrix}\right\}^{-1} \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \\ \mathbb{I}_{2t}\left(\theta^0\right) \tilde{\mathbf{f}}_t \end{bmatrix} \left(\sum_{i=1}^{N} w_i e_{it}\right)\right\}.
\end{aligned}
$$

By Lemma A.2, $\tilde{\mathbf{f}}_t \xrightarrow{p} \tilde{\mathbf{H}}_1^{-1} \mathbf{f}_t^0$. Assumption A1 ensures that $\left(\mathbf{F}^0 \mathbf{F}^{0\prime}/T\right) \xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{f}}^0$ as $T \to \infty$, where $\boldsymbol{\Sigma}_{\mathbf{f}}^0$ is a positive definite matrix. Following arguments analogous to those in Proposition 1 in Bai (2003), $\left(\mathbf{B}_1^{0\prime} \tilde{\mathbf{B}}_1^{\tilde{P}}/N\right) \xrightarrow{p} \mathbf{Q}_{\mathbf{B}_{11}}^0$, where $\mathbf{Q}_{\mathbf{B}_{11}}^0$ is an invertible matrix and it is unique by Assumption C1. By Lemma A.3 in Bai (2003), $\tilde{\mathbf{V}}_1 \xrightarrow{p} \mathbf{V}_1^0$, where $\mathbf{V}_1^0$ is a $P^0 \times P^0$ positive definite matrix. It follows that $\tilde{\mathbf{H}}_1 \xrightarrow{p} \mathbf{H}_1^0 = \boldsymbol{\Sigma}_{\mathbf{f}}^0 \mathbf{Q}_{\mathbf{B}_{11}}^0 \left(\mathbf{V}_1^0\right)^{-1}$, where $\mathbf{H}_1^0$ is a $P^0 \times P^0$ invertible matrix and it is unique by Assumption C1. This implies that $\tilde{\mathbf{f}}_t \xrightarrow{p} \left(\mathbf{H}_1^0\right)^{-1} \mathbf{f}_t^0$ and thus

$$
\sum_{t=1}^{T} \mathbb{I}_{jt}\left(\theta^0\right) \tilde{\mathbf{f}}_t \tilde{\mathbf{f}}_t' \xrightarrow{p} \left(\mathbf{H}_1^0\right)^{-1} \left[\sum_{t=1}^{T} \mathbb{I}_{jt}\left(\theta^0\right) \mathbf{f}_t^0 \mathbf{f}_t^{0\prime}\right] \left(\mathbf{H}_1^{0\prime}\right)^{-1}, \quad j = 1, 2.
$$

Taking into account Lemma A.1, it follows that

$$
\begin{aligned}
\widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}}\left(\theta^0\right) &= \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \left(\mathbf{H}_1^0\right)^{-1} \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \left(\mathbf{H}_1^{0\prime}\right)^{-1} & \mathbf{0}_{P^0} \\ \mathbf{0}_{P^0} & \mathbb{I}_{2t}\left(\theta^0\right) \left(\mathbf{H}_1^0\right)^{-1} \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \left(\mathbf{H}_1^{0\prime}\right)^{-1} \end{bmatrix}\right\}^{-1} \\
&\quad \times \left\{\sum_{t=1}^{T} \begin{bmatrix} \mathbb{I}_{1t}\left(\theta^0\right) \left(\mathbf{H}_1^0\right)^{-1} \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0\right) \\ \mathbb{I}_{1t}\left(\theta^0\right) \left(\mathbf{H}_1^0\right)^{-1} \mathbf{f}_t^0 \mathbf{f}_t^{0\prime} \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{2i}^0\right) \end{bmatrix}\right\} + o_p(1) \\
&= \begin{bmatrix} \mathbf{H}_1^{0\prime} \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0\right) \\ \mathbf{H}_1^{0\prime} \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{2i}^0\right) \end{bmatrix} + o_p(1).
\end{aligned}
$$

From Lemma A.3, we thus have

$$
\widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}}\left(\theta^0\right) = \left[\widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)', \widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)'\right]' \xrightarrow{p} \left\{\left[\mathbf{H}_1^{0\prime} \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0\right)\right]', \left[\mathbf{H}_1^{0\prime} \mathbf{L}' \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0\right)\right]'\right\}'.
$$

Adding and subtracting terms,

$$
\begin{aligned}
\bar{R}_{\mathbf{w}t} &= \mathbb{I}_{1t}\left(\theta^0\right)\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)' \mathbf{f}_t^0 + \mathbb{I}_{2t}\left(\theta^0\right)\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{2i}^0\right)' \mathbf{f}_t^0 + \bar{e}_{\mathbf{w}t} \\
&= \mathbb{I}_{1t}\left(\theta^0\right)\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)' \tilde{\mathbf{H}}_1 \tilde{\mathbf{f}}_t + \mathbb{I}_{2t}\left(\theta^0\right)\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{2i}^0\right)' \tilde{\mathbf{H}}_1 \tilde{\mathbf{f}}_t + \bar{e}_{\mathbf{w}t} \\
&\quad + \mathbb{I}_{1t}\left(\theta^0\right)\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)' \tilde{\mathbf{H}}_1 \left(\tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0 - \tilde{\mathbf{f}}_t\right) \\
&\quad + \mathbb{I}_{2t}\left(\theta^0\right)\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{2i}^0\right)' \tilde{\mathbf{H}}_1 \left(\tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0 - \tilde{\mathbf{f}}_t\right).
\end{aligned}
$$

The least squares estimator $\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)$ then is

$$
\begin{aligned}
\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right) &= \left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'\right]^{-1}\left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t \bar{R}_{\mathbf{w}t}\right] \\
&= \left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'\right]^{-1}\left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'\right]\tilde{\mathbf{H}}_1'\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right) \\
&\quad + \left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'\right]^{-1}\left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t \bar{e}_{\mathbf{w}t}\right] \\
&\quad + \left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'\right]^{-1}\left[\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)' \tilde{\mathbf{H}}_1\left(\tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0 - \tilde{\mathbf{f}}_t\right)\right]
\end{aligned}
$$

so that

$$
\begin{aligned}
& T^{1/2}\left[\widehat{\widetilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right) - \tilde{\mathbf{H}}_1'\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)\right] \\
= & \left[T^{-1}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'\right]^{-1}\left[T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t \bar{e}_{\mathbf{w}t}\right] \\
& + \left[T^{-1}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'\right]^{-1}\left[T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\sum_{i=1}^N w_i \boldsymbol{\beta}_{1i}^0\right)' \tilde{\mathbf{H}}_1\left(\tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0 - \tilde{\mathbf{f}}_t\right)\right].
\end{aligned}
$$

Notice that

$$
\begin{aligned}
T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t \bar{e}_{\mathbf{w}t} &= T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0 \bar{e}_{\mathbf{w}t} + T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\left(\tilde{\mathbf{f}}_t - \tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0\right)\bar{e}_{\mathbf{w}t} \\
&= \tilde{\mathbf{H}}_1^{-1}\left[T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0 \bar{e}_{\mathbf{w}t}\right] + o_p(1)
\end{aligned}
$$

by Lemma A.4(a), since

$$
\begin{aligned}
T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\left(\tilde{\mathbf{f}}_t - \tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0\right)\bar{e}_{\mathbf{w}t} &= T^{1/2}\frac{1}{T}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\left(\tilde{\mathbf{f}}_t - \tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0\right)\bar{e}_{\mathbf{w}t} \\
&= O_p\left(\frac{T^{1/2}}{\min\{N,T\}}\right) \\
&= o_p(1)
\end{aligned}
$$

when $\sqrt{T}/N \to 0$ as $N, T \to \infty$, as ensured by Assumption C2. Consider

$$
\begin{aligned}
T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0 \bar{e}_{\mathbf{w}t} &= T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0\left(\sum_{i=1}^N w_i e_{it}\right) \\
&= \sum_{i=1}^N w_i\left[T^{-1/2}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0 e_{it}\right].
\end{aligned}
$$

By Assumption C3, $T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}\xrightarrow{d}\mathcal{N}\left[\mathbf{0},\boldsymbol{\Omega}_{11}^0\left(\theta^0,\theta^0\right)\right]$, where $\boldsymbol{\Omega}_{11}^0\left(\theta^0,\theta^0\right)$ is positive definite. Assumption C2 and Lemma A.4(b) imply

$$\begin{aligned}T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\sum_{i=1}^{N}w_i\boldsymbol{\beta}_{1i}^0\right)'\tilde{\mathbf{H}}_1\left(\tilde{\mathbf{H}}_1^{-1}\mathbf{f}_t^0-\tilde{\mathbf{f}}_t\right) &= T^{1/2}\frac{1}{T}\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\sum_{i=1}^{N}w_i\boldsymbol{\beta}_{1i}^0\right)'\left(\mathbf{f}_t^0-\tilde{\mathbf{H}}_1\tilde{\mathbf{f}}_t\right)\\ &= O_p\left(\frac{T^{1/2}}{\min\{N,T\}}\right)\\ &= o_p(1).\end{aligned}$$

As $N,T\to\infty$, it thus follows that

$$T^{1/2}\left[\widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)-\tilde{\mathbf{H}}_1'\left(\sum_{i=1}^{N}w_i\boldsymbol{\beta}_{1i}^0\right)\right]\xrightarrow{d}\mathcal{N}\left[\mathbf{0},\mathbf{H}_1^{0\prime}\boldsymbol{\Sigma}_{\mathbf{f}11}^0\left(\theta^0,\theta^0\right)^{-1}\boldsymbol{\Omega}_{11}^0\left(\theta^0,\theta^0\right)\boldsymbol{\Sigma}_{\mathbf{f}11}^0\left(\theta^0,\theta^0\right)^{-1}\mathbf{H}_1^0\right].$$

