# Least Squares Estimation of Large Dimensional Threshold Factor Models*

Daniele Massacci

Einaudi Institute for Economics and Finance

September 8, 2015

**Abstract**

Large dimensional factor models are estimated under the maintained assumption that the loadings do not change over time. This paper studies least squares estimation of large dimensional factor models subject to regime shifts in the loadings parameterized according to the threshold principle. We estimate the unknown threshold value by concentrated least squares, and factors and loadings by principal components. The estimator for the threshold value is superconsistent, with convergence rate that depends on the times series and the cross-sectional dimensions of the available panel, and it does not affect the estimator for factors and loadings: this has the same convergence rate as in factor models with time-invariant loadings. We further propose model selection criteria robust to the threshold effect. Empirical application of the model documents an increase in connectedness in financial variables during periods of high economic policy uncertainty.

**JEL classification:** C13, C33, C52, G10.

**Keywords:** Large Factor Model, Threshold Effect, Least Squares Estimation, Model Selection, Connectedness.

*I really am highly indebted to Marco Lippi for introducing me to factor models and for several enlightening conversations. Kind hospitality by the Department of Finance at Bocconi University is acknowledged. This paper benefits from comments from seminar participants at University of Rome Tor Vergata and Bocconi University; from suggestions from conference participants at the Vienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance, and at the IAAE 2015 Annual Conference; and from conversations with Domenico Giannone, Alessandro Giovannelli, Hashem Pesaran and Stefano Soccorsi. Errors and omissions are my own responsibility. Financial support from the Associazione Borsisti Marco Fanno and from UniCredit and Universities Foundation is gratefully acknowledged. Address correspondence to Daniele Massacci, EIEF, Via Sallustiana 62, 00187 Roma, Italy. Tel.: +39 06 4792 4974. E-mail: dm355@cantab.net.

# 1 Introduction

Factor models are widely used tools to explain the common variations in large scale macroeconomic and financial data. An extensive literature analyzes factor models under the maintained assumption of constant loadings over the entire sample period: see Connor and Korajczyk (1986, 1988, 1993), Forni *et al.* (2000, 2004, 2015), Forni and Lippi (2001), Bai and Ng (2002), Stock and Watson (2002), and Bai (2003) for seminal contributions on linear factor models. Economic models are however unlikely to have constant parameters over time and factor models with time-dependent loadings are called for. Time-dependence in the loadings may be easily implemented through a change-point mechanism: this may be parameterized as either a structural break or a regime shift driven by the threshold principle, depending on the underlying data generating process.

Structural breaks in the loadings may arise as a consequence of events such as technological or policy changes. Several important contributions deal with large dimensional factor models subject to loadings instabilities. Breitung and Eickmeier (2011) show that ignoring breaks leads to overestimation of the number of factors and develop statistical tests for the null hypothesis of stability in the factor loadings. Bates *et al.* (2013) study the robustness properties of the principal components estimator of the factors under neglected loadings instability. Chen *et al.* (2014), Han and Inoue (2014) and Yamamoto and Tanaka (2015) develop further statistical tools to detect breaks. Chen (2015) considers least squares estimation of the break date. Cheng *et al.* (2015) propose shrinkage estimation of large dimensional factor models with structural breaks.

Regime shift representations of the dependent variables are suitable when "history repeats", as in the case of financial returns (Timmermann (2008), and Ang and Timmermann (2012)). Ng and Wright (2013) introduce a threshold mechanism within large dimensional factor models to simulate data and investigate the effects of nonlinearities on business cycle dynamics[1]. We take Ng and Wright (2013) intuition as a starting point and propose a large dimensional factor model with regime changes in the loadings governed by the threshold principle. We let the threshold value be unknown and focus on estimation and model selection. To the very best of our knowledge, we are the first to tackle this problem: we therefore believe this paper provides a valuable contribution to the literature on large dimensional factor models.

Let $R^0$ be the true number of factors. Under the maintained assumption that $R^0$ is known, we

[1] See Ng and Wright (2013), p. 1147.

propose to estimate the threshold value by concentrated least squares, and factors and loadings by principal components (Hansen (2000), and Bai and Ng (2002)). We obtain a number of novel theoretical results. Let $N$ and $T$ denote the cross-sectional and time series dimensions, respectively. We first provide sufficient conditions to ensure that our model is identified from a linear factor model: formally, for $0.5 < \alpha^0 \leq 1$, we require that *at least* a fraction $O\left(N^{\alpha^0}\right)$ of the $N$ cross-sectional units experiences a regime shift in the loadings, so that the shift resists to the aggregation induced by the principal components estimator. We then show that the estimator for the threshold parameter is consistent at a rate equal to $N^{\left(2\alpha^0-1\right)}T$: this depends on the time series dimension $T$ and on the number of cross-sectional units $N^{\alpha^0}$ subject to the threshold effect. The convergence rate monotonically increases in $\alpha^0$ and it is such that $T < N^{\left(2\alpha^0-1\right)}T \leq NT$: this shows the direct relationship between identification of the model and convergence rate of the estimator for the threshold. As a consequence of this superconsistency property, we finally show that the principal components estimator for both regime-specific loadings and factors have convergence rate equal to $C_{NT} = \min\left\{\sqrt{N}, \sqrt{T}\right\}$: despite the threshold effect, the convergence rate $C_{NT}$ is equal to the one derived in Bai and Ng (2002) for linear factor models.

We next consider the case in which the true number of factors $R^0$ no longer is known and has to be estimated. Breitung and Eickmeier (2011) show that structural instability in the loadings leads to a factor representation with a higher dimensional factor space: due to an analogy argument, the same issue arises when a regime shift drives time variation in the loadings. Since the convergence rate $C_{NT}$ of the estimator for loadings and factors is the same as in linear factor models, we make Bai and Ng (2002) information criteria robust to the threshold effect by accounting for the induced higher dimensional factor space representation: this is a further theoretical contribution of our paper.

We finally provide an application to illustrate the potential usefulness of our methodology for applied work. We show how our theoretical framework may be used to measure connectedness in financial markets (Acharya *et al.* (2010), Adrian and Brunnermeier (2011), Billio *et al.* (2012), Engle and Kelly (2012), and Diebold and Yilmaz (2014)). We extend Billio *et al.* (2012) measure based on principal components analysis to allow for regime-specific connectedness. Using Baker *et al.* (2013) index of economic policy uncertainty as threshold variable, we show that connectedness in financial markets increases during periods of high uncertainty: this result is the empirical contribution of our paper and it may be relevant for risk measurement and management.

The large dimensional threshold factor model we consider and the factor model with structural instability are complementary specifications that allow for a change-point within a large dimensional factor structure. The analogy between the two models suggests that our approach may be used for estimation and model selection in factor models with structural instabilities: this is considered in Massacci (2015), which therefore contributes to an existing and growing literature (Breitung and Eickmeier (2011), Bates *et al.* (2013), Chen *et al.* (2014), Han and Inoue (2014), Yamamoto and Tanaka, Chen (2015) and Cheng *et al.* (2015)).

The remainder of the paper is organized as follows. Section 2 describes the threshold factor model. Section 3 deals with estimation. Section 4 proposes information criteria to determine the number of factors. Section 5 performs a comprehensive Monte Carlo analysis. Section 6 provides an empirical application. Finally, Section 7 concludes. Appendix A provides technical proofs.

Concerning notation, $\mathbb{I}(\cdot)$ denotes the indicator function; for a given scalar $A$, $\mathbf{I}_A$ and $\mathbf{0}_A$ are the $A \times A$ identity matrix and zero matrix, respectively.

## 2 The Approximate Threshold Factor Model

We consider the model

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

where $T$ denotes the time series dimension of the available sample; $\mathbf{x}_t = (x_{1t}, \ldots, x_{Nt})' \in \Re^N$ is the $N \times 1$ vector of observable dependent variables; $\mathbf{f}_t = (f_{1t}, \ldots, f_{Rt})' \in \Re^R$ is the $R \times 1$ vector of latent factors; $z_t \in \Re$ is an observable covariate and $\theta$ is the unknown threshold value; $\mathbf{e}_t = (e_{1t}, \ldots, e_{Nt})' \in \Re^N$ is the $N \times 1$ vector of idiosyncratic errors; $\mathbf{\Lambda}_j = (\boldsymbol{\lambda}_{j1}, \ldots, \boldsymbol{\lambda}_{jN})'$ is the $N \times R$ matrix of factor loadings with $i-th$ row defined as $\boldsymbol{\lambda}_{ji} = (\lambda_{ji1}, \ldots, \lambda_{jiR})'$, for $j = 1, 2$ and $i = 1, \ldots, N$.

The model in (1) belongs to the class of threshold models proposed in Tong and Lim (1980): see Tsay (1989, 1998), Chan (1993) and Hansen (1996, 2000) for methodological contributions; and Hansen (2011) for a survey of the literature. According to the threshold principle introduced in Pearson (1900), the regime prevailing at time $t$ depends on the position of $z_t$ with respect to the unknown threshold value $\theta$. Ng and Wright (2013) simulate data from a large dimensional threshold factor model to investigate the effects of nonlinearities on business cycle dynamics[2]: we explicitly focus on estimation and model selection. Our results apply to the case in which the threshold variable is a more general linear combination of exogenous covariates (Massacci (2014)).

The model in (1) extends large dimensional linear factor models to allow for a threshold effect. Given Assumption C3 stated in Section 3.1 below, we follow Chamberlain and Rothschild (1983) and allow for some degree of correlation in the idiosyncratic components within each regime: (1) then is an *approximate threshold factor model*; it is more general than an *exact threshold factor model*, which would extend the arbitrage pricing theory of Ross (1976) and would not allow for any correlation in the idiosyncratic components in any regime.

# 3 Estimation

As in Stock and Watson (2002), we study estimation of (1) under the assumption that the true number of factors $R^0$ (i.e., the true dimension of $\mathbf{f}_t$) is known. We extend the theory in Bai and Ng (2002) based on principal components estimation to allow for concentrated least squares estimation, as motivated in Hansen (2000) for threshold regressions. The plan is as follows: Section 3.1 states the assumptions; Section 3.2 deals with identification; Section 3.3 describes the principal components estimator; Section 3.4 proves the consistency of the estimator; and Section 3.5 derives the convergence rates.

## 3.1 Assumptions

We group the assumptions into three sets, depending on the role they play to identify and estimate the model, and to derive the convergence rates. 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 $\boldsymbol{\Lambda}_j^0 = \left(\boldsymbol{\lambda}_{j1}^0, \ldots, \boldsymbol{\lambda}_{jN}^0\right)'$, $\theta^0$ and $\mathbf{f}_t^0$ the true values of $\boldsymbol{\Lambda}_j$, $\theta$ and $\mathbf{f}_t$, respectively. Define $\mathbf{f}_{jt}^0(\theta) = \mathbb{I}_{jt}(\theta)\mathbf{f}_t^0$, for $j = 1, 2$ and $t = 1, \ldots, T$, and let $\boldsymbol{\delta}_i^0 = \boldsymbol{\lambda}_{2i}^0 - \boldsymbol{\lambda}_{1i}^0$, for $i = 1, \ldots, N$.

### 3.1.1 Identification

**Assumption I - Threshold Factor Model.** For $0.5 < \alpha^0 \leq 1$, $\boldsymbol{\delta}_i^0 \neq \mathbf{0}$ for $i = 1, \ldots, N^{\alpha^0}$, and $\sum_{i=N^{\alpha^0}+1}^{N} \boldsymbol{\delta}_i^0 = O(1)$.

[2] See Ng and Wright (2013), p. 1147.

Assumption I requires that *at least* a fraction $O\left(N^{\alpha^0}\right)$ of the $N$ series experiences a threshold effect, for $0.5 < \alpha^0 \leq 1$: this follows up on Bates *et al.* (2013), who show that if at most $O\left(N^{0.5}\right)$ series undergo a break then the principal components estimator as applied to the misspecified linear model achieves the same Bai and Ng (2002) convergence rate. Assumption I ensures that enough series experience a regime shift so that (1) is identified from a linear factor model when factors and loadings are estimated by principal components. As shown in Theorem 3.4 below, $\alpha^0$ affects the convergence rate of the estimator for $\theta^0$: the higher the former, the faster the latter. In this paper we do not aim at estimating $\alpha^0$ and leave this interesting issue to future research.

#### 3.1.2 Consistency

**Assumption C1 - Factors.** $\mathrm{E}\left\|\mathbf{f}_t^0\right\|^4 < \infty$; for $j = 1, 2$, $T^{-1}\sum_{t=1}^T \mathbf{f}_{jt}^0(\theta)\,\mathbf{f}_{jt}^0\left(\theta^0\right)' \xrightarrow{p} \mathbf{\Sigma}_{j\mathbf{f}}^0\left(\theta,\theta^0\right)$ as $T \to \infty$ for all $\theta$ and some positive definite matrix $\mathbf{\Sigma}_{j\mathbf{f}}^0\left(\theta,\theta^0\right)$ such that $\mathbf{\Sigma}_{j\mathbf{f}}^0\left(\theta^0,\theta^0\right) - \mathbf{\Sigma}_{j\mathbf{f}}^0\left(\theta,\theta^0\right)$ is positive definite for all $\theta \neq \theta^0$.

**Assumption C2 - Factor Loadings.** For $j = 1, 2$ and $i = 1, \ldots, N$, $\left\|\boldsymbol{\lambda}_{ji}^0\right\| \leq \bar{\lambda} < \infty$, and $\left\|\mathbf{\Lambda}_j^{0\prime}\mathbf{\Lambda}_j^0/N - \mathbf{D}_{\mathbf{\Lambda}_j}^0\right\| \to 0$ as $N \to \infty$ for some $R^0 \times R^0$ positive definite matrix $\mathbf{D}_{\mathbf{\Lambda}_j}^0$.

**Assumption C3 - Time and Cross-Section Dependence and Heteroscedasticity.** There exists a positive $M < \infty$ such that for $j = 1, 2$, for all $\theta$ and for all $(N, T)$,

**(a)** $\mathrm{E}(e_{it}) = 0$ and $\mathrm{E}\left|e_{it}\right|^8 \leq M$;

**(b)** $\mathrm{E}\left[\mathbb{I}_{jt}(\theta)\,\mathbb{I}_{jv}(\theta)\,e_{it}e_{iv}\right] = \tau_{jitv}(\theta)$ with $\left|\tau_{jitv}(\theta)\right| \leq \left|\tau_{jtv}\right|$ for some $\tau_{jtv}$ and for all $i$, and $T^{-1}\sum_{t=1}^T\sum_{v=1}^T\left|\tau_{jtv}\right| \leq M$;

**(c)** $\mathrm{E}\left[T^{-1}\sum_{t=1}^T \mathbb{I}_{jt}(\theta)\,e_{it}e_{lt}\right] = \sigma_{jil}(\theta)$, $\left|\sigma_{jll}(\theta)\right| \leq M$ for all $l$, and $N^{-1}\sum_{i=1}^N\sum_{l=1}^N\left|\sigma_{jil}(\theta)\right| \leq M$;

**(d)** $\mathrm{E}\left|T^{-1/2}\sum_{t=1}^T \mathbb{I}_{jt}(\theta)\,e_{it}e_{lt} - \mathrm{E}\left[\mathbb{I}_{jt}(\theta)\,e_{it}e_{lt}\right]\right|^4 \leq M$ for every $(i, l)$.

**Assumption C4 - Weak Dependence between $\mathbf{f}_t^0$, $z_t$ and $e_{it}$.** There exists some positive constant $M < \infty$ such that for all $\theta$ and for all $(N, T)$,

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

Assumptions C1 to C4 are the natural extensions of Assumptions A to D imposed on linear factor models in Bai and Ng (2002) and accommodate the threshold effect. Assumption C1 restricts the sequences $\left\{\mathbf{f}_t^0\right\}_{t=1}^T$ and $\left\{z_t\right\}_{t=1}^T$ so that appropriate second moments exist; it also imposes full rank conditions that exclude multicollinearity in the factors. According to Assumption C2, factor loadings are nonstochastic and each factor has a nonnegligible effect on the variance of $\mathbf{x}_t$ within each regime. Under Assumption C3, limited degrees of time-series and cross-section dependence in the idiosyncratic components as well as heteroscedasticity are allowed. Finally, Assumption C4 provides an upper bound to the degree of dependence between the factors, $z_t$ and the idiosyncratic components: Assumption C4 is stronger than Assumption D in Bai and Ng (2002), which only bounds the dependence between the factors and the idiosyncratic components.

#### 3.1.3 Convergence Rates

Define $\mathbf{D}_{\mathbf{f}}^0(\theta)=\mathrm{E}\left(\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\left|z_t=\theta\right.\right)$ and denote by $f_{\mathcal{Z}}(z_t)$ the density function of $z_t$.

**Assumption CR - Stationarity, Moment Bound, Continuity and Full Rank.**

**(a)** $\left\{\mathbf{f}_t^0, z_t, \mathbf{e}_t\right\}_{t=1}^T$ is strictly stationary, ergodic and $\rho-$mixing, with $\rho-$mixing coefficients satisfying $\sum_{m=1}^{\infty}\rho_m^{1/2}<\infty$;

**(b)** $\mathrm{E}\left(\left\|\mathbf{f}_t^0 e_{it}\right\|^4\left|z_t=\theta\right.\right)\leq C$ for some $C<\infty$ and for $i=1,\ldots,N$, and $f_{\mathcal{Z}}(\theta)\leq\bar{f}<\infty$;

**(c)** $f_{\mathcal{Z}}(\theta)$ and $\mathbf{D}_{\mathbf{f}}^0(\theta)$ are continuous at $\theta=\theta^0$;

**(d)** $\boldsymbol{\delta}_i^{0\prime}\mathbf{D}_{\mathbf{f}}^0\left(\theta^0\right)\boldsymbol{\delta}_i^0>0$, $i=1,\ldots,N^{\alpha^0}$ and $0.5<\alpha^0\leq 1$, and $\sum_{i=N^{\alpha^0}+1}^{N}\boldsymbol{\delta}_i^{0\prime}\mathbf{D}_{\mathbf{f}}^0\left(\theta^0\right)\boldsymbol{\delta}_i^0=O(1)$; $f_{\mathcal{Z}}(\theta)>0$ for all $\theta$.

Assumption CR is analogous to Assumption 1 in Hansen (2000). Assumption CR(a) restricts the memory of the sequence $\left\{\mathbf{f}_t^0, z_t, \mathbf{e}_t\right\}_{t=1}^T$; it excludes trends and integrated processes. Assumption CR(b) gives a conditional moment bound. Assumption CR(c) imposes a continuous support on $z_t$. The full-rank condition in Assumption CR(d) strengthens Assumption I and rules out the "continuous threshold" set up of Chan and Tsay (1998), which arises in the one-factor model when the scalar factor $f_t^0$ equals the threshold variable $z_t$ and $\theta^0=0$: in this case, $\delta_i^0\mathrm{E}\left(f_t^0 f_t^0\left|z_t=\theta^0\right.\right)\delta_i^0=\delta_i^0\mathrm{E}\left(f_t^0 f_t^0\left|f_t^0=\theta^0\right.\right)\delta_i^0=0$, for $i=1,\ldots,N$, and Assumption CR(d) is violated.