Taking into account Lemma A.3, in a similar way it may be proved that

$$T^{1/2}\left[\widehat{\widehat{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)-\tilde{\mathbf{H}}_1'\mathbf{L}'\left(\sum_{i=1}^{N}w_i\boldsymbol{\beta}_{1i}^0\right)\right]\xrightarrow{d}\mathcal{N}\left[\mathbf{0},\mathbf{H}_1^{0\prime}\boldsymbol{\Sigma}_{\mathbf{f}22}^0\left(\theta^0,\theta^0\right)^{-1}\boldsymbol{\Omega}_{22}^0\left(\theta^0,\theta^0\right)\boldsymbol{\Sigma}_{\mathbf{f}22}^0\left(\theta^0,\theta^0\right)^{-1}\mathbf{H}_1^0\right]$$

as $N,T\to\infty$. Finally, notice that

$$\begin{aligned}&\hat{\mathbf{M}}\left(\theta^0,\theta^0\right)^{-1}\hat{\boldsymbol{\Omega}}\left(\theta^0,\theta^0\right)\hat{\mathbf{M}}\left(\theta^0,\theta^0\right)^{-1}\\ \xrightarrow{p}\ &\left\{\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{f}11}^0\left(\theta^0,\theta^0\right) & \mathbf{0}_{P^0}\\ \mathbf{0}_{P^0} & \boldsymbol{\Sigma}_{\mathbf{f}22}^0\left(\theta^0,\theta^0\right)\end{bmatrix}\left(\mathbf{I}_2\otimes\mathbf{H}_1^{0\prime}\right)^{-1}\right\}^{-1}\\ &\times\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\boldsymbol{\Omega}^0\left(\theta^0,\theta^0\right)\left(\mathbf{I}_2\otimes\mathbf{H}_1^{0\prime}\right)^{-1}\\ &\times\left\{\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{f}11}^0\left(\theta^0,\theta^0\right) & \mathbf{0}_{P^0}\\ \mathbf{0}_{P^0} & \boldsymbol{\Sigma}_{\mathbf{f}22}^0\left(\theta^0,\theta^0\right)\end{bmatrix}\left(\mathbf{I}_2\otimes\mathbf{H}_1^{0\prime}\right)^{-1}\right\}^{-1}\\ =\ &\left(\mathbf{I}_2\otimes\mathbf{H}_1^{0\prime}\right)\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{f}11}^0\left(\theta^0,\theta^0\right) & \mathbf{0}_{P^0}\\ \mathbf{0}_{P^0} & \boldsymbol{\Sigma}_{\mathbf{f}22}^0\left(\theta^0,\theta^0\right)\end{bmatrix}^{-1}\boldsymbol{\Omega}^0\left(\theta^0,\theta^0\right)\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{f}11}^0\left(\theta^0,\theta^0\right) & \mathbf{0}_{P^0}\\ \mathbf{0}_{P^0} & \boldsymbol{\Sigma}_{\mathbf{f}22}^0\left(\theta^0,\theta^0\right)\end{bmatrix}^{-1}\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right),\end{aligned}$$

which completes the proof of the theorem. ■

Proof of Theorem 3.2. The proof of the Theorem is analogous to the proof of Theorem 5.1 in Massacci (2017) and we sketch the main steps. Similarly to Lemma A.11 in Massacci (2017), for each θ

$$\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{\mathbf{f}}_t(\theta)\tilde{\bar{e}}_{\mathbf{w}t}-\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{bmatrix}\mathbb{I}_{1t}(\theta)\\ \mathbb{I}_{2t}(\theta)\end{bmatrix}\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}\right\|=o_p(1).$$

To prove that $\hat{\mathbf{k}}(\theta)\Rightarrow\mathbf{k}^0(\theta)$ we show that $\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}T^{-1/2}\sum_{t=1}^{T}\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}\Rightarrow\mathbf{k}^0(\theta)$: this follows if $\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}T^{-1/2}\sum_{t=1}^{T}\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}$ is stochastically equicontinuous. We use Application 4 of Theorem 1 in Doukhan *et al.* (1995). The summands $\mathbf{k}_t^0(\theta)=\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}$ satisfy the required

$\beta-$mixing decay rate under Assumption D1(a). By Assumption B1,

$$\begin{aligned}\sup_{\theta}\left\|\mathbf{k}_t^0(\theta)\right\| &= \sup_{\theta}\left\|\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}\right\| \\ &= \sup_{\theta}\left\|\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0\left(\sum_{i=1}^N w_i e_{it}\right)\right\| O(1) \\ &\leq \left\{\sum_{i=1}^N \sup_{\theta}\left\|\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0 e_{it}\right\|\right\} O\left(\frac{1}{N}\right) \\ &\leq \left\{\sum_{i=1}^N \sup_{\theta}\left\|\mathbb{I}_{1t}(\theta)\mathbf{f}_t^0 e_{it}\right\| + \sup_{\theta}\left\|\mathbb{I}_{2t}(\theta)\mathbf{f}_t^0 e_{it}\right\|\right\} O\left(\frac{1}{N}\right) \\ &\leq \left\{\sum_{i=1}^N \max_{j=1,2}\left[\sup_{\theta}\left\|\mathbb{I}_{jt}(\theta)\mathbf{f}_t^0 e_{it}\right\|\right]\right\} O\left(\frac{1}{N}\right):\end{aligned}$$

by Schwarz's inequality and Assumptions D1(b) and D1(c)

$$\mathrm{E}\left|\max_{j=1,2}\left[\sup_{\theta}\left\|\mathbb{I}_{jt}(\theta)\mathbf{f}_t^0 e_{it}\right\|\right]\right|^{2\xi} \leq \left\{\mathrm{E}\left|\max_{j=1,2}\left[\sup_{\theta}\left\|\mathbb{I}_{jt}(\theta)\mathbf{f}_t^0 e_{it}\right\|\right]\right|^{4\xi}\right\}^{1/2}\left(\mathrm{E}\left|e_{it}\right|^{4\xi}\right)^{1/2} < \infty,$$

and $\sup_{\theta}\left\|\mathbf{k}_t^0(\theta)\right\|$ is $\mathcal{L}_{2\xi}$ bounded. For some $G<\infty$ and for all θ, there is some $\bar{\theta}$ such that $\left|\theta-\bar{\theta}\right| \leq G\cdot\mathbb{N}(\zeta)^{-1}$. Set $\mathbb{N}(\zeta) = M^{1/\gamma}G\zeta^{-1/\gamma}$ and notice that

$$\begin{aligned}&\left[\mathrm{E}\left\|\mathbf{k}_t^0(\theta)-\mathbf{k}_t^0(\bar{\theta})\right\|^{2\xi}\right]^{1/(2\xi)} \\ =& \left\{\mathrm{E}\left\|\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0\bar{e}_{\mathbf{w}t} - \left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}\left[\mathbb{I}_{1t}(\bar{\theta}),\mathbb{I}_{2t}(\bar{\theta})\right]'\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}\right\|^{2\xi}\right\}^{1/(2\xi)} \\ =& \left\{\mathrm{E}\left\|\left\{\left[\mathbb{I}_{1t}(\theta)-\mathbb{I}_{1t}(\bar{\theta})\right]\mathbf{f}_t^0\bar{e}_{\mathbf{w}t},\left[\mathbb{I}_{2t}(\theta)-\mathbb{I}_{2t}(\bar{\theta})\right]\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}\right\}'\right\|^{2\xi}\right\}^{1/(2\xi)} O(1) \\ \leq& \left\{\sum_{i=1}^N\left\{\mathrm{E}\left\|\left\{\left[\mathbb{I}_{1t}(\theta)-\mathbb{I}_{1t}(\bar{\theta})\right]\mathbf{f}_t^0 e_{it},\left[\mathbb{I}_{2t}(\theta)-\mathbb{I}_{2t}(\bar{\theta})\right]\mathbf{f}_t^0 e_{it}\right\}'\right\|^{2\xi}\right\}^{1/(2\xi)}\right\} O\left(\frac{1}{N}\right) \\ \leq& \left\{\sum_{i=1}^N\left\{\mathrm{E}\left|\max_{j=1,2}\left\|\left[\mathbb{I}_{1t}(\theta)-\mathbb{I}_{1t}(\bar{\theta})\right]\mathbf{f}_t^0 e_{it}\right\|\right|^{2\xi}\right\}^{1/(2\xi)}\right\} O\left(\frac{1}{N}\right).\end{aligned}$$

By Assumption D2,

$$\left\{\mathrm{E}\left|\max_{j=1,2}\left\|\left[\mathbb{I}_{1t}(\theta)-\mathbb{I}_{1t}(\bar{\theta})\right]\mathbf{f}_t^0 e_{it}\right\|\right|^{2\xi}\right\}^{1/(2\xi)} \leq M\left|\theta-\bar{\theta}\right|^{\gamma} \leq M\cdot G^{\gamma}\cdot\mathbb{N}(\zeta)^{-\gamma} = \zeta,$$

and $\mathbb{N}(\zeta)$ satisfies the definition of bracketing numbers: the log of $\mathbb{N}(\zeta)$ can be shown to be integrable as in the proof of Theorem 1 in Hansen (1996). It follows that $\left(\mathbf{I}_2\otimes\mathbf{H}_1^0\right)^{-1}T^{-1/2}\sum_{t=1}^T\left[\mathbb{I}_{1t}(\theta),\mathbb{I}_{2t}(\theta)\right]'\mathbf{f}_t^0\bar{e}_{\mathbf{w}t}$ is stochastically equicontinuous and then $\hat{\mathbf{k}}^{\mathbb{H}_0}(\theta) \Rightarrow \mathbf{k}^0(\theta)$. Assumption D3 and similar steps as in the proof of Theorem 1 in Hansen (1996) complete the proof of the theorem. ■

Proof of Theorem 3.3. Starting from (a), given Assumptions A1-A4 and under $\mathcal{H}_1$, $\lim_{N,T\to\infty}\Pr\left(\tilde{P}=P^0+P_{\boldsymbol{\delta}}^0\right)=1$ by Proposition 3.1. By Assumption A1, Proposition 1 in Bai (2003), and Lemma A.3 in Bai (2003), $\tilde{\mathbf{H}}_{\boldsymbol{\delta}}\left(\theta^0\right)\xrightarrow{p}\mathbf{H}_{\boldsymbol{\delta}}^0\left(\theta^0\right)$, where $\mathbf{H}_{\boldsymbol{\delta}}^0\left(\theta^0\right)$ is a $\left(P^0+P_{\boldsymbol{\delta}}^0\right)\times\left(P^0+P_{\boldsymbol{\delta}}^0\right)$ invertible matrix and it is unique by Assumption E1. By arguments similar to those in the proof of Proposition 1 in Chen *et al.* (2014), it follows that

$$\tilde{\mathbf{f}}_t = \tilde{\mathbf{H}}_{\boldsymbol{\delta}}\left(\theta^0\right)^{-1}\mathbf{f}_{\boldsymbol{\delta}t}^0\left(\theta^0\right) + O_p\left(C_{NT}^{-1}\right),$$

which implies that

$$\tilde{\mathbf{f}}_t\left(\theta^0\right)=\left[\begin{array}{c}\mathbb{I}_{1t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\\ \mathbb{I}_{2t}\left(\theta^0\right)\tilde{\mathbf{f}}_t\end{array}\right]=\left[\mathbf{I}_2\otimes\tilde{\mathbf{H}}_{\boldsymbol{\delta}}\left(\theta^0\right)\right]^{-1}\left[\begin{array}{c}\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\\ \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\end{array}\right]+O_p\left(C_{NT}^{-1}\right).$$