### 3.2 Identification

Let $\boldsymbol{\Delta}^0 = \left(\boldsymbol{\delta}_1^0, \ldots, \boldsymbol{\delta}_N^0\right)'$ and write the data generating process of $\mathbf{x}_t$ as $\mathbf{x}_t = \boldsymbol{\Lambda}_1^0 \mathbf{f}_t^0 + \mathbb{I}_{2t}\left(\theta^0\right) \boldsymbol{\Delta}^0 \mathbf{f}_t^0 + \mathbf{e}_t$. Define $\mathbf{F}^0 = \left(\mathbf{f}_1^0, \ldots, \mathbf{f}_T^0\right)$ and denote $\tilde{\boldsymbol{\Lambda}}_1 = \left(\tilde{\boldsymbol{\lambda}}_{11}, \ldots, \tilde{\boldsymbol{\lambda}}_{1N}\right)'$ the principal components estimator for $\boldsymbol{\Lambda}_1^0$ from the misspecified linear factor model $\mathbf{x}_t = \boldsymbol{\Lambda}_1 \mathbf{f}_t + \mathbf{e}_t$. Let $\tilde{\mathbf{V}}_1$ be the $R^0 \times R^0$ diagonal matrix of the first $R^0$ largest eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}} = (NT)^{-1} \sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t'$ in decreasing order: the underlying optimization problem requires the normalization $N^{-1} \tilde{\boldsymbol{\Lambda}}_1' \tilde{\boldsymbol{\Lambda}}_1 = \mathbf{I}_{R^0}$. The following theorem states the properties of $\tilde{\boldsymbol{\Lambda}}_1$.

**Theorem 3.1** *There exists a $R^0 \times R^0$ rotation matrix $\tilde{\mathbf{H}}_1$ with* rank $\left[\tilde{\mathbf{H}}_1\right] = R^0$ *such that*

$$B_{NT}^2 \left(\frac{1}{N} \sum_{i=1}^N \left\| \tilde{\boldsymbol{\lambda}}_{1i} - \tilde{\mathbf{H}}_1' \boldsymbol{\lambda}_{1i}^0 \right\|^2\right) = O_p(1),$$

*as $N, T \to \infty$, where*

$$B_{NT} = \min\left\{\sqrt{N}, \sqrt{T}, N^{1-\alpha^0}\right\}$$

*and*

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

Theorem 3.1 shows that the average squared deviations between the loadings estimated under the null hypothesis of linearity and those that lie in the true loading space vanish as $N, T \to \infty$ at a rate equal to $B_{NT}^2$, which drives identification. Under Assumption I, the model in (1) is identified from the standard linear factor model as the rate of convergence $N^{1-\alpha^0}$ of the principal components estimator is slower than it would be under correct linear model specification: the model in (1) would not be identified from a linear factor model if $0 \leq \alpha^0 \leq 0.5$, since in this case $B_{NT}^2 = \min\{N, T\}$, as derived in Bai and Ng (2002). If $\alpha^0 = 1$ and all cross-sectional units are subject to threshold effect, $B_{NT}^2 = 1$ and the principal components estimator from the misspecified linear model is asymptotically biased. As proved in Theorem 3.4, the parameter $\alpha^0$ regulates the convergence rate of the estimator for the unknown threshold value $\theta^0$: this result shows the connection between identification strength and estimation precision.

### 3.3 Principal Components Estimation

We estimate the common factors and factor loadings by principal components, and the unknown threshold $\theta^0$ by concentrated least squares: see Bai and Ng (2002) and Hansen (2000), respectively. Define the

$N \times 2R^0$ matrix of loadings $\mathbf{\Lambda} = (\mathbf{\Lambda}_1, \mathbf{\Lambda}_2)$ and the $R^0 \times T$ matrix of factors $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_T)$. Let $\mathbf{\Lambda}^0 = \left(\mathbf{\Lambda}_1^0, \mathbf{\Lambda}_2^0\right)$ be the true value of $\mathbf{\Lambda}$. The objective function in terms of $\mathbf{\Lambda}$, $\mathbf{F}$ and $\theta$ is the sum of squared residuals (divided by $NT$)

$$S(\mathbf{\Lambda}, \mathbf{F}, \theta) = (NT)^{-1} \sum_{t=1}^{T} \left[\mathbf{x}_t - \mathbb{I}_{1t}(\theta)\mathbf{\Lambda}_1\mathbf{f}_t - \mathbb{I}_{2t}(\theta)\mathbf{\Lambda}_2\mathbf{f}_t\right]' \left[\mathbf{x}_t - \mathbb{I}_{1t}(\theta)\mathbf{\Lambda}_1\mathbf{f}_t - \mathbb{I}_{2t}(\theta)\mathbf{\Lambda}_2\mathbf{f}_t\right] : \quad (2)$$

the estimators $\hat{\mathbf{\Lambda}} = \left(\hat{\mathbf{\Lambda}}_1, \hat{\mathbf{\Lambda}}_2\right)$, $\hat{\mathbf{F}} = \left(\hat{\mathbf{f}}_1, \ldots, \hat{\mathbf{f}}_T\right)$ and $\hat{\theta}$ for $\mathbf{\Lambda}^0$, $\mathbf{F}^0$and $\theta^0$, respectively, with $\hat{\mathbf{\Lambda}}_j = \left(\hat{\boldsymbol{\lambda}}_{j1}, \ldots, \hat{\boldsymbol{\lambda}}_{jN}\right)'$, for $j = 1, 2$, jointly solve

$$\hat{\mathbf{\Lambda}}, \hat{\mathbf{F}}, \hat{\theta} = \arg\min_{\mathbf{\Lambda}, \mathbf{F}, \theta} S(\mathbf{\Lambda}, \mathbf{F}, \theta).$$

For given $\mathbf{\Lambda}$ and $\theta$, and subject to $N^{-1}\left(\mathbf{\Lambda}_j'\mathbf{\Lambda}_j\right) = \mathbf{I}_{R^0}$, for $j = 1, 2$, from (2) we have

$$\hat{\mathbf{f}}_t(\mathbf{\Lambda}, \theta) = N^{-1}\left[\mathbb{I}_{1t}(\theta)\mathbf{\Lambda}_1 + \mathbb{I}_{2t}(\theta)\mathbf{\Lambda}_2\right]' \mathbf{x}_t, \quad t = 1, \ldots, T : \quad (3)$$

replacing $\mathbf{f}_t$ in (2) with $\hat{\mathbf{f}}_t(\mathbf{\Lambda}, \theta)$ obtained in (3) leads to the concentrated objective function

$$S_{\mathbf{F}}(\mathbf{\Lambda}, \theta) = (NT)^{-1} \sum_{t=1}^{T} \mathbf{x}_t' \left\{\mathbf{I}_N - N^{-1}\left[\mathbb{I}_{1t}(\theta)\mathbf{\Lambda}_1\mathbf{\Lambda}_1' + \mathbb{I}_{2t}(\theta)\mathbf{\Lambda}_2\mathbf{\Lambda}_2'\right]\right\} \mathbf{x}_t, \quad (4)$$

and the estimators for $\mathbf{\Lambda}^0$ and $\theta^0$ jointly solve

$$\hat{\mathbf{\Lambda}}, \hat{\theta} = \arg\min_{\mathbf{\Lambda}, \theta} S_{\mathbf{F}}(\mathbf{\Lambda}, \theta).$$

From (4), the estimator for $\mathbf{\Lambda}^0$ for given $\theta$ is defined as

$$\hat{\mathbf{\Lambda}}(\theta) = \left[\hat{\mathbf{\Lambda}}_1(\theta), \hat{\mathbf{\Lambda}}_2(\theta)\right] = \arg\max_{\mathbf{\Lambda}} V_{\mathbf{F}}(\mathbf{\Lambda}, \theta), \quad (5)$$

where

$$\begin{aligned} V_{\mathbf{F}}(\mathbf{\Lambda}, \theta) &= (NT)^{-1} \sum_{t=1}^{T} \left\{\mathbf{x}_t' \left[\mathbb{I}_{1t}(\theta)(\mathbf{\Lambda}_1\mathbf{\Lambda}_1') + \mathbb{I}_{2t}(\theta)(\mathbf{\Lambda}_2\mathbf{\Lambda}_2')\right] \mathbf{x}_t\right\} \\ &= (NT)^{-1} \left\{\mathrm{tr}\left\{\mathbf{\Lambda}_1' \left[\sum_{t=1}^{T} \mathbb{I}_{1t}(\theta)\mathbf{x}_t\mathbf{x}_t'\right] \mathbf{\Lambda}_1\right\} + \mathrm{tr}\left\{\mathbf{\Lambda}_2' \left[\sum_{t=1}^{T} \mathbb{I}_{2t}(\theta)\mathbf{x}_t\mathbf{x}_t'\right] \mathbf{\Lambda}_2\right\}\right\}. \end{aligned}$$

The problem

$$\max_{\mathbf{\Lambda}} V_{\mathbf{F}}\left(\mathbf{\Lambda},\theta\right)$$

is equivalent to

$$\max_{\mathbf{\Lambda}}\left[\mathbf{\Lambda}_1'\hat{\mathbf{\Sigma}}_{1\mathbf{x}}\left(\theta\right)\mathbf{\Lambda}_1+\mathbf{\Lambda}_2'\hat{\mathbf{\Sigma}}_{2\mathbf{x}}\left(\theta\right)\mathbf{\Lambda}_2\right], \tag{6}$$

where

$$\hat{\mathbf{\Sigma}}_{j\mathbf{x}}\left(\theta\right)=\left[\left(NT\right)^{-1}\sum_{t=1}^{T}\mathbb{I}_{jt}\left(\theta\right)\mathbf{x}_t\mathbf{x}_t'\right],\quad j=1,2: \tag{7}$$

for $j=1,2$, and for given $\theta$, the estimator for $\mathbf{\Lambda}_j^0$ solving the problem in (6) is $\hat{\mathbf{\Lambda}}_j\left(\theta\right)$, where $\hat{\mathbf{\Lambda}}_j\left(\theta\right)$ is equal to $\sqrt{N}$ times the $N\times R^0$ matrix of eigenvectors of $\hat{\mathbf{\Sigma}}_{j\mathbf{x}}\left(\theta\right)$ corresponding to its largest $R^0$ eigenvalues. Replacing $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ in (4) with $\hat{\mathbf{\Lambda}}_1\left(\theta\right)$ and $\hat{\mathbf{\Lambda}}_2\left(\theta\right)$ leads to the concentrated sum of squared residuals (divided by $NT$)

$$S_{\mathbf{F\Lambda}}\left(\theta\right)=\left(NT\right)^{-1}\sum_{t=1}^{T}\mathbf{x}_t'\left\{\mathbf{I}_N-N^{-1}\left[\mathbb{I}_{1t}\left(\theta\right)\hat{\mathbf{\Lambda}}_1\left(\theta\right)\hat{\mathbf{\Lambda}}_1\left(\theta\right)'+\mathbb{I}_{2t}\left(\theta\right)\hat{\mathbf{\Lambda}}_2\left(\theta\right)\hat{\mathbf{\Lambda}}_2\left(\theta\right)'\right]\right\}\mathbf{x}_t: \tag{8}$$

the estimator for $\theta^0$ then solves

$$\hat{\theta}=\arg\min_{\theta} S_{\mathbf{F\Lambda}}\left(\theta\right).$$

Given $\hat{\theta}$, the estimator for $\mathbf{\Lambda}_j^0$ is $\hat{\mathbf{\Lambda}}_j=\hat{\mathbf{\Lambda}}_j\left(\hat{\theta}\right)$, for $j=1,2$. Finally, given $\hat{\theta}$ and $\hat{\mathbf{\Lambda}}=\left(\hat{\mathbf{\Lambda}}_1,\hat{\mathbf{\Lambda}}_2\right)$, from (3)

$$\hat{\mathbf{f}}_t=\hat{\mathbf{f}}_t\left(\hat{\mathbf{\Lambda}},\hat{\theta}\right)=N^{-1}\left[\mathbb{I}_{1t}\left(\hat{\theta}\right)\hat{\mathbf{\Lambda}}_1+\mathbb{I}_{2t}\left(\hat{\theta}\right)\hat{\mathbf{\Lambda}}_2\right]'\mathbf{x}_t,\quad t=1,\ldots,T.$$

### 3.4 Consistency

From Theorem 3.1 the two regimes described in (1) are separately identified under Assumption 1. Define the $R^0\times T$ matrices of regime-specific factors $\mathbf{F}_j^0\left(\theta\right)=\left[\mathbf{f}_{j1}^0\left(\theta\right),\ldots,\mathbf{f}_{jT}^0\left(\theta\right)\right]$, for $j=1,2$, such that $\mathbf{F}_1^0\left(\theta\right)+\mathbf{F}_2^0\left(\theta\right)=\left(\mathbf{f}_1^0,\ldots,\mathbf{f}_T^0\right)=\mathbf{F}^0$ for any $\theta$, and $\mathbf{F}_1^0\left(\theta^0\right)\mathbf{F}_2^0\left(\theta^0\right)'=\mathbf{0}_{R^0}$. Let $\hat{\mathbf{H}}_{jj}\left(\theta\right)$ and $\hat{\mathbf{H}}_{jm}\left(\theta\right)$ be the rotation matrices

$$\hat{\mathbf{H}}_{jj}\left(\theta\right)=\frac{\mathbf{F}_j^0\left(\theta^0\right)\mathbf{F}_j^0\left(\theta\right)'}{T}\frac{\mathbf{\Lambda}_j^{0\prime}\hat{\mathbf{\Lambda}}_j\left(\theta\right)}{N}\hat{\mathbf{V}}_j\left(\theta\right)^{-1},\quad j=1,2, \tag{9}$$

$$\hat{\mathbf{H}}_{mj}\left(\theta\right)=\frac{\mathbf{F}_m^0\left(\theta^0\right)\mathbf{F}_j^0\left(\theta\right)'}{T}\frac{\mathbf{\Lambda}_m^{0\prime}\hat{\mathbf{\Lambda}}_j\left(\theta\right)}{N}\hat{\mathbf{V}}_j\left(\theta\right)^{-1},\quad j,m=1,2,\quad j\neq m, \tag{10}$$

where $\hat{\mathbf{V}}_j(\theta)$ is the $R^0 \times R^0$ diagonal matrix of the first $R^0$ largest eigenvalues of $\hat{\mathbf{\Sigma}}_{j\mathbf{x}}(\theta)$ defined in (7) in decreasing order: for $\theta = \theta^0$ notice that $\hat{\mathbf{H}}_{jj}(\theta)$ and $\hat{\mathbf{H}}_{mj}(\theta)$ reduce to

$$\hat{\mathbf{H}}_{jj}\left(\theta^0\right) = \frac{\mathbf{F}_j^0\left(\theta^0\right)\mathbf{F}_j^0\left(\theta^0\right)'}{T}\frac{\mathbf{\Lambda}_j^{0\prime}\hat{\mathbf{\Lambda}}_j\left(\theta^0\right)}{N}\hat{\mathbf{V}}_j\left(\theta^0\right)^{-1}, \quad \hat{\mathbf{H}}_{mj}\left(\theta^0\right) = \mathbf{0}_{R^0} \quad j,m = 1,2, \quad j \neq m,$$

and $\hat{\mathbf{H}}_{jj}\left(\theta^0\right)$ becomes a regime-specific rotation matrix analogous to the one derived in Bai and Ng (2002) for linear factor models[3]. The following theorem shows the bias of the principal components estimator induced by the presence of regimes when $\theta \neq \theta^0$.

**Theorem 3.2** *There exist $R^0 \times R^0$ matrices $\hat{\mathbf{H}}_{jj}(\theta)$ and $\hat{\mathbf{H}}_{mj}(\theta)$ as defined in* (9) *and* (10)*, respectively, with* $\text{rank}\left[\hat{\mathbf{H}}_{jj}(\theta)\right] = R^0$ *for all $\theta$, and* $\text{rank}\left[\hat{\mathbf{H}}_{mj}(\theta)\right] = R^0$ *for $\theta \neq \theta^0$, and $C_{NT} = \min\left\{\sqrt{N}, \sqrt{T}\right\}$, such that*

$$C_{NT}^2\left[\frac{1}{N}\sum_{i=1}^N\left\|\hat{\boldsymbol{\lambda}}_{ji}(\theta) - \hat{\mathbf{H}}_{jj}(\theta)'\boldsymbol{\lambda}_{ji}^0 - \hat{\mathbf{H}}_{mj}(\theta)'\boldsymbol{\lambda}_{mi}^0\right\|^2\right] = O_p(1), \quad \forall\theta, \quad j,m = 1,2, \quad j \neq m.$$

Theorem 3.2 shows that the presence of regimes adds the asymptotic bias $\hat{\mathbf{H}}_{mj}(\theta)'\boldsymbol{\lambda}_{mi}^0$ to the principal components estimator $\hat{\boldsymbol{\lambda}}_{ji}(\theta)$ for the space $\hat{\mathbf{H}}_{jj}(\theta)'\boldsymbol{\lambda}_{ji}^0$ spanned by $\boldsymbol{\lambda}_{ji}^0$. As in linear factor models, the rate of convergence is equal to $C_{NT}^2 = \min\{N, T\}$ and therefore depends on the panel structure. Taking into account (10), it follows that for $\theta = \theta^0$,

$$C_{NT}^2\left[\frac{1}{N}\sum_{i=1}^N\left\|\hat{\boldsymbol{\lambda}}_{ji}\left(\theta^0\right) - \hat{\mathbf{H}}_{jj}\left(\theta^0\right)'\boldsymbol{\lambda}_{ji}^0\right\|^2\right] = O_p(1), \quad j = 1,2, \tag{11}$$

which extends the result in Theorem 1 in Bai and Ng (2002) to accommodate the presence of regimes when the threshold $\theta^0$ is known.

Theorem 3.2 plays a key role in proving the following theorem, which states the consistency of $\hat{\theta}$ as an estimator for $\theta^0$.

**Theorem 3.3** *Under Assumptions I and C1-C4, $\hat{\theta} \xrightarrow{p} \theta^0$ as $N, T \to \infty$.*

Theorems 3.2 and 3.3 imply a number of results analogous to those collected in Theorem 1 in Stock and Watson (2002): these are stated in Corollary 3.1 below.

[3]See Bai and Ng (2002), p. 213.

**Corollary 3.1** *For $j = 1, 2$, and under Assumptions I and C1-C4, as $N, T \to \infty$:*

**(a)** $\hat{\boldsymbol{\lambda}}_{ji}\left(\hat{\theta}\right) \xrightarrow{p} \hat{\mathbf{H}}_{jj}\left(\theta^0\right)' \boldsymbol{\lambda}_{ji}^0$;

**(b)** $\hat{\mathbf{f}}_t \xrightarrow{p} \left[\mathbb{I}_{1t}\left(\theta^0\right)\hat{\mathbf{H}}_{11}\left(\theta^0\right)^{-1} + \mathbb{I}_{2t}\left(\theta^0\right)\hat{\mathbf{H}}_{22}\left(\theta^0\right)^{-1}\right]\mathbf{f}_t^0$;

**(c)** $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{ji}\left(\hat{\theta}\right) - \hat{\mathbf{H}}_{jj}\left(\theta^0\right)' \boldsymbol{\lambda}_{ji}^0\right\|^2 \xrightarrow{p} 0$;

**(d)** $\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{\mathbf{f}}_t - \left[\mathbb{I}_{1t}\left(\theta^0\right)\hat{\mathbf{H}}_{11}\left(\theta^0\right)^{-1} + \mathbb{I}_{2t}\left(\theta^0\right)\hat{\mathbf{H}}_{22}\left(\theta^0\right)^{-1}\right]\mathbf{f}_t^0\right\|^2 \xrightarrow{p} 0.$

## 3.5 Convergence Rates

The following theorem states the convergence rates of the concentrated least squares estimator for the threshold $\theta^0$ and of the principal components estimator for the loadings.