From (9), write $R_{it}=\boldsymbol{\beta}^{0\prime}_{1i}\mathbf{f}^0_t+\boldsymbol{\delta}^{0\prime}_{\boldsymbol{\delta}2i}\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)+e_{it}$, for $i=1,\ldots,N$ and $t=1,\ldots,T$; we thus have

$$\begin{aligned}
\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}\left(\theta^0\right) &= \left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1}\left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\bar{R}_{\mathbf{w}t}\right]\\
&= \left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1}\left\{\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\left\{\sum_{i=1}^N w_i\left[\boldsymbol{\beta}^{0\prime}_{1i}\mathbf{f}^0_t+\boldsymbol{\delta}^{0\prime}_{\boldsymbol{\delta}2i}\mathbf{f}^0_{\boldsymbol{\delta}2t}\left(\theta^0\right)+e_{it}\right]\right\}\right\}\\
&= \left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1}\\
&\quad\times\left\{\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\left\{\sum_{i=1}^N w_i\left\{\mathbb{I}_{1t}\left(\theta^0\right)\left[\mathbf{f}^{0\prime}_t\boldsymbol{\beta}^0_{1i}+\mathbf{f}^0_{\boldsymbol{\delta}2t}\left(\theta^0\right)'\boldsymbol{\delta}^0_{\boldsymbol{\delta}2i}\right]+\mathbb{I}_{2t}\left(\theta^0\right)\left[\mathbf{f}^{0\prime}_t\boldsymbol{\beta}^0_{1i}+\mathbf{f}^0_{\boldsymbol{\delta}2t}\left(\theta^0\right)'\boldsymbol{\delta}^0_{\boldsymbol{\delta}2i}\right]\right\}\right\}\right\}\\
&\quad+\left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1}\left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\left(\sum_{i=1}^N w_ie_{it}\right)\right]\\
&= \left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1}\left\{\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\left[\begin{array}{c}\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\\ \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\end{array}\right]'\right\}\left[\sum_{i=1}^N w_i\left(\begin{array}{c}\boldsymbol{\beta}^0_{1i}\\ \mathbf{0}\end{array}\right)',\sum_{i=1}^N w_i\left(\begin{array}{c}\boldsymbol{\beta}^0_{1i}\\ \boldsymbol{\delta}^0_{\boldsymbol{\delta}2i}\end{array}\right)'\right]'\\
&\quad+\left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\tilde{\mathbf{f}}_t\left(\theta^0\right)'\right]^{-1}\left[\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta^0\right)\left(\sum_{i=1}^N w_ie_{it}\right)\right].
\end{aligned}$$

Taking into account Lemma A.1,

$$\begin{aligned}
\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}\left(\theta^0\right) &= \left[\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)',\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)'\right]'\\
&\xrightarrow{p} \left\{\sum_{t=1}^T\left[\mathbf{I}_2\otimes\mathbf{H}^0_{\boldsymbol{\delta}}\left(\theta^0\right)\right]^{-1}\left[\begin{array}{c}\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\\ \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\end{array}\right]\left[\begin{array}{c}\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\\ \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\end{array}\right]'\left[\mathbf{I}_2\otimes\mathbf{H}^0_{\boldsymbol{\delta}}\left(\theta^0\right)'\right]^{-1}\right\}^{-1}\\
&\quad\times\left\{\sum_{t=1}^T\left[\mathbf{I}_2\otimes\mathbf{H}^0_{\boldsymbol{\delta}}\left(\theta^0\right)\right]^{-1}\left[\begin{array}{c}\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\\ \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\end{array}\right]\left[\begin{array}{c}\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\\ \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}^0_{\boldsymbol{\delta}t}\left(\theta^0\right)\end{array}\right]'\right\}\left[\sum_{i=1}^N w_i\left(\begin{array}{c}\boldsymbol{\beta}^0_{1i}\\ \mathbf{0}\end{array}\right)',\sum_{i=1}^N w_i\left(\begin{array}{c}\boldsymbol{\beta}^0_{1i}\\ \boldsymbol{\delta}^0_{\boldsymbol{\delta}2i}\end{array}\right)'\right]'\\
&= \left[\mathbf{I}_2\otimes\mathbf{H}^0_{\boldsymbol{\delta}}\left(\theta^0\right)'\right]\left[\sum_{i=1}^N w_i\left(\begin{array}{c}\boldsymbol{\beta}^0_{1i}\\ \mathbf{0}\end{array}\right)',\sum_{i=1}^N w_i\left(\begin{array}{c}\boldsymbol{\beta}^0_{1i}\\ \boldsymbol{\delta}^0_{\boldsymbol{\delta}2i}\end{array}\right)'\right]'.
\end{aligned}$$

Under Assumption B2, it follows that

$$\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)-\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)\xrightarrow{p}\mathbf{H}^0_{\boldsymbol{\delta}}\left(\theta^0\right)'\sum_{i=1}^N w_i\left(\begin{array}{c}\mathbf{0}\\ \boldsymbol{\delta}^0_{\boldsymbol{\delta}2i}\end{array}\right)=\mathbf{c}\neq\mathbf{0}.$$

As for (b), from (6) we have

$$\begin{aligned}
\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}\left(\theta^0\right) &= T\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}\left(\theta^0\right)'\mathbf{G}\left[\mathbf{G}'\hat{\mathbf{M}}\left(\theta^0,\theta^0\right)^{-1}\hat{\boldsymbol{\Omega}}\left(\theta^0,\theta^0\right)\hat{\mathbf{M}}\left(\theta^0,\theta^0\right)^{-1}\mathbf{G}\right]^{-1}\mathbf{G}'\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}}\left(\theta^0\right)\\
&= T\left[\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)-\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)\right]'\left[\mathbf{G}'\hat{\mathbf{M}}\left(\theta^0,\theta^0\right)^{-1}\hat{\boldsymbol{\Omega}}\left(\theta^0,\theta^0\right)\hat{\mathbf{M}}\left(\theta^0,\theta^0\right)^{-1}\mathbf{G}\right]^{-1}\left[\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)-\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)\right].
\end{aligned}$$

As $N,T\to\infty$: $\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}1}\left(\theta^0\right)-\widehat{\tilde{\boldsymbol{\beta}}}_{\mathbf{w}2}\left(\theta^0\right)\xrightarrow{p}\mathbf{c}\neq\mathbf{0}$ from (a); $\hat{\mathbf{M}}\left(\theta^0,\theta^0\right)\xrightarrow{p}\mathbf{M}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}\left(\theta^0,\theta^0\right)$ and $\hat{\boldsymbol{\Omega}}\left(\theta^0,\theta^0\right)\xrightarrow{p}\boldsymbol{\Omega}^0_{\mathbf{H}^0_{\boldsymbol{\delta}}}\left(\theta^0,\theta^0\right)$ by Assumption E2. We thus have that $\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}\left(\theta^0\right)\xrightarrow{p}\infty$ as $N,T\to\infty$. The consistency of $\sup\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}$ follows from the consistency of $\widehat{LM}^{\mathrm{HAC}}_{\mathbf{w}}\left(\theta^0\right)$. ■

Proof of Theorem 3.4. Following the proof of Theorem 2 in Corradi and Swanson (2014), to prove (a) and (b) it is sufficient to notice that $\mathrm{E}^*\left[\tilde{\mathbf{f}}_t\left(\theta\right)\widetilde{\bar{e}}_{\mathbf{w}t}u^*_{bt}\right]=T^{-1}\sum_{t=1}^T\tilde{\mathbf{f}}_t\left(\theta\right)\widetilde{\bar{e}}_{\mathbf{w}t}+O_p\left(T^{-1/2}\right)$, where E^* is the expected value calculated with respect to the bootstrap probability measure conditional upon the original data. ■

Proof of Theorem 3.5. Consider the loss function in (4). From the proof of Lemma A.9 in Massacci (2017), $S\left(\tilde{\mathbf{B}}_1^P, \tilde{\mathbf{F}}^P\right) - S\left(\tilde{\mathbf{B}}_1, \tilde{\mathbf{F}}\right) = O_p\left(C_{NT}^{-2}\right)$, for $P \geq P^0$, where $\tilde{\mathbf{B}}_1 = \tilde{\mathbf{B}}_1^{\tilde{P}}$ and $\tilde{\mathbf{F}} = \tilde{\mathbf{F}}^{\tilde{P}}$. Using arguments similar to those used in the proof of Theorem 3.1 to prove that $\tilde{\mathbf{H}}_1 \xrightarrow{p} \mathbf{H}_1^0$, it follows that $\tilde{\mathbf{H}}_1^P \xrightarrow{p} \mathbf{H}_1^{0,P}$, where $\mathbf{H}_1^{0,P}$ is a $P^0 \times P$ unique matrix such that $\mathbf{H}_1^{0,P^0} = \mathbf{H}_1^0$, where $\mathbf{H}_1^0$ is defined in Theorem 3.1: the generalized inverse $\tilde{\mathbf{H}}_1^{P+}$ of $\tilde{\mathbf{H}}_1^P$ is such that $\tilde{\mathbf{H}}_1^P \cdot \tilde{\mathbf{H}}_1^{P+} = \mathbf{I}_P$. The proof is completed by suitably generalizing the proof of Theorem 3.1. ■

Proof of Lemma A.1. Under Assumptions A3(a) and B1

$$\mathrm{E}\left(\bar{e}_{\mathbf{w}t}\right) = \mathrm{E}\left(\sum_{i=1}^N w_i e_{it}\right) = O\left(\frac{1}{N}\right)\sum_{i=1}^N \mathrm{E}\left(e_{it}\right) = 0.$$

By Assumptions A3(c) and B1,

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^T \mathrm{Var}\left(\bar{e}_{\mathbf{w}t}\right) &= \frac{1}{T}\sum_{t=1}^T \mathrm{E}\left[\left(\sum_{i=1}^N w_i e_{it}\right)^2\right] \\
&= \frac{1}{T}\sum_{t=1}^T \mathrm{E}\left[\left(\sum_{i=1}^N w_i e_{it}\right)\left(\sum_{l=1}^N w_l e_{lt}\right)\right] \\
&= O\left(\frac{1}{N^2}\right)\sum_{i=1}^N\sum_{l=1}^N \mathrm{E}\left(\frac{1}{T}\sum_{t=1}^T e_{it}e_{lt}\right) \\
&= O\left(\frac{1}{N^2}\right)\sum_{i=1}^N\sum_{l=1}^N \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^T \mathbb{I}_{1t}\left(\theta^0\right) e_{it}e_{lt}\right] + O\left(\frac{1}{N^2}\right)\sum_{i=1}^N\sum_{l=1}^N \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^T \mathbb{I}_{2t}\left(\theta^0\right) e_{it}e_{lt}\right] \\
&\leq O\left(\frac{1}{N}\right)\left[\frac{1}{N}\sum_{i=1}^N\sum_{l=1}^N \left|\sigma_{1il}\left(\theta^0\right)\right| + \frac{1}{N}\sum_{i=1}^N\sum_{l=1}^N \left|\sigma_{2il}\left(\theta^0\right)\right|\right] \\
&= o\left(1\right),
\end{aligned}$$

which implies that $\mathrm{Var}\left(\bar{e}_{\mathbf{w}t}\right) = o\left(1\right)$. Since $\mathrm{E}\left(\bar{e}_{\mathbf{w}t}\right) = 0$ and $\mathrm{Var}\left(\bar{e}_{\mathbf{w}t}\right) \to 0$ as $N \to \infty$, then $\bar{e}_{\mathbf{w}t} \xrightarrow{p} 0$ as $N \to \infty$: this completes the proof of the lemma. ■