**Theorem 3.4** *Under Assumptions I, C1-C4 and CR,*

$$N^{\left(2\alpha^0-1\right)}T\left(\hat{\theta} - \theta^0\right) = O_p(1)$$

*and*

$$C_{NT}^2\left[\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{ji}\left(\hat{\theta}\right) - \hat{\mathbf{H}}_{jj}\left(\theta^0\right)' \boldsymbol{\lambda}_{ji}^0\right\|^2\right] = O_p(1), \quad j = 1, 2.$$

Theorem 3.4 states the superconsistency of $\hat{\theta}$ as an estimator for $\theta^0$: it extends to an infinite dimensional system the result in Chan (1993) seminal contribution. The convergence rate $N^{\left(2\alpha^0-1\right)}T$ of $\hat{\theta}$ depends on the time series dimension $T$ *and* the number of cross-sectional units $N^{\alpha^0}$ subject to threshold effect: the rate $N^{\left(2\alpha^0-1\right)}T$ monotonically increases in $\alpha^0$; since $0.5 < \alpha^0 \leq 1$ by Assumption I, then $T < N^{\left(2\alpha^0-1\right)}T \leq NT$; $N^{\left(2\alpha^0-1\right)}T$ is unknown since $\alpha^0$ is unknown. The higher $\alpha^0$, the stronger identification of (1) from a linear factor model, and the faster the convergence rate of $\hat{\theta}$ to $\theta^0$: this shows the connection between identification and estimation. When $\alpha^0 = 1$, all cross-sectional units are subject to threshold effect and the convergence rate is $NT$. Theorem 3.4 implies that the principal components estimator for the loadings has the same convergence rate derived in Bai and Ng (2002) in the case of linear factor models: the estimator for the threshold therefore does not affect the estimator of the loadings. Corollary 3.2 below follows from Theorem 3.4.

**Corollary 3.2** *Under Assumptions I, C1-C4 and CR,*

$$C_{NT}^2 \left[ \frac{1}{T} \sum_{t=1}^{T} \left\| \hat{\mathbf{f}}_t - \left[ \mathbb{I}_{1t}\left(\theta^0\right) \hat{\mathbf{H}}_{11}\left(\theta^0\right)^{-1} + \mathbb{I}_{2t}\left(\theta^0\right) \hat{\mathbf{H}}_{22}\left(\theta^0\right)^{-1} \right] \mathbf{f}_t^0 \right\|^2 \right] = O_p\left(1\right).$$

Corollary 3.2 shows that the convergence rate $C_{NT}$ also applies to the principal components estimator for the factors; it also shows that the rotation induced by $\hat{\mathbf{f}}_t$ around $\mathbf{f}_t^0$ depends upon the regime. Corollary 3.2 justifies the robust Bai and Ng (2002) information criteria proposed in Section 4.

# 4 Determining the Number of Factors

We now consider the case in which the true number of factors $R^0$ in (1) (i.e., the true dimension of $\mathbf{f}_t^0$) no longer is known and has to be determined. Breitung and Eickmeier (2011) show that neglecting structural breaks in the factor loadings inflates the estimated number of factors. Given the analogy between factor models with structural instability and (1), the latter suffers from the same problem. We rely on Corollary 3.2 and suggest a simple way to robustify Bai and Ng (2002) selection criteria to account for the threshold effect.

Given (1) and for fixed number of factors $R$, the loss function in (2) generalizes to

$$S\left(\mathbf{\Lambda}^R, \mathbf{F}^R, \theta\right) = \left(NT\right)^{-1} \sum_{t=1}^{T} \left[\mathbf{x}_t - \mathbb{I}_{1t}\left(\theta\right) \mathbf{\Lambda}_1^R \mathbf{f}_t^R - \mathbb{I}_{2t}\left(\theta\right) \mathbf{\Lambda}_2^R \mathbf{f}_t^R\right]' \left[\mathbf{x}_t - \mathbb{I}_{1t}\left(\theta\right) \mathbf{\Lambda}_1^R \mathbf{f}_t^R - \mathbb{I}_{2t}\left(\theta\right) \mathbf{\Lambda}_2^R \mathbf{f}_t^R\right], \tag{12}$$

where $\mathbf{\Lambda}^R = \left(\mathbf{\Lambda}_1^R, \mathbf{\Lambda}_2^R\right)$, $\mathbf{F}^R = \left(\mathbf{f}_1^R, \ldots, \mathbf{f}_T^R\right)$, and where the superscript $R$ denotes the dependence on the number of factors. The loss function in (12) depends on $\theta$. From Theorem 3.4, it easily follows that for any *a priori* chosen number of factors $R = \bar{R}$ such that $\bar{R} \geq R^0$, the estimator $\hat{\theta}^{\bar{R}}$ for $\theta^0$ is such that $N^{\left(2\alpha^0-1\right)} T \left(\hat{\theta}^{\bar{R}} - \theta^0\right) = O_p\left(1\right)$, with $\hat{\theta}^{R_0} = \hat{\theta}$ (see Lemma A.9 in Appendix A.3): in practice, $\bar{R}$ may be chosen as discussed below. Given the convergence rate in Corollary 3.2, this naturally suggests generalizing Bai and Ng (2002) criteria by first setting $\theta = \hat{\theta}^{\bar{R}}$ in (12) to then select $\hat{R}$ factor within each mutually exclusive regime, and therefore $\left(\hat{R} + \hat{R}\right)$ factors in total.

Let $\hat{\mathbf{\Lambda}}^R\left(\theta\right)$ and $\hat{\mathbf{F}}^R\left(\theta\right)$ be the estimators for $\mathbf{\Lambda}^R$ and $\mathbf{F}^R$, respectively, for any $\theta$. Given the loss function in (12), and following Bai and Ng (2002), we want penalty functions $g\left(N, T\right)$ to obtain criteria of the form

$$PC\left(R,R\right)=S\left[\hat{\mathbf{\Lambda}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\mathbf{F}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\theta}^{\bar{R}}\right]+\left(R+R\right)\cdot g\left(N,T\right),$$

which consistently estimate the number of factors $R^0$ in each regime and therefore $\left(R^0+R^0\right)$ factors in total: the criterion $PC\left(R,R\right)$ accounts for the fact that the threshold effect leads to a factor representation with a higher dimensional factor space, namely to a representation with $\left(R^0+R^0\right)$ factors. Given a bounded integer $R^{\max}\geq R^0$, the true numbers of factors $R^0$ is estimated as

$$\hat{R}=\arg\min_{1\leq R\leq R^{\max}}PC\left(R,R\right):$$

given the convergence rate $C_{NT}$ in Corollary 3.2, this leads to the threshold effect robust Bai and Ng (2002) information criteria

$$IC_{p1}\left(R,R\right)=\ln S\left[\hat{\mathbf{\Lambda}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\mathbf{F}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\theta}^{\bar{R}}\right]+\left(R+R\right)\left(\frac{N+T}{NT}\right)\ln\left(\frac{NT}{N+T}\right),$$

$$IC_{p2}\left(R,R\right)=\ln S\left[\hat{\mathbf{\Lambda}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\mathbf{F}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\theta}^{\bar{R}}\right]+\left(R+R\right)\left(\frac{N+T}{NT}\right)\ln\left(C_{NT}^{2}\right),\tag{13}$$

$$IC_{p3}\left(R,R\right)=\ln S\left[\hat{\mathbf{\Lambda}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\mathbf{F}}^{R}\left(\hat{\theta}^{\bar{R}}\right),\hat{\theta}^{\bar{R}}\right]+\left(R+R\right)\left[\frac{\ln\left(C_{NT}^{2}\right)}{C_{NT}^{2}}\right].$$

In practice, to obtain the estimator $\hat{\theta}^{\bar{R}}$ for $\theta^0$, we may set $\bar{R}=R^{\max}$. The following theorem states the validity of the proposed information criteria.

**Theorem 4.1** *Under Assumptions I, C1-C4 and CR, the criteria* $IC_{p1}\left(R,R\right)$*,* $IC_{p2}\left(R,R\right)$ *and* $IC_{p3}\left(R,R\right)$ *defined in* (13) *consistently estimate the number of factors* $R^0$*.*

# 5 Monte Carlo Analysis

We now assess the empirical validity of the theoretical results, which relate to the consistency and the convergence rates of the estimators, and to the model selection criteria: the experiments are described in Sections 5.1 and 5.2, respectively, and the results are discussed in Section 5.3.

## 5.1 Consistency and Convergence Rates

In line with the results in Section 3, we assume a known number of factors. As in Breitung and Eickmeier (2011), we analyze the case of a one-factor model. From (1) we simulate data using the Data Generating Process (DGP)

$$x_{it}^{s} = \mathbb{I}\left(z_t^s \leq \theta^0\right)\lambda_{1i}^0 f_t^{0s} + \mathbb{I}\left(z_t^s > \theta^0\right)\lambda_{2i}^0 f_t^{0s} + e_{it}^s, \quad i = 1,\ldots,N, \quad t = 1,\ldots,T,$$

where $s = 1,\ldots,S$ refers to the replication, $N$ and $T$ are the cross-sectional and time series dimensions, respectively, and $S$ is the total number of replications. We run the experiment in Ox 7.01 (see Doornik, 2012). We set $S = 2000$, $N = 25, 50, 100$ and $T = 100, 200, 400$. In generating the data, we define $\delta_i^0 = \lambda_{2i}^0 - \lambda_{1i}^0$: we set $\delta_i^0 > 0$ for $i = 1,\ldots,\left[N^{\alpha^0}\right]$ and $\delta_i^0 = 0$ for $i = \left[N^{\alpha^0}\right] + 1,\ldots,N$, where $[\cdot]$ denotes the integer part of the argument. We fix the factor loadings $\lambda_{1i}^0$ and $\lambda_{2i}^0$ and the threshold parameter $\theta^0$ throughout the replications, with $\lambda_{1i}^0 \sim \mathcal{N}(1,1)$ for $i = 1,\ldots,N$ as in the Monte Carlo experiment in Breitung and Eickmeier (2011), and $\theta^0 = 2$. We control for: $(i)$ the number of cross-sectional units $\left[N^{\alpha^0}\right]$ subject to a regime change, with $0.5 < \alpha^0 \leq 1$; $(ii)$ the magnitude of the threshold effect $\delta_i^0$; and $(iii)$ the proportion of observations of $z_t^s$ below $\theta^0$.

We generate the factor $f_t^{0s}$ as

$$f_t^{0s} = \rho_f f_{t-1}^{0s} + \left(1 - \rho_f^2\right)^{1/2} \epsilon_{ft}^s, \quad f_{-50}^{0s} = 0, \quad \epsilon_{ft}^s \sim \mathrm{IID}\mathcal{N}(0,1), \quad t = -49,\ldots,0,\ldots,T, \tag{14}$$

with $\rho_f$ fixed in repeated samples and drawn as $\rho_f \sim \mathcal{U}(0.05, 0.95)$, so that $\mathrm{E}\left(f_t^{0s}\right) = 0$ and $\mathrm{Var}\left(f_t^{0s}\right) = 1$. We generate $z_t^s$ as

$$z_t^s = \mu_z(1-\rho_z) + \rho_z z_{t-1}^s + \left(1-\rho_z^2\right)^{1/2}\epsilon_{zt}^s, \quad z_{-50}^s = \mu_z, \quad \epsilon_{zt}^s \sim \mathrm{IID}\mathcal{N}(0,1), \quad t = -49,\ldots,0,\ldots,T, \tag{15}$$

where $\mu_z$ and $\rho_z$ are fixed in repeated samples, with $\rho_z \sim \mathcal{U}(0.05, 0.95)$ and $\mu_z$ calibrated as described below, so that $\mathrm{E}(z_t^s) = \mu_z$ and $\mathrm{Var}(z_t^s) = 1$. We generate the idiosyncratic components $e_{it}^s$ as

$$e_{it}^s = \rho_e e_{i,t-1}^s + \sigma_{ii}^{1/2}\left(1-\rho_e^2\right)^{1/2}\epsilon_{e_i t}^s, \quad e_{i,-50}^s = 0, \quad \epsilon_{e_i t}^s \sim \mathrm{IID}\mathcal{N}(0,1), \quad i = 1,\ldots,N, \quad t = -49,\ldots,0,\ldots,T, \tag{16}$$

where $\rho_e$ and $\sigma_{ii}$ are fixed in repeated samples with $\rho_e \sim \mathcal{U}(0.05, 0.95)$ and $\sigma_{ii} \sim \chi(1)$, for $i = 1, \ldots, N$: in this way $\mathrm{Var}(e_{it}^s) = \sigma_{ii}$ for $i = 1, \ldots, N$, and $N^{-1} \sum_{i=1}^N \sigma_{ii} \to 1$ as $N \to \infty$.

We control for the number of cross-sectional units $\left[N^{\alpha^0}\right]$ subject to a regime change by setting $\alpha^0 = 0.60, 1.00$. As for the threshold effect, we set $\delta_i^0 = 0.25, 1.00, 1.75$ for $i = 1, \ldots, \left[N^{\alpha^0}\right]$. Finally, we let $\pi^0 = \mathrm{P}\left(z_t^s \leq \theta^0\right)$ and obtain the expression for $\mu_z$ in closed form as a function of $\pi^0$ as

$$\pi^0 = \mathrm{P}\left(z_t^s \leq \theta^0\right) = \mathrm{P}\left(z_t^s - \mu_z \leq \theta^0 - \mu_z\right) = \mathbf{\Phi}\left(\theta^0 - \mu_z\right) \Leftrightarrow \mu_z = \theta^0 - \mathbf{\Phi}^{-1}\left(\pi^0\right):$$

we consider five values for $\pi^0$, namely $\pi^0 = 0.15, 0.30, 0.50, 0.70, 0.85$.

To reduce the effect induced by the initial values $f_{-50}^{0s} = 0$, $z_{-50}^s = \mu_z$ and $e_{i,-50}^s = 0$, we discard the first 50 observations in the DGPs for $f_t^{0s}$, $z_t^s$ and $e_{it}^s$. We estimate factor and factor loadings as detailed in Section 3.3. Given the convergence rates derived in Theorem 3.4, the estimator for $\theta^0$ is asymptotically independent of that for $\lambda_{1i}^0$, $\lambda_{2i}^0$ and $f_t^{0s}$. As in Tong and Lim (1980), Tsay (1989) and Kapetanios (2000), we estimate $\theta^0$ by grid search: we implement the algorithm by selecting 19 equally spaced quantiles of the empirical distribution function of $z_t^s$, namely $\{5\%, 10\%, 15\%, \ldots, 85\%, 90\%, 95\%\}$, and the true value $\theta^0 = 2$. Given the concentrated least squares estimator $\hat{\theta}^s$ for $\theta^0$, we estimate factor and loadings by principal components. We assess the performance of $\hat{\theta}^s$ by computing

$$\text{bias} = S^{-1} \sum_{s=1}^{S} \left(\hat{\theta}^s - \theta^0\right), \quad \text{RMSE} = \sqrt{S^{-1} \sum_{s=1}^{S} \left(\hat{\theta}^s - \theta^0\right)^2}.$$

Finally, given the estimator $\hat{c}_{it}^s = \mathbb{I}\left(z_t^s \leq \hat{\theta}^s\right) \hat{\lambda}_{1i}^s \hat{f}_t^s + \mathbb{I}\left(z_t^s > \hat{\theta}^s\right) \hat{\lambda}_{2i}^s \hat{f}_t^s$ for the generally defined common component $c_{it}^{0s} = \mathbb{I}\left(z_t^s \leq \theta^0\right) \lambda_{1i}^0 f_t^{0s} + \mathbb{I}\left(z_t^s > \theta^0\right) \lambda_{2i}^0 f_t^{0s}$, we report

$$\text{MSE} = S^{-1} \sum_{s=1}^{S} \left[(NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\hat{c}_{it}^s - c_{it}^{0s}\right)^2\right].$$

## 5.2 Model Selection

We simulate the data using the DGP

$$x_{it}^s = \mathbb{I}\left(z_t^s \leq \theta^0\right)\left(\lambda_{11i}^0 f_{1t}^{0s} + \lambda_{12i}^0 f_{2t}^{0s}\right) + \mathbb{I}\left(z_t^s > \theta^0\right)\left(\lambda_{21i}^0 f_{1t}^{0s} + \lambda_{22i}^0 f_{2t}^{0s}\right) + e_{it}^s, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T,$$

with $\lambda_{11i}^0 \sim \mathcal{N}(1,1)$, $\lambda_{12i}^0 \sim \mathcal{N}(1,1)$, $\lambda_{21i}^0 = \lambda_{11i}^0 + \delta_i^0$ and $\lambda_{22i}^0 = \lambda_{12i}^0 + \delta_i^0$. We set $\delta_i^0 = 0.25, 1.00, 1.75$ for $i = 1, \ldots, [N^{\alpha_0}]$, and $\delta_i^0 = 0$ for $i = [N^{\alpha_0}] + 1, \ldots, N$, with $\alpha_0 = 0.60$. The factors $f_{1t}^{0s}$ and $f_{2t}^{0s}$ are generated as an AR(1) process analogous to (14); $z_t^s$ and $e_{it}^s$ are as in (15) and (16), respectively. The model has $R^0 = 2$ factors and it is estimated with $R^{\max} = 8$. We assess the model selection criteria in 13 by reporting the average number of estimated factors over the 2000 replications.

### 5.3 Results

The results from the Monte Carlo analysis are collected in three tables: Tables 1 and 2 focus upon consistency and convergence rates of the estimators; model selection criteria are assessed in Table 3.

Table 1 about here

Table 2 about here

Table 3 about here

Table 1 displays results for the concentrated least squares estimator $\hat{\theta}$ for $\theta^0 = 2$ when $\alpha^0 = 0.60$ (Panel A) and $\alpha^0 = 1.00$ (Panel B). Given Theorems 3.1 and 3.4, a higher $\alpha^0$ leads to stronger identification of $\theta^0$ and faster convergence rate of $\hat{\theta}$ to $\theta^0$, respectively: in line with these theoretical results, the Monte Carlo outcomes show that the RMSE of $\hat{\theta}$ when $\alpha^0 = 1.00$ generally is lower than the homologous value when $\alpha^0 = 0.60$. The RMSE tends to decrease with $N$, $T$ and $\delta_i^0 > 0$. The RMSE generally is minimized at $\pi^0 = 0.50$ and monotonically increases as $\pi^0$ tends to 0.15 and 0.85: this behavior is mitigated the higher $N$, $T$, $\delta_i^0 > 0$ and $\alpha^0$ are, as $\theta^0$ is more strongly identified and more precisely estimated and the frequency of the regimes plays a less important role. The bias displays a pattern somehow similar to that of the RMSE.