Proof of Lemma A.2. The result follows from Theorem 3.1 in Massacci (2017) and the proof is omitted. ■

Proof of Lemma A.3. Under $\mathcal{H}_0$ and Assumption B1,

$$\sum_{i=1}^N w_i \boldsymbol{\beta}_{2i}^0 = \sum_{i=1}^{\lfloor N^{0.5}\rfloor} w_i \boldsymbol{\beta}_{2i}^0 + \sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i \boldsymbol{\beta}_{2i}^0 = \sum_{i=1}^{\lfloor N^{0.5}\rfloor} w_i \boldsymbol{\beta}_{2i}^0 + \mathbf{L}' \sum_{i=\lfloor N^{0.5}\rfloor+1}^N w_i \boldsymbol{\beta}_{1i}^0.$$

Assumptions A2 and B1 imply that

$$\begin{aligned}
\left\|\sum_{i=1}^{\lfloor N^{0.5}\rfloor} w_i \boldsymbol{\beta}_{2i}^0\right\| &\leq O\left(\frac{1}{N}\right)\left\|\sum_{i=1}^{\lfloor N^{0.5}\rfloor} \boldsymbol{\beta}_{2i}^0\right\| \\
&\leq O\left(\frac{1}{N}\right)\sum_{i=1}^{\lfloor N^{0.5}\rfloor} \left\|\boldsymbol{\beta}_{2i}^0\right\| \\
&\leq O\left(\frac{1}{N}\right)\sum_{i=1}^{\lfloor N^{0.5}\rfloor} \bar{\beta} \\
&\leq \bar{\lambda} O\left(\frac{\lfloor N^{0.5}\rfloor}{N}\right) \\
&= o\left(1\right).
\end{aligned}$$

Assumptions A2 and B1 also imply that

$$\begin{aligned}
\mathbf{L}' \sum_{i=\lfloor N^{0.5} \rfloor+1}^{N} w_i \boldsymbol{\beta}_{1i}^0 &= \mathbf{L}' \sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0 - \mathbf{L}' \sum_{i=1}^{\lfloor N^{0.5} \rfloor} w_i \boldsymbol{\beta}_{1i}^0 \\
&= \mathbf{L}' \sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0 - \mathbf{L}' O\left(\frac{1}{N}\right) \left(\sum_{i=1}^{\lfloor N^{0.5} \rfloor} \boldsymbol{\beta}_{1i}^0\right) \\
&= \mathbf{L}' \sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0 + o(1).
\end{aligned}$$

It thus follows that $\left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{2i}^0\right) \to \mathbf{L}' \left(\sum_{i=1}^{N} w_i \boldsymbol{\beta}_{1i}^0\right)$ as $N \to \infty$, which completes the proof of the lemma. ■

Proof of Lemma A.4. The proofs of Lemmas A.4(a) and A.4(b) are similar to the proofs of Lemmas A.1(iv) and A.1(iii) in Bai and Ng (2006), respectively, and they are omitted. ■

References

[1] Ang, A., and G. Bekaert (2002), "International Asset Allocation with Regime Shifts," *Review of Financial Studies*, 15 (4), 1137 – 1187.

[2] Ang, A., and A. Timmermann (2012), "Regime Changes and Financial Markets," *Annual Review of Financial Economics*, 4, 313 – 337.

[3] Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013), "Value and Momentum Everywhere," *Journal of Finance*, 68 (3), 929 – 985.

[4] Baele, L., G. Bekaert, and K. Inghelbrecht (2010), "The Determinants of Stock and Bond Return Comovements," *Review of Financial Studies*, 23 (6), 2374 – 2428.

[5] Bai, J. (2003), "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71 (1), 135 – 171.

[6] Bai, J. and K. Li (2012), "Statistical Analysis of Factor Models of High Dimension," *Annals of Statistics*, 40 (1), 436 – 465.

[7] Bai, J., and S. Ng (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70 (1), 191 – 221.

[8] Bai, J., and S. Ng (2006), "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions," *Econometrica*, 74 (4), 1133 – 1150.

[9] Bai, J., and P. Perron (1998), "Estimation and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66 (1), 47 – 78.

[10] Baker, S. R., N. Bloom, and S. J. Davis (2016), "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, 131 (4), 1593 – 1636.

[11] Baltagi, B. H., C. Kao, and F. Wang (2017), "Identification and Estimation of a Large Factor Model with Structural Instability," *Journal of Econometrics*, 197 (1), 87 – 100.

[12] Baltagi, B. H., C. Kao, and F. Wang (2020), "Estimating and Testing High Dimensional Factor Models with Multiple Structural Changes," *Journal of Econometrics*, available at https://doi.org/10.1016/j.jeconom.2020.04.005.

[13] Barigozzi, M., H. Cho, and P. Fryzlewicz (2018), "Simultaneous Multiple Change-Point and Factor Analysis for High-Dimensional Time Series," *Journal of Econometrics*, 206 (1), 187 – 225.

[14] Barigozzi, M., M. Hallin, and S. Soccorsi (2019), "Identification of Global and Local Shocks in International Financial Markets via General Dynamic Factor Models," *Journal of Financial Econometrics*, 17 (3), 462 – 494.

[15] Barigozzi, M., and L. Trapani (2020), "Sequential Testing for Structural Stability in Approximate Factor Models," *Stochastic Processes and their Applications*, 130 (8), 5149–5187.

[16] Bates, B. J., M. Plagborg-Møller, J. H. Stock, and M. W. Watson (2013), "Consistent Factor Estimation in Dynamic Factor Models with Structural Instability," *Journal of Econometrics*, 177 (2), 289 – 304.

[17] Breitung, J., and S. Eickmeier (2011), "Testing for Structural Breaks in Dynamic Factor Models," *Journal of Econometrics*, 163 (1), 71 – 84.

[18] Brogaard, J., L. Dai, P. T. H. Ngo, and B. Zhang (2020), "Global Political Uncertainty and Asset Prices," *Review of Financial Studies*, 33 (4), 1737 – 1780.

[19] Cai, Y., and J. Stander (2020), "The Threshold GARCH Model: Estimation and Density Forecasting for Financial Returns," *Journal of Financial Econometrics*, 18 (2), 395 – 424.

[20] Chamberlain, G., and M. Rothschild (1983), "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, 51 (5), 1281 – 1304.

[21] Chan, K. S., B. Hansen, and A. Timmermann (2017), "Guest Editors' Introduction: Regime Switching and Threshold Models," *Journal of Business and Economic Statistics*, 35 (2), 159 – 161.

[22] Chen, L., J. J. Dolado, and J. Gonzalo (2014), "Detecting Big Structural Breaks in Large Factor Models," *Journal of Econometrics*, 180 (1), 30 – 48.

[23] Cheng, X., L. Zhipeng, and F. Schorfheide (2016), "Shrinkage Estimation of High-Dimensional Factor Models with Structural Instabilities," *Review of Economic Studies*, 83 (4), 1511 – 1543.

[24] Connor, G., and R. A. Korajczyk (1986), "Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis," *Journal of Financial Economics*, 15 (3), 373 – 394.

[25] Connor, G., and R. A. Korajczyk (1988), "Risk and Return in an Equilibrium APT: Application of a New Test Methodology," *Journal of Financial Economics*, 21 (2), 255 – 289.

[26] Connor, G., and R. A. Korajczyk (1993), "A Test for the Number of Factors in an Approximate Factor Model," *Journal of Finance*, 48 (4), 1263 – 1288.

[27] Corradi, V., and N. R. Swanson (2014), "Testing for Structural Stability of Factor Augmented Forecasting Models," *Journal of Econometrics*, 182 (1), 100 – 118.

[28] Diebold, F. X., and Rudebusch, G. D. (1996), "Measuring Business Cycles: A Modern Perspective," *Review of Economics and Statistics*, 78 (1), 67 – 77.

[29] Doukhan, P., P. Massart, and E. Rio (1995), "Invariance Principles for Absolutely Regular Empirical Processes," *Annales de l.I. H. P., Section B*, 31 (2), 393 – 427.

[30] Fama, E. F., and K. R. French (1993), "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33 (1), 3 – 56.

[31] Forbes, K. J., and R. Rigobon (2002), "No Contagion, Only Interdependence: Measuring Stock Market Comovements," *Journal of Finance*, 57 (5), 2223 – 2261.

[32] Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2017), "Dynamic Factor Models with Infinite-Dimensional Factor Spaces: Asymptotic Analysis," *Journal of Econometrics*, 199 (1), 74 – 92.

[33] Giglio, S., and D. Xiu (2019), "Asset Pricing with Omitted Factors", Working Paper, Yale School of Management and Booth School of Business.

[34] Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993), "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48 (5), 1779 – 1801.

[35] Guidolin, M., and M. Pedio (2019), "Forecasting and Trading Monetary Policy Effects on the Riskless Yield Curve with Regime Switching Nelson–Siegel Models," *Journal of Economic Dynamics and Control*, 107, available at https://doi.org/10.1016/j.jedc.2019.103723.

[36] Guidolin, M., and A. Timmermann (2007), "Asset Allocation under Multivariate Regime Switching," *Journal of Economic Dynamics and Control*, 31 (11), 3503 – 3544.

[37] Guidolin, M., and A. Timmermann (2008), "International Asset Allocation under Regime Switching, Skew, and Kurtosis Preferences," *Review of Financial Studies*, 21 (2), 889 – 935.

[38] Han, X., and A. Inoue (2015), "Tests for Parameter Instability in Dynamic Factor Models," *Econometric Theory*, 31 (5), 1117 – 1152.

[39] Hansen, B. E. (1996), "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis", *Econometrica*, 64 (2), 413 – 430.

[40] Henkel, S. J., J. S. Martin, and F. Nardari (2011), "Time-Varying Short-Horizon Predictability," *Journal of Financial Economics*, 99 (3), 560 – 580.

[41] Kim, C.-J. (1994), "Dynamic linear models with Markov-switching," *Journal of Econometrics*, 60 (1 – 2), 1 – 22.