Table 2 shows the MSE of the common components when $\alpha^0 = 0.60$ (Panels A) and $\alpha^0 = 1.00$ (Panels B). We assess the empirical validity of Theorem 3.4 by considering both unfeasible and feasible estimators, the former and the latter being obtained by setting $\theta = \theta^0$ and $\theta = \hat{\theta}$, respectively. In line with Theorem 3.4, the MSE of the feasible estimator converges to that of the unfeasible counterpart as both $N$ and $T$ increase. The MSE monotonically decreases in $N$ and $T$, and in $\delta_i^0 > 0$ for $N = 25$, whereas it does not exhibit any systematically noticeable difference between $\alpha^0 = 0.60$ and $\alpha^0 = 1.00$.

Table 3 collects results for the model selection criteria $IC_{p1}(R,R)$, $IC_{p2}(R,R)$ and $IC_{p3}(R,R)$ (Panels A, B and C, respectively) proposed in (13) when $\alpha^0 = 0.60$. The criteria $IC_{p1}(R,R)$ and $IC_{p2}(R,R)$ display a similar behavior, with the latter having a hedge over the former: they tend to overestimate the number of factors for the small cross-sectional dimensions $N = 25, 50$, whereas they perform well for $N = 100$. Conversely, $IC_{p3}(R,R)$ has inferior empirical performance than $IC_{p1}(R,R)$ and $IC_{p2}(R,R)$. Finally, unfeasible and feasible estimators give similar results in terms of model selection performance.

In conclusion, the Monte Carlo findings corroborate the theoretical results stated in Theorems 3.1 and 3.4. They also confirm the validity of the information criteria in (13), with $IC_{p1}(R,R)$ and $IC_{p2}(R,R)$ having better finite sample properties than $IC_{p3}(R,R)$.

# 6 Empirical Application

This section provides an application to illustrate the potential usefulness of our contribution for applied work. We show how our framework may be used to measure connectedness in multivariate nonlinear dynamic systems, with a focus on financial variables: a threshold factor specification is suitable when "history repeats", as in the case of financial markets, which undergo regime shifts (Timmermann (2008), and Ang and Timmermann (2012)). In what follows, Section 6.1 proposes a measure of connectedness, Section 6.2 describes the data and the empirical model, and Section 6.3 presents the results.

## 6.1 Measure of Connectedness

Connectedness is central to risk measurement and management. There exist several measures of connectedness, which are based on different underlying metrics: examples are the marginal expected shortfall of Acharya *et al.* (2010), the CoVaR of Adrian and Brunnermeier (2011), the equicorrelation approach of Engle and Kelly (2012), and the network approach of Diebold and Yilmaz (2014). In line with our methodological contribution, we focus on the principal components approach of Billio *et al.* (2012).

Given the sequence of $N \times 1$ vectors $\{\mathbf{x}_t\}_{t=1}^T$, let $\{\omega_r\}_{r=1}^N$ be the sequence of eigenvalues of the $N \times N$ covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}} = (NT)^{-1} \sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t'$. In relation to financial markets, Billio *et al.* (2012) quantify the degree of connectedness amongst the elements of $\mathbf{x}_t$ as the risk associated to the first $R$ eigenvalues in relation to the overall risk of the system. Formally, they measure connectedness through[4]

$$C(R) = \frac{\sum_{r=1}^{R} \omega_r}{\sum_{r=1}^{N} \omega_r} :$$

by construction $C(R)$ is increasing in $R$; for given $R$, a higher $C(R)$ denotes a higher degree of connectedness amongst the underlying variables. The measure $C(R)$ is a powerful tool to capture the degree of connectedness amongst random variables. However, it suffers from two main drawbacks. First, the number of eigenvalues $R$ is chosen *a priori* and not according to a selection criterion. Second, $C(R)$ refers to the entire time series dimension $T$ and is unable to detect variations in connectedness induced by a threshold effect. Financial markets experience regimes shifts (Timmermann (2008), and Ang and Timmermann (2012)): the measure $C(R)$ may not provide an accurate description of the dynamics in connectedness of the variables of interest[5]. Our methodology allows us to build a measure of connectedness that accommodates a regime shift and that relies on the optimally selected number of eigenvalues.

Let $\{\omega_{jr}\}_{r=1}^{N}$ be the sequence of eigenvalues of the $N \times N$ covariance matrix $\hat{\boldsymbol{\Sigma}}_{j\mathbf{x}}(\theta)$ defined in (7) in decreasing order, for $j = 1, 2$. We generalize $C(R)$ and measure connectedness through

$$C_j\left(\hat{R}\right) = \frac{\sum_{r=1}^{\hat{R}} \omega_{jr}}{\sum_{r=1}^{N} \omega_{jr}}, \quad j = 1, 2. \tag{17}$$

Compared to $C(R)$, the measure $C_j\left(\hat{R}\right)$ has two distinctive features: it quantifies connectedness within each regime; and the number of eigenvalues $\hat{R}$ is optimally determined according to the criteria in (13).

## 6.2 Data and Model Specification

We construct the vector of dependent variables from the updated monthly financial dataset employed in Jurado *et al.* (2015) and, on a quarterly frequency, in Ludvigson and Ng (2007)[6]: this consists of a panel of 147 series related to the U.S. financial markets, as detailed in Ludvigson and Ng (2007). As customary in the literature, all variables are standardized so to have zero mean and unit variance.

The choice of the threshold variable is a function of the research question. We investigate how

[4] Billio *et al.* (2012) refer to $C(R)$ as to the Cumulative Risk Fraction.

[5] Billio *et al.* (2012) measure the dynamic degree of connectedness in financial returns by computing $C(R)$ over rolling windows.

[6] I am very grateful to Sydney Ludvigson for providing me with the updated version of the dataset I am using in the paper.

economic policy uncertainty affects connectedness amongst financial variables. As a threshold variable we use the lagged index of economic policy uncertainty proposed in Baker *et al.* (2013)[7]: a higher value of the index denotes a higher level of uncertainty.

Due to data availability issues, we perform the empirical analysis over the period running from January 1985 to December 2014, a total of 360 observations. The threshold variable has mean, standard deviation, maximum and minimum equal to 107.640, 32.566, 245.127 and 57.203, respectively.

We consider $R^{\max} = 10$ and estimate the change-point by setting $\bar{R} = R^{\max}$; we then construct a grid for the change-point with lowest and highest values equal to 5% and 95%, respectively, and step equal to 0.5%. The number of factors are selected according to the criteria in (13).

### 6.3 Results

Results are collected in Table 4. The point estimate of the threshold $\theta^0$ is $\hat{\theta} = 131.413$: this splits the sample into a low and a high economic policy uncertainty regimes, with frequencies equal to $\hat{\pi} = 0.783$ and $1 - \hat{\pi} = 0.217$, respectively. The criteria $IC_{p1}(R, R)$ and $IC_{p2}(R, R)$ select $\hat{R} = 3$ factors, with corresponding connectedness measures $C_1\left(\hat{R}\right) = 0.678$ and $C_2\left(\hat{R}\right) = 0.865$. Conversely, $IC_{p3}(R, R)$ selects $\hat{R}_1 = 6$ factors: this is consistent with the results from the Monte Carlo analysis discussed in Section 5.3, which show that $IC_{p3}(R, R)$ tends to overestimate the number of factors in finite samples. Our empirical results therefore show that connectedness amongst financial variables increases with economic policy uncertainty. This result is likely to be relevant for risk measurement and management.

## 7 Conclusions

We study least squares estimation of large dimensional factor models subject to a regime shift in the loadings induced by the threshold effect. Our methodology is appealing as it allows to handle the general situation in which the threshold parameter is unknown and has to be estimated. We show that the concentrated least squares estimator of the threshold value is superconsistent: the convergence rate depends on the time series dimension and on the number of cross-sectional units subject to threshold effect. The principal components estimators for the factors and the loadings have the same convergence rate they have in linear factor models: this result allows to robustify Bai and Ng (2002) model selection

[7]The index is made available at http://www.policyuncertainty.com/ .

criteria by simply accounting for the fact that the threshold effect leads to a factor representation with a higher dimensional factor space. In an application, our procedure allows to document an increase in connectedness amongst financial variables during periods of high economic policy uncertainty: this result is likely to be relevant for risk measurement and management.

This work may be extended along several directions. We see as a priority the application of our methodology to large dimensional factor models subject to structural instability: this is the focus of current research.

# A Proofs of Theorems

## A.1 Proofs of Results in Section 3.4

We rely on the following lemmas.

**Lemma A.1** *Under Assumptions I and C1-C3, there exists some positive constant $M < \infty$ such that for all $\theta$, all $(N, T)$ and $j = 1, 2$:*

**(a)** $N^{-1} \sum_{i=1}^{N} \sum_{l=1}^{N} \sigma_{jil}^2 (\theta) \leq M$;

**(b)** $\mathrm{E}\left\{ N^{-2} \sum_{i=1}^{N} \sum_{l=1}^{N} \left[ T^{-1} \sum_{t=1}^{T} \mathbb{I}_{jt}(\theta) x_{it} x_{lt} \right]^2 \right\} \leq M$;

**(c)** $\mathrm{E}\left[ N^{-1} \sum_{i=1}^{N} \left\| T^{-1/2} \sum_{t=1}^{T} \mathbb{I}_{jt}(\theta) e_{it} \boldsymbol{\lambda}_{ji}^0 \right\|^2 \right] \leq M$.