[42] Kim, C.-J., and C. R. Nelson (1998), "Business Cycle Turning Points, a New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching," *Review of Economics and Statistics*, 80 (2), 188 – 201.

[43] Lehmann, N. B., and D. M. Modest (2005), "Diversification and the Optimal Construction of Basis Portfolios," *Management Science*, 51 (4), 581 – 598.

[44] Liu, X., and E. Y. Chen (2019), "Helping Effects Against Curse of Dimensionality in Threshold Factor Models for Matrix Time Series," Working Paper, San Diego State University and Princeton University.

[45] Liu, X., and R. Chen (2020), "Threshold Factor Models for High-Dimensional Time Series," *Journal of Econometrics*, 216 (1), 53 – 70.

[46] Lustig, H., N. Roussanov, and A. Verdelhan (2011), "Common Risk Factors in Currency Markets," *Review of Financial Studies*, 24 (11), 3731 – 3777.

[47] Ma, S., and L. Su (2018), "Estimation of Large Dimensional Factor Models with an Unknown Number of Breaks," *Journal of Econometrics*, 207 (1), 1 – 29.

[48] Massacci, D. (2017), "Least Squares Estimation of Large Dimensional Threshold Factor Models," *Journal of Econometrics*, 197 (1), 101 – 129.

[49] Massacci, D. (2019), "Unstable Diffusion Indexes: With an Application to Bond Risk Premia," *Oxford Bulletin of Economics and Statistics*, 81 (6), 1376 – 1400.

[50] Newey, W. K., and K. D. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autororrelation Consistent Covariance Matrix," *Econometrica*, 55 (3), 703 – 708.

[51] Pearson, K. (1900), "On the Correlation of Characters not Quantitatively Measurable," *Royal Society Philosophical Transactions*, Series A, 195, 1 – 47.

[52] Pesaran, M. H. (2006), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74 (4), 967 – 1012.

[53] Pukthuanthong, K., and R. Roll (2009), "Global Market Integration: An Alternative Measure and Its Application," *Journal of Financial Economics*, 94 (2), 214 – 232.

[54] Smith, S. C., and A. Timmermann (2020), "Break Risk," *Review of Financial Studies*, Forthcoming.

[55] Stock, J. H., and M. W. Watson (2002), "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97 (460), 1167 – 1179.

[56] Su, L., and X. Wang (2017), "On Time-Varying Factor Models: Estimation and Testing," *Journal of Econometrics*, 198 (1), 84 – 101.

[57] Trapani, L. (2018), "A Randomised Sequential Procedure to Determine the Number of Factors," *Journal of the American Statistical Association*, 113 (523), 1341 – 1349.

[58] Yamamoto, Y., and S. Tanaka (2015), "Testing for Factor Loading Structural Change under Common Breaks," *Journal of Econometrics*, 189 (1), 187 − 206.

Table 1: Experiment 1

Panel A: $\widehat{LM}_{\mathbf{f}}\left(\theta^0\right)$ Statistic									
N		50				100			
$\delta_{\mathbf{f}}$		0.00	0.25	1.00	1.75	0.00	0.25	1.00	1.75
T	π^0								
100	0.15	0.0485	0.0285	0.0030	0.0005	0.0530	0.0280	0.0065	0.0010
	0.30	0.0535	0.0360	0.0135	0.0045	0.0450	0.0305	0.0090	0.0015
	0.50	0.0475	0.0475	0.0305	0.0140	0.0555	0.0520	0.0300	0.0115
	0.70	0.0540	0.0655	0.0955	0.0820	0.0465	0.0640	0.0830	0.0690
	0.85	0.0540	0.0865	0.1750	0.2100	0.0485	0.0800	0.1515	0.1790
200	0.15	0.0545	0.0290	0.0020	0.0000	0.0545	0.0335	0.0015	0.0000
	0.30	0.0500	0.0395	0.0085	0.0015	0.0505	0.0390	0.0080	0.0005
	0.50	0.0490	0.0465	0.0350	0.0180	0.0515	0.0530	0.0225	0.0105
	0.70	0.0635	0.0735	0.1025	0.0945	0.0555	0.0680	0.0850	0.0665
	0.85	0.0485	0.0850	0.1815	0.2240	0.0480	0.0810	0.1495	0.1775
400	0.15	0.0490	0.0275	0.0040	0.0000	0.0490	0.0260	0.0035	0.0000
	0.30	0.0570	0.0365	0.0105	0.0015	0.0505	0.0355	0.0085	0.0005
	0.50	0.0510	0.0525	0.0435	0.0280	0.0585	0.0560	0.0330	0.0110
	0.70	0.0560	0.0770	0.1170	0.1205	0.0485	0.0610	0.0765	0.0670
	0.85	0.0460	0.0790	0.1760	0.2385	0.0515	0.0760	0.1535	0.1985
Panel B: $\widehat{LM}_{\mathbf{w}}\left(\theta^0\right)$ Statistic									
N		50				100			
$\delta_{\mathbf{f}}$		0.00	0.25	1.00	1.75	0.00	0.25	1.00	1.75
T	π^0								
100	0.15	0.0490	0.0475	0.0465	0.0405	0.0540	0.0545	0.0520	0.0500
	0.30	0.0545	0.0545	0.0545	0.0500	0.0515	0.0505	0.0530	0.0500
	0.50	0.0445	0.0470	0.0500	0.0490	0.0485	0.0460	0.0480	0.0485
	0.70	0.0540	0.0525	0.0570	0.0545	0.0570	0.0520	0.0525	0.0530
	0.85	0.0505	0.0475	0.0490	0.0450	0.0565	0.0535	0.0525	0.0500
200	0.15	0.0495	0.0520	0.0555	0.0520	0.0490	0.0500	0.0480	0.0480
	0.30	0.0490	0.0495	0.0525	0.0510	0.0460	0.0455	0.0475	0.0475
	0.50	0.0485	0.0475	0.0425	0.0500	0.0470	0.0455	0.0440	0.0460
	0.70	0.0550	0.0550	0.0540	0.0575	0.0495	0.0480	0.0495	0.0505
	0.85	0.0500	0.0495	0.0520	0.0540	0.0460	0.0460	0.0495	0.0490
400	0.15	0.0460	0.0460	0.0450	0.0435	0.0480	0.0475	0.0490	0.0480
	0.30	0.0445	0.0440	0.0450	0.0420	0.0545	0.0555	0.0550	0.0580
	0.50	0.0505	0.0505	0.0470	0.0440	0.0535	0.0530	0.0510	0.0555
	0.70	0.0485	0.0485	0.0480	0.0520	0.0585	0.0555	0.0500	0.0505
	0.85	0.0595	0.0565	0.0530	0.0535	0.0475	0.0460	0.0445	0.0500

This table presents size results for the $\widehat{LM}_{\mathbf{f}}\left(\theta^0\right)$ statistic of Massacci (2017) and the $\widehat{LM}_{\mathbf{w}}\left(\theta\right)$ statistic as defined in (8) for $\theta = \theta^0$. The DGP is detailed in Section 5.1.1. The size is computed over $S = 2000$ replications according to (10), with $\hat{\mathcal{T}}^s = \widehat{LM}^s_{\mathbf{f}}\left(\theta^0\right), \widehat{LM}^s_{\mathbf{w}}\left(\theta^0\right)$.

Table 2: Experiment 2

Panel A: $\widehat{LM}_{\mathbf{w}}^{\mathrm{HAC}}\left(\theta^0\right)$ Statistic

N			50									100							
			Size			**Power**						**Size**			**Power**				
			$\delta_{\mathbf{f}}$	0.00	1.75	$\delta_{\mathbf{f}}$	0.00		1.75			$\delta_{\mathbf{f}}$	0.00	1.75	$\delta_{\mathbf{f}}$	0.00		1.75	
			δ^0	0.00	0.00	δ^0	0.25	1.00	0.25	1.00		δ^0	0.00	0.00	δ^0	0.25	1.00	0.25	1.00
T	**DGP**	π^0																	
		0.15		0.0235	0.0220		0.5860	0.6855	0.2185	0.5235			0.0220	0.0190		0.6825	0.6975	0.3570	0.5565
100	$\mathrm{CSI}_{\mathbf{e}}$	0.30		0.0405	0.0345		0.9800	0.9955	0.7780	0.9760			0.0420	0.0375		0.9970	0.9975	0.9210	0.9835
		0.50		0.0475	0.0410		1.0000	1.0000	0.9885	1.0000			0.0425	0.0390		1.0000	1.0000	0.9995	1.0000
		0.15		0.0220	0.0225		0.4670	0.6805	0.1255	0.4705			0.0175	0.0175		0.6160	0.6815	0.2410	0.5230
100	$\mathrm{CSD}_{\mathbf{e}}$	0.30		0.0360	0.0325		0.9440	0.9955	0.5220	0.9660			0.0395	0.0335		0.9905	0.9970	0.7915	0.9765
		0.50		0.0440	0.0410		0.9955	1.0000	0.9050	1.0000			0.0450	0.0400		0.9995	1.0000	0.9925	1.0000
		0.15		0.0195	0.0220		0.4810	0.6790	0.1235	0.4630			0.0170	0.0180		0.6225	0.6825	0.2365	0.5220
100	$\mathrm{CSDH}_{\mathbf{e}}$	0.30		0.0380	0.0345		0.9470	0.9945	0.5315	0.9630			0.0420	0.0355		0.9905	0.9970	0.7945	0.9785
		0.50		0.0480	0.0430		0.9970	1.0000	0.9045	1.0000			0.0395	0.0410		1.0000	1.0000	0.9925	1.0000
		0.15		0.0375	0.0285		0.9720	0.9885	0.7070	0.9690			0.0390	0.0350		0.9875	0.9875	0.9100	0.9805
200	$\mathrm{CSI}_{\mathbf{e}}$	0.30		0.0485	0.0380		0.9995	1.0000	0.9940	1.0000			0.0480	0.0475		1.0000	1.0000	0.9995	1.0000
		0.50		0.0500	0.0520		1.0000	1.0000	0.9995	1.0000			0.0445	0.0425		1.0000	1.0000	1.0000	1.0000
		0.15		0.0410	0.0365		0.9265	0.9890	0.4700	0.9540			0.0330	0.0340		0.9775	0.9875	0.7395	0.9750
200	$\mathrm{CSD}_{\mathbf{e}}$	0.30		0.0470	0.0410		0.9990	1.0000	0.9155	1.0000			0.0440	0.0460		1.0000	1.0000	0.9945	1.0000
		0.50		0.0510	0.0450		1.0000	1.0000	0.9975	1.0000			0.0490	0.0450		1.0000	1.0000	1.0000	1.0000
		0.15		0.0455	0.0375		0.9240	0.9885	0.4715	0.9525			0.0315	0.0320		0.9785	0.9880	0.7450	0.9735
200	$\mathrm{CSDH}_{\mathbf{e}}$	0.30		0.0460	0.0375		0.9990	1.0000	0.9145	1.0000			0.0465	0.0475		1.0000	1.0000	0.9960	1.0000
		0.50		0.0550	0.0465		1.0000	1.0000	0.9985	1.0000			0.0495	0.0475		1.0000	1.0000	1.0000	1.0000
		0.15		0.0380	0.0375		1.0000	1.0000	0.9870	1.0000			0.0375	0.0380		1.0000	1.0000	1.0000	1.0000
400	$\mathrm{CSI}_{\mathbf{e}}$	0.30		0.0525	0.0430		1.0000	1.0000	1.0000	1.0000			0.0530	0.0475		1.0000	1.0000	1.0000	1.0000
		0.50		0.0480	0.0410		1.0000	1.0000	1.0000	1.0000			0.0510	0.0490		1.0000	1.0000	1.0000	1.0000
		0.15		0.0385	0.0425		0.9995	1.0000	0.8765	1.0000			0.0420	0.0390		1.0000	1.0000	0.9905	1.0000
400	$\mathrm{CSD}_{\mathbf{e}}$	0.30		0.0500	0.0490		1.0000	1.0000	0.9985	1.0000			0.0475	0.0510		1.0000	1.0000	1.0000	1.0000
		0.50		0.0460	0.0415		1.0000	1.0000	1.0000	1.0000			0.0470	0.0455		1.0000	1.0000	1.0000	1.0000
		0.15		0.0380	0.0440		0.9995	1.0000	0.8770	1.0000			0.0395	0.0410		1.0000	1.0000	0.9910	1.0000
400	$\mathrm{CSDH}_{\mathbf{e}}$	0.30		0.0505	0.0500		1.0000	1.0000	0.9980	1.0000			0.0475	0.0515		1.0000	1.0000	1.0000	1.0000
		0.50		0.0440	0.0445		1.0000	1.0000	1.0000	1.0000			0.0480	0.0455		1.0000	1.0000	1.0000	1.0000

Table 2-*Continued*: Experiment 2

Panel B: $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ Statistic																		
N			50								100							
			Size			Power					Size			Power				
			$\delta_{\mathbf{f}}$	0.00	1.75	$\delta_{\mathbf{f}}$	0.00		1.75		$\delta_{\mathbf{f}}$	0.00	1.75	$\delta_{\mathbf{f}}$	0.00		1.75	
			δ^0	0.00	0.00	δ^0	0.25	1.00	0.25	1.00	δ^0	0.00	0.00	δ^0	0.25	1.00	0.25	1.00
T	**DGP**	π^0																
100	$\text{CSI}_{\mathbf{e}}$	0.50		0.0075	0.0085		0.9935	1.0000	0.8345	0.9880		0.0120	0.0100		0.9990	0.9995	0.9730	0.9980
	$\text{CSD}_{\mathbf{e}}$	0.50		0.0070	0.0085		0.9340	0.9985	0.5565	0.9860		0.0105	0.0100		0.9950	0.9995	0.8670	0.9960
	$\text{CSDH}_{\mathbf{e}}$	0.50		0.0085	0.0100		0.9325	0.9985	0.5655	0.9845		0.0110	0.0100		0.9960	0.9995	0.8635	0.9955
200	$\text{CSI}_{\mathbf{e}}$	0.50		0.0240	0.0195		1.0000	1.0000	1.0000	1.0000		0.0200	0.0180		1.0000	1.0000	1.0000	1.0000
	$\text{CSD}_{\mathbf{e}}$	0.50		0.0215	0.0215		1.0000	1.0000	0.9800	1.0000		0.0180	0.0190		1.0000	1.0000	0.9995	1.0000
	$\text{CSDH}_{\mathbf{e}}$	0.50		0.0220	0.0210		1.0000	1.0000	0.9790	1.0000		0.0180	0.0170		1.0000	1.0000	0.9995	1.0000
400	$\text{CSI}_{\mathbf{e}}$	0.50		0.0240	0.0300		1.0000	1.0000	1.0000	1.0000		0.0275	0.0260		1.0000	1.0000	1.0000	1.0000
	$\text{CSD}_{\mathbf{e}}$	0.50		0.0260	0.0260		1.0000	1.0000	1.0000	1.0000		0.0295	0.0270		1.0000	1.0000	1.0000	1.0000
	$\text{CSDH}_{\mathbf{e}}$	0.50		0.0265	0.0275		1.0000	1.0000	1.0000	1.0000		0.0290	0.0225		1.0000	1.0000	1.0000	1.0000
1000	$\text{CSI}_{\mathbf{e}}$	0.50		0.0470	0.0450		1.0000	1.0000	1.0000	1.0000		0.0445	0.0425		1.0000	1.0000	1.0000	1.0000
	$\text{CSD}_{\mathbf{e}}$	0.50		0.0505	0.0440		1.0000	1.0000	1.0000	1.0000		0.0365	0.0415		1.0000	1.0000	1.0000	1.0000
	$\text{CSDH}_{\mathbf{e}}$	0.50		0.0505	0.0420		1.0000	1.0000	1.0000	1.0000		0.0375	0.0420		1.0000	1.0000	1.0000	1.0000

This table presents size and power results for the $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ statistic as defined in (6) for $\theta = \theta^0$ (Panel A), and for the $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ statistic defined in (7) (Panel B). The DGP is detailed in Section 5.2.1. $\text{CSI}_{\mathbf{e}}$ denotes time homoskedastic and cross-sectionally independent idiosyncratic components. $\text{CSD}_{\mathbf{e}}$ denotes time homoskedastic and cross-sectionally dependent idiosyncratic components. $\text{CSDH}_{\mathbf{e}}$ denotes time heteroskedastic and cross-sectionally dependent idiosyncratic components. The results are computed over $S = 2000$ replications according to (10), with $\hat{\mathcal{T}}^s = \widehat{LM}_{\mathbf{w}}^{\text{HAC},s}(\theta^0), \sup \widehat{LM}_{\mathbf{w}}^{\text{HAC},s}$.

Table 3: Experiment 3

Panel A: $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta^0)$ Statistic															
N	50								100						
	Size	**Power**							**Size**	**Power**					
		α^0	0.60		0.80		1.00			α^0	0.60		0.80		1.00
		δ_i^0	0.25	1.00	0.25	1.00	0.25	1.00		δ_i^0	0.25	1.00	0.25	1.00	0.25
T															
100	0.0420		0.1620	0.8235	0.4965	0.9915	0.9925	1.0000	0.0340		0.0910	0.3015	0.7245	0.9940	1.0000
200	0.0495		0.2975	0.9860	0.8175	1.0000	1.0000	1.0000	0.0490		0.1650	0.5640	0.9770	0.9995	1.0000
400	0.0485		0.5730	1.0000	0.9930	1.0000	1.0000	1.0000	0.0490		0.2800	0.8640	1.0000	1.0000	1.0000

Note: the last column of the N = 100 block (α⁰ = 1.00, δ⁰ = 1.00) is given below:

T	1.00
100	1.0000
200	1.0000
400	1.0000

Panel B: $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ Statistic																
N	50								100							
	Size	**Power**							**Size**	**Power**						
		α^0	0.60		0.80		1.00			α^0	0.60		0.80		1.00	
		δ_i^0	0.25	1.00	0.25	1.00	0.25	1.00		δ_i^0	0.25	1.00	0.25	1.00	0.25	1.00
T																
100	0.0095		0.0305	0.5140	0.1770	0.9460	0.9020	0.9995	0.0090		0.0215	0.0865	0.3530	0.9640	0.9915	0.9995
200	0.0155		0.1350	0.9365	0.5885	0.9985	0.9995	1.0000	0.0195		0.0570	0.3285	0.8705	0.9995	1.0000	1.0000
400	0.0275		0.3500	1.0000	0.9485	1.0000	1.0000	1.0000	0.0300		0.1235	0.7025	0.9985	1.0000	1.0000	1.0000
1000	0.0355		0.8390	1.0000	1.0000	1.0000	1.0000	1.0000	0.0425		0.4215	0.9905	1.0000	1.0000	1.0000	1.0000

This table presents size and power results for the $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ statistic as defined in (6) for $\theta = \theta^0$ (Panel A), and for the $\sup \widehat{LM}_{\mathbf{w}}^{\text{HAC}}$ statistic defined in (7) (Panel B). The DGP is detailed in Section 5.3.1. The results are computed over $S = 2000$ replications according to (10), with $\hat{\mathcal{T}}^s = \widehat{LM}_{\mathbf{w}}^{\text{HAC},s}(\theta^0), \sup \widehat{LM}_{\mathbf{w}}^{\text{HAC},s}$.

Table 4: Experiment 4

N	50						100					
	Size, $P^0 = 2$		**Power,** $P^0 = 1$				**Size,** $P^0 = 2$		**Power,** $P^0 = 1$			
δ^0	0.00		0.25		1.00		0.00		0.25		1.00	
$\bar{P}$	P^0+2	P^0+4	$2P^0+2$	$2P^0+4$	$2P^0+2$	$2P^0+4$	P^0+2	P^0+4	$2P^0+2$	$2P^0+4$	$2P^0+2$	$2P^0+4$
T												
100	0.0365	0.0330	0.9800	0.9605	0.9975	0.9600	0.0315	0.0280	0.9995	0.9965	1.0000	0.9955
200	0.0495	0.0410	1.0000	1.0000	0.9995	0.9980	0.0455	0.0410	1.0000	1.0000	1.0000	1.0000
400	0.0490	0.0430	1.0000	1.0000	1.0000	1.0000	0.0425	0.0380	1.0000	1.0000	1.0000	1.0000

This table presents size and power results for the $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ statistic as defined in (6) for $\theta = \theta^0$. The DGP is detailed in Section 5.4. The results are computed over $S = 2000$ replications according to (10), with $\hat{\mathcal{T}}^s = \widehat{LM}_{\mathbf{w}}^{\text{HAC},s}(\theta^0)$.

Table 5: Experiment 5

N	50									100								
	Size			**Power**						**Size**			**Power**					
δ^0	0.00			0.25			1.00			0.00			0.25			1.00		
d.o.f.	20	15	10	20	15	10	20	15	10	20	15	10	20	15	10	20	15	10
T																		
100	0.0325	0.0470	0.0415	0.9720	0.9665	0.9295	0.9990	0.9995	0.9995	0.0470	0.0380	0.0470	0.9990	0.9980	0.9965	1.0000	1.0000	1.0000
200	0.0485	0.0460	0.0485	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0440	0.0420	0.0440	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
400	0.0480	0.0560	0.0480	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0490	0.0455	0.0495	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

This table presents size and power results for the $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ statistic as defined in (6) for $\theta = \theta^0$. The DGP is detailed in Section 5.5. The results are computed over $S = 2000$ replications according to (10), with $\hat{\mathcal{T}}^s = \widehat{LM}_{\mathbf{w}}^{\text{HAC},s}(\theta^0)$.

Table 6: Experiment 6

N	100										200									
	Size					**Power**					**Size**					**Power**				
P^0	2	4	6	8	10	2	4	6	8	10	2	4	6	8	10	2	4	6	8	10
T																				
200	0.0465	0.0435	0.0440	0.0340	0.0285	1.0000	1.0000	1.0000	1.0000	1.0000	0.0480	0.3850	0.0380	0.0345	0.0340	1.0000	1.0000	1.0000	1.0000	1.0000
400	0.0465	0.0465	0.0325	0.0350	0.0345	1.0000	1.0000	1.0000	1.0000	1.0000	0.0550	0.0450	0.0435	0.0480	0.0450	1.0000	1.0000	1.0000	1.0000	1.0000

This table presents size and power results for the $\widehat{LM}_{\mathbf{w}}^{\text{HAC}}(\theta)$ statistic as defined in (6) for $\theta = \theta^0$. The DGP is detailed in Section 5.6. The results are computed over $S = 2000$ replications according to (10), with $\hat{\mathcal{T}}^s = \widehat{LM}_{\mathbf{w}}^{\text{HAC},s}(\theta^0)$.

Table 7: Industry Portfolios, Portfolio Weights, Descriptive Statistics and Correlations, January 1985 - February 2020

Panel A: Descriptive Statistics, EPU_{t-1}				
	$\text{EPU}_{t-1} \leq \hat{\theta}$		$\text{EPU}_{t-1} > \hat{\theta}$	
	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_1$	$\hat{f}_2$
Std. dev	0.0047	0.9076	0.0048	1.1728
Median	0.0206	-0.2795	0.0205	0.3720
Maximum	0.0287	4.4201	0.0311	0.9663
Minimum	0.0070	-0.5778	0.0103	-5.8593
Skewness	-0.3535	3.6119	-0.0181	-3.5119
Kurtosis	0.3210	14.1876	-0.3180	14.3721
Panel B: Descriptive Statistics, MOM_{t-1}				
	$\text{MOM}_{t-1} \leq \hat{\theta}$		$\text{MOM}_{t-1} > \hat{\theta}$	
	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_1$	$\hat{f}_2$
Std. dev	0.0059	0.3012	0.0044	2.2883
Median	0.0204	0.1217	0.0204	-0.7432
Maximum	0.0319	0.2947	0.0298	11.9920
Minimum	0.0077	-1.2963	0.0095	-1.6747
Skewness	-0.0942	-2.8913	-0.1972	3.8302
Kurtosis	-0.5745	9.4633	-0.0833	16.8458
Panel C: Correlations				
	EPU_{t-1}		MOM_{t-1}	
	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_1$	$\hat{f}_2$
Correlations	0.8999	-0.9897	0.9601	-0.9213

For industry portfolios as described in Section 6.1 this table presents: descriptive statistics for the $\hat{P}$ sequences of $N \times 1$ portfolio weights associated to the estimated factors $\hat{f}_p$, for $p = 1, \ldots, \hat{P}$, for the regimes $z_t \leq \hat{\theta}$ and $z_t > \hat{\theta}$, with $z_t = \text{EPU}_{t-1}$ (Panel A) and $z_t = \text{MOM}_{t-1}$ (Panel B); the $\hat{P}$ correlations between the sequences of $N \times 1$ portfolio weights associated to the estimated factor $\hat{f}_p$ for $z_t \leq \hat{\theta}$ and $z_t > \hat{\theta}$, for $p = 1, \ldots, \hat{P}$ and $z_t = \text{EPU}_{t-1}, \text{MOM}_{t-1}$ (Panel C).

Table 8: Combined Portfolios, Portfolio Weights, Descriptive Statistics and Correlations, January 1985 - February 2020

Panel A: Descriptive Statistics, EPU_{t-1}

	$\text{EPU}_{t-1} \leq \hat{\theta}$						$\text{EPU}_{t-1} > \hat{\theta}$					
	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$	$\hat{f}_5$	$\hat{f}_6$	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$	$\hat{f}_5$	$\hat{f}_6$
Std. dev	0.0005	0.0908	0.0570	0.0907	0.0530	4.0025	0.0005	0.1048	0.1524	0.0408	0.0675	1.4731
Median	0.0038	0.0207	0.0143	0.0117	0.0063	-0.0360	0.0038	0.0189	0.0074	0.0139	0.0157	0.1844
Maximum	0.0055	0.1922	0.1238	0.1196	0.2167	26.8906	0.0057	0.2004	0.4470	0.0984	0.1065	5.4602
Minimum	0.0021	-0.1727	-0.2357	-1.1892	-0.2249	-12.9026	0.0021	-0.1976	-0.7004	-0.1740	-0.7450	-9.0756
Skewness	0.1037	-0.1984	-1.1687	-9.4994	-0.4892	0.9269	0.4650	-0.1929	-1.1760	-1.3893	-6.6107	-0.8285
Kurtosis	1.0420	-1.1628	2.1174	119.5639	4.3656	7.7509	1.7420	-1.1593	4.3950	2.9939	64.6351	7.3264

Panel B: Descriptive Statistics, MOM_{t-1}

	$\text{MOM}_{t-1} \leq \hat{\theta}$							$\text{MOM}_{t-1} > \hat{\theta}$						
	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$	$\hat{f}_5$	$\hat{f}_6$	$\hat{f}_7$	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$	$\hat{f}_5$	$\hat{f}_6$	$\hat{f}_7$
Std. dev	0.0006	0.0715	0.3369	0.0611	0.0392	0.1926	0.3035	0.0005	0.1112	0.0851	0.0710	0.0502	0.3417	0.4541
Median	0.0039	0.0125	-0.0434	0.0120	0.0088	0.0135	0.0211	0.0039	0.0238	0.0166	0.0143	0.0129	-0.0160	0.0213
Maximum	0.0058	0.1514	1.1887	0.4828	0.0810	0.6596	1.4846	0.0053	0.2232	0.2152	0.0906	0.1632	1.7766	1.8771
Minimum	0.0017	-0.1454	-1.0738	-0.2698	-0.3393	-0.8820	-1.6469	0.0023	-0.2022	-0.3450	-0.8914	-0.2211	-1.9223	-1.5460
Skewness	0.1809	-0.1623	0.2666	1.4244	-3.5272	-0.4191	-0.6154	0.1857	-0.1951	-0.9045	-8.5428	-1.0323	-0.0455	0.1832
Kurtosis	1.7691	-1.1458	1.4097	16.9446	24.3900	2.3385	8.2574	1.1597	-1.1873	1.9456	101.6855	3.3228	5.7637	2.0536

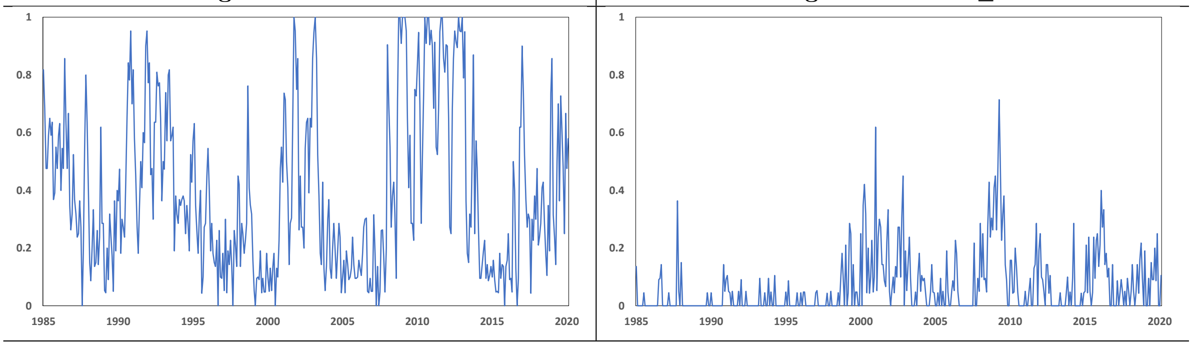
Panel C: Correlations

	EPU_{t-1}						MOM_{t-1}						
	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$	$\hat{f}_5$	$\hat{f}_6$	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$	$\hat{f}_5$	$\hat{f}_6$	$\hat{f}_7$
Correlations	0.8392	0.9743	-0.7521	0.0217	0.0611	0.3610	0.9163	0.9726	-0.8793	-0.4171	0.4167	-0.5527	-0.5605

For the combined portfolios as described in Section 6.1 this table presents: descriptive statistics for the $\hat{P}$ sequences of $N \times 1$ portfolio weights associated to the estimated factors $\hat{f}_p$, for $p = 1, \ldots, \hat{P}$, for the regimes $z_t \leq \hat{\theta}$ and $z_t > \hat{\theta}$, with $z_t = \text{EPU}_{t-1}$ (Panel A) and $z_t = \text{MOM}_{t-1}$ (Panel B); the $\hat{P}$ correlations between the sequences of $N \times 1$ portfolio weights associated to the estimated factor $\hat{f}_p$ for $z_t \leq \hat{\theta}$ and $z_t > \hat{\theta}$, for $p = 1, \ldots, \hat{P}$ and $z_t = \text{EPU}_{t-1}, \text{MOM}_{t-1}$ (Panel C).

Figure 1: Industry Portfolios, Monthly Frequencies of Regimes, January 1985 - February 2020

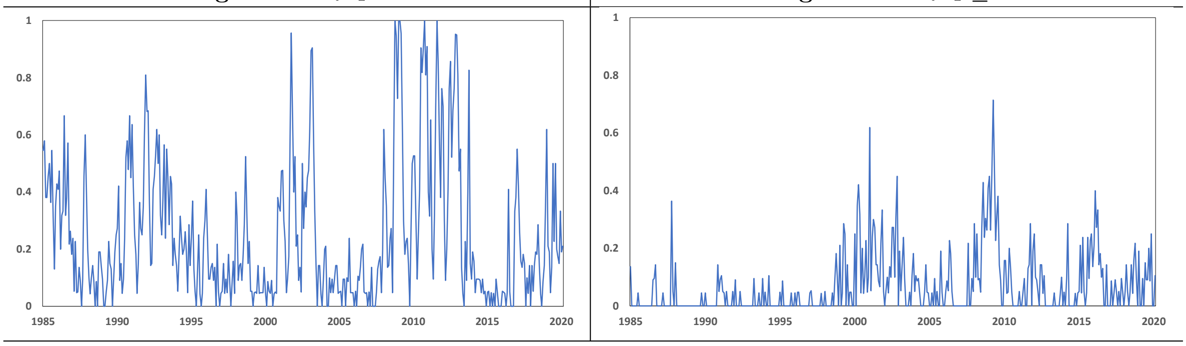
Regime $\text{EPU}_{t-1} > \hat{\theta}$	**Regime** $\text{MOM}_{t-1} \leq \hat{\theta}$



This figure shows the monthly averages of the daily indicator functions $\mathbb{I}\left(\text{EPU}_{t-1} > \hat{\theta}\right)$ and $\mathbb{I}\left(\text{MOM}_{t-1} \leq \hat{\theta}\right)$ (see left and right panel, respectively) for the models with industry portfolios as described in Section 6.1. The sample period is from January 1985 to February 2020, a total of 422 monthly observations.

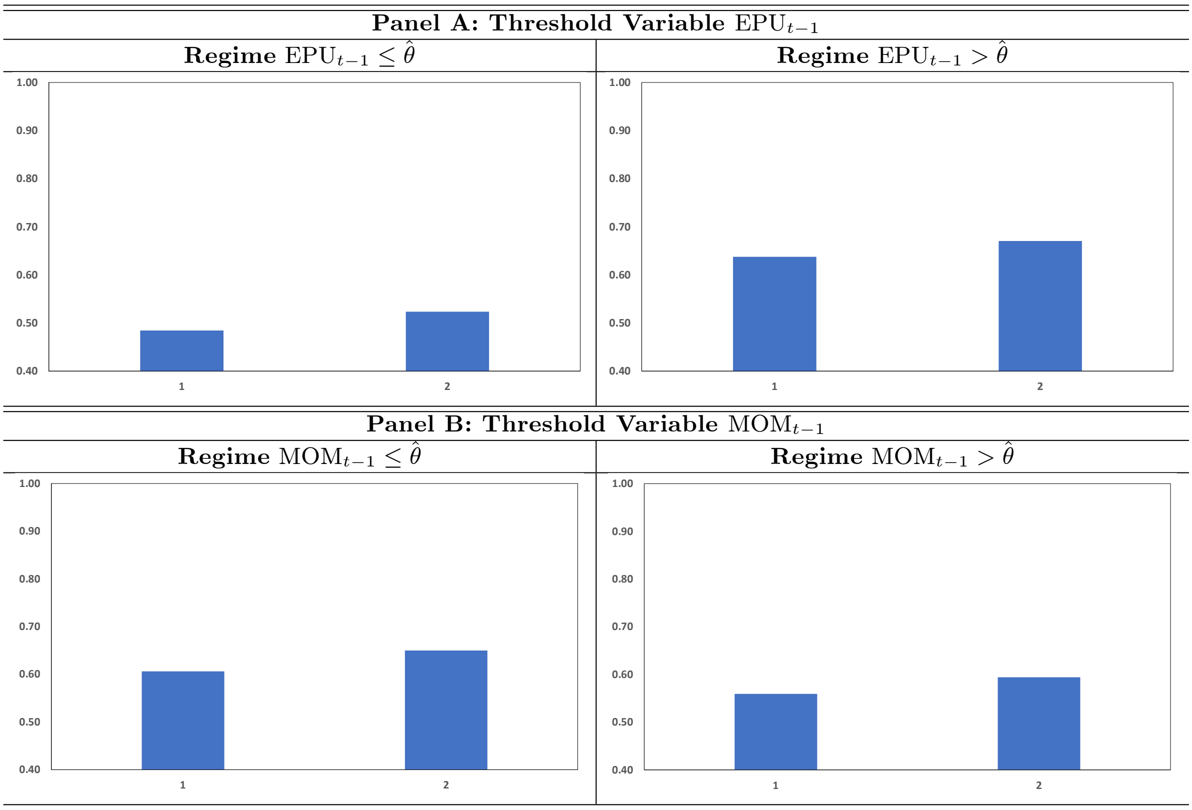
Figure 2: Combined Portfolios, Monthly Frequencies of Regimes, January 1985 - February 2020

Regime $\text{EPU}_{t-1} > \hat{\theta}$	**Regime** $\text{MOM}_{t-1} \leq \hat{\theta}$



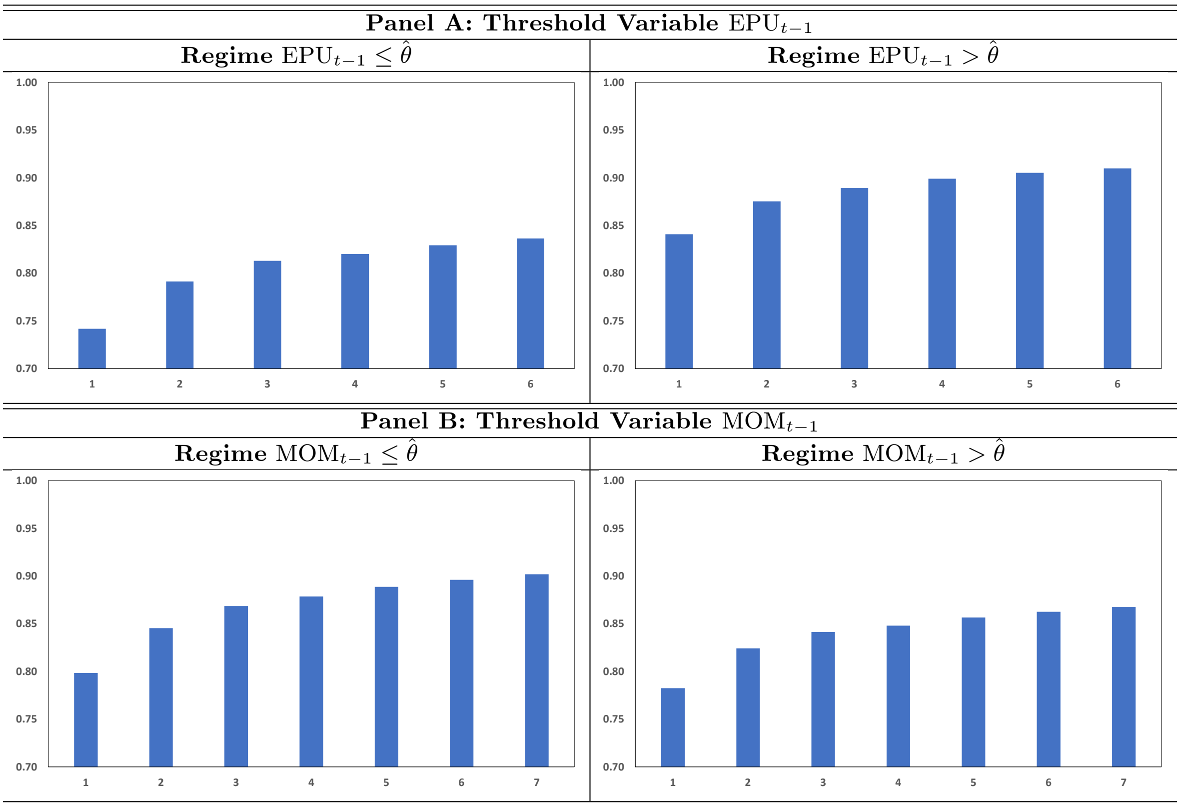
This figure shows the monthly averages of the daily indicator functions $\mathbb{I}\left(\text{EPU}_{t-1} > \hat{\theta}\right)$ and $\mathbb{I}\left(\text{MOM}_{t-1} \leq \hat{\theta}\right)$ (see left and right panel, respectively) for the models with combined portfolios as described in Section 6.1. The sample period is from January 1985 to February 2020, a total of 422 monthly observations.

Figure 3: Industry Portfolios, Average Cumulated R − squared



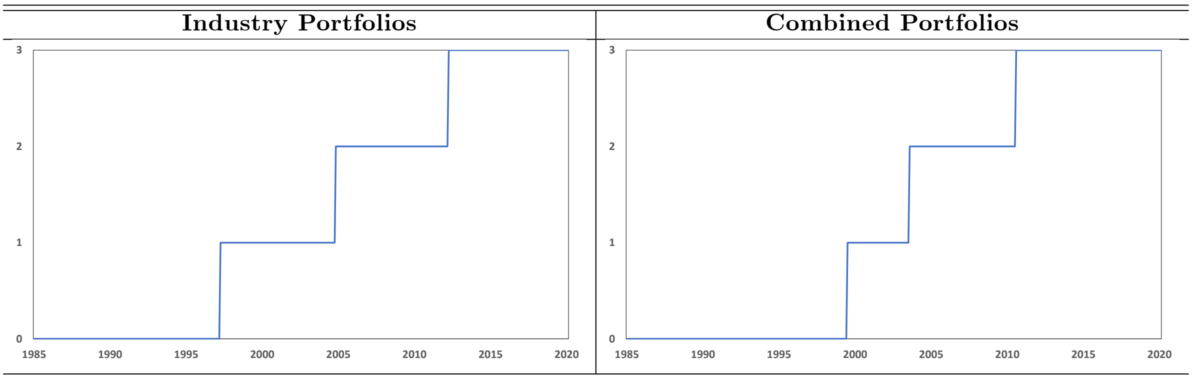
For the models with industry portfolios as described in Section 6.1, this figure shows the average cumulated R − squared for the $\hat{P} = 2$ estimated factors for threshold variables $z_t = \text{EPU}_{t-1}$ (Panel A) and $z_t = \text{MOM}_{t-1}$ (Panel B), and regimes $z_t \leq \hat{\theta}$ and $z_t > \hat{\theta}$.

Figure 4: Combined Portfolios, Average Cumulated R − squared



For the models with combined portfolios as described in Section 6.1, this figure shows the average cumulated R − squared for the $\hat{P} = 6$ and $\hat{P} = 7$ estimated factors for threshold variables $z_t = \mathrm{EPU}_{t-1}$ (Panel A) and $z_t = \mathrm{MOM}_{t-1}$ (Panel B), respectively, and regimes $z_t \leq \hat{\theta}$ and $z_t > \hat{\theta}$.

Figure 5: Industry and Combined Portfolios, Estimated Break Dates, January 1985 - February 2020



This figure shows the estimated break dates for the models with industry and combined portfolios as described in Section 6.1. The sample period is from January 1985 to February 2020.

The Data Analytics for Finance and Macro (DAFM) Research Centre
Forecasting trends is more important now than it has ever been. Our quantitative financial and macroeconomic research helps central banks and statistical agencies better understand their markets. Having our base in central London means we can focus on issues that really matter to the City. Our emphasis on digitalisation and the use of big data analytics can help people understand economic behaviour and adapt to changes in the economy.

Find out more at kcl.ac.uk/dafm

KING'S BUSINESS SCHOOL

Data Analytics for Finance & Macro Research Centre