**Lemma A.2** *Given $\hat{\mathbf{H}}_{jj}(\theta)$ and $\hat{\mathbf{H}}_{mj}(\theta)$ defined in* (9) *and* (10)*, respectively, for $j = 1, 2$, and $j \neq m$, and for any $\theta$,*

$$S_{\mathbf{F}\boldsymbol{\Lambda}}(\theta) - S_{\mathbf{F}}\left[ \boldsymbol{\Lambda}_1^0 \hat{\mathbf{H}}_{11}(\theta) + \boldsymbol{\Lambda}_2^0 \hat{\mathbf{H}}_{21}(\theta), \boldsymbol{\Lambda}_2^0 \hat{\mathbf{H}}_{22}(\theta) + \boldsymbol{\Lambda}_1^0 \hat{\mathbf{H}}_{12}(\theta), \theta \right] = O_p\left( C_{NT}^{-1} \right).$$

**Lemma A.3** *There exists a $\tau(\theta) > 0$ such that*

$$\operatorname{p\lim} \inf_{N,T \to \infty} S_{\mathbf{F}}\left[ \boldsymbol{\Lambda}_1^0 \hat{\mathbf{H}}_{11}(\theta) + \boldsymbol{\Lambda}_2^0 \hat{\mathbf{H}}_{21}(\theta), \boldsymbol{\Lambda}_2^0 \hat{\mathbf{H}}_{22}(\theta) + \boldsymbol{\Lambda}_1^0 \hat{\mathbf{H}}_{12}(\theta), \theta \right] - S_{\mathbf{F}}\left( \boldsymbol{\Lambda}_1^0, \boldsymbol{\Lambda}_2^0, \theta^0 \right) = \tau(\theta), \quad \forall \theta \neq \theta^0.$$

**Proof of Theorem 3.1.** As defined in Section 3.2, $\tilde{\mathbf{V}}_1$ is the $R^0 \times R^0$ diagonal matrix of the first $R^0$ largest eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}} = (NT)^{-1} \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'$ in decreasing order, and $\tilde{\boldsymbol{\Lambda}}_1$ is the estimator for $\boldsymbol{\Lambda}_1^0$ in the true data generating process $\mathbf{x}_t = \boldsymbol{\Lambda}_1^0 \mathbf{f}_t^0 + \mathbb{I}_{2t}(\theta^0) \boldsymbol{\Delta}^0 \mathbf{f}_t^0 + \mathbf{e}_t$ from the misspecified model $\mathbf{x}_t = \boldsymbol{\Lambda}_1 \mathbf{f}_t + \mathbf{e}_t$: the equality $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}} \tilde{\boldsymbol{\Lambda}}_1 = \tilde{\boldsymbol{\Lambda}}_1 \tilde{\mathbf{V}}_1$ then holds by the definitions of eigenvectors and eigenvalues. Applying the normalisation $N^{-1} \tilde{\boldsymbol{\Lambda}}_1' \tilde{\boldsymbol{\Lambda}}_1 = \mathbf{I}_{R^0}$ to implement the principal components estimator, it follows that $N^{-1} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_{1i} \right\|^2 = O_p(1)$. By Lemma A.3 in Bai (2003), $\tilde{\mathbf{V}}_1 \xrightarrow{p} \mathbf{V}_1$ where $\mathbf{V}_1$ is a positive definite matrix: we then focus on $\left\| \tilde{\mathbf{V}}_1 \left( \tilde{\boldsymbol{\lambda}}_{1i} - \tilde{\mathbf{H}}_1' \boldsymbol{\lambda}_{1i}^0 \right) \right\|^2$. Theorem 3.1 relies on the identity

$$\begin{aligned} \tilde{\mathbf{V}}_1 \left( \tilde{\boldsymbol{\lambda}}_{1i} - \tilde{\mathbf{H}}_1' \boldsymbol{\lambda}_{1i}^0 \right) &= N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \sigma_{1il}(\theta^0) + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \varkappa_{1il}(\theta^0) \\ &\quad + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \sigma_{2il}(\theta^0) + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \varkappa_{2il}(\theta^0) \\ &\quad + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \varphi_{il} + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \varphi_{li} + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \vartheta_{il} + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \vartheta_{li} + N^{-1} \sum_{l=1}^{N} \tilde{\boldsymbol{\lambda}}_{1l} \psi_{il}, \end{aligned}$$

where

$$
\begin{gathered}
\varkappa_{jil}(\theta) = T^{-1}\sum_{t=1}^{T}\mathbb{I}_{jt}(\theta)\, e_{it}e_{lt} - \sigma_{jil}(\theta), \quad j=1,2, \\
\varphi_{jil}(\theta) = T^{-1}\sum_{t=1}^{T}\mathbb{I}_{jt}(\theta)\left[\mathbb{I}_{1t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1i}^{0\prime}\mathbf{f}_t^{0} + \mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{2i}^{0\prime}\mathbf{f}_t^{0}\right] e_{lt}, \quad j=1,2, \\
\varphi_{il} = \varphi_{1il}(\theta) + \varphi_{2il}(\theta) = T^{-1}\sum_{t=1}^{T}\left[\mathbb{I}_{1t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1i}^{0\prime}\mathbf{f}_t^{0} + \mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{2i}^{0\prime}\mathbf{f}_t^{0}\right] e_{lt}, \\
\varphi_{jli}(\theta) = T^{-1}\sum_{t=1}^{T}\mathbb{I}_{jt}(\theta)\left[\mathbb{I}_{1t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1l}^{0\prime}\mathbf{f}_t^{0} + \mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{2l}^{0}\mathbf{f}_t^{0}\right]' e_{it}, \quad j=1,2, \\
\varphi_{li} = \varphi_{1li}(\theta) + \varphi_{2li}(\theta) = T^{-1}\sum_{t=1}^{T}\left[\mathbb{I}_{1t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1l}^{0\prime}\mathbf{f}_t^{0} + \mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{2l}^{0}\mathbf{f}_t^{0}\right]' e_{it}, \\
\vartheta_{il} = T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1i}^{0\prime}\mathbf{f}_t^{0}\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^{0}, \quad \vartheta_{li} = T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1l}^{0\prime}\mathbf{f}_t^{0}\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_i^{0}, \\
\psi_{il} = T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^{0}\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^{0}.
\end{gathered} \tag{18}
$$

The matrix $\tilde{\mathbf{H}}_1$ depends on $N$ and $T$: this dependence is implicitly suppressed to keep notation simple. Notice that

$$
\left\|\tilde{\mathbf{H}}_1\right\| \leq \left\|\frac{\mathbf{F}^{0}\mathbf{F}^{0\prime}}{T}\right\| \left\|\frac{\boldsymbol{\Lambda}_1^{0\prime}\boldsymbol{\Lambda}_1^{0}}{N}\right\|^{1/2} \left\|\frac{\tilde{\boldsymbol{\Lambda}}_1'\tilde{\boldsymbol{\Lambda}}_1}{N}\right\|^{1/2} \left\|\tilde{\mathbf{V}}_1^{-1}\right\| = O_p(1),
$$

by Assumptions C1 and C2. By Loève's inequality,

$$
N^{-1}\sum_{i=1}^{N}\left\|\tilde{\mathbf{V}}_1\left(\tilde{\boldsymbol{\lambda}}_{1i} - \tilde{\mathbf{H}}_1'\boldsymbol{\lambda}_{1i}^{0}\right)\right\|^2 \leq 9N^{-1}\sum_{i=1}^{N}\left[\tilde{\sigma}_{1i\cdot}(\theta) + \tilde{\varkappa}_{1i\cdot}(\theta) + \tilde{\sigma}_{2i\cdot}(\theta) + \tilde{\varkappa}_{2i\cdot}(\theta) + \tilde{\varphi}_{i\cdot} + \tilde{\varphi}_{\cdot i} + \tilde{\vartheta}_{i\cdot} + \tilde{\vartheta}_{\cdot i} + \tilde{\psi}_{i\cdot}\right],
$$

where

$$
\begin{gathered}
\tilde{\sigma}_{ji\cdot}(\theta) = N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\sigma_{jil}(\theta)\right\|^2, \quad \tilde{\varkappa}_{ji\cdot}(\theta) = N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\varkappa_{jil}(\theta)\right\|^2, \quad j=1,2, \\
\tilde{\varphi}_{i\cdot} = N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\varphi_{il}\right\|^2, \quad \tilde{\varphi}_{\cdot i} = N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\varphi_{li}\right\|^2, \\
\tilde{\vartheta}_{i\cdot} = N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\vartheta_{il}\right\|^2, \quad \tilde{\vartheta}_{\cdot i} = N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\vartheta_{li}\right\|^2, \\
\tilde{\psi}_{i\cdot} = N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\psi_{il}\right\|^2.
\end{gathered}
$$

We first consider $\tilde{\sigma}_{1i\cdot}(\theta)$: $\tilde{\sigma}_{2i\cdot}(\theta)$ is analogous and omitted. We have

$$
\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\sigma_{1il}(\theta)\right\|^2 \leq \left(\sum_{l=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_{1l}\right\|^2\right)\left[\sum_{l=1}^{N}\sigma_{1il}^2(\theta)\right]
$$

and

$$
N^{-1}\sum_{i=1}^{N}\tilde{\sigma}_{1i\cdot}(\theta) \leq N^{-1}\left(N^{-1}\sum_{l=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_{1l}\right\|^2\right)\left[N^{-1}\sum_{i=1}^{N}\sum_{l=1}^{N}\sigma_{1il}^2(\theta)\right] = O_p\left(N^{-1}\right)
$$

by Lemma A.1(a). As for $\tilde{\varkappa}_{ji\cdot}(\theta)$, for $j=1$ ($j=2$ is analogous),

$$
\begin{aligned}
\sum_{i=1}^{N}\tilde{\varkappa}_{1i\cdot}(\theta) &= N^{-2}\sum_{i=1}^{N}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\varkappa_{1il}(\theta)\right\|^2 \\
&= N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\sum_{q=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}'\tilde{\boldsymbol{\lambda}}_{1q}\varkappa_{1il}(\theta)\varkappa_{1iq}(\theta) \\
&\leq \left[N^{-2}\sum_{l=1}^{N}\sum_{q=1}^{N}\left(\tilde{\boldsymbol{\lambda}}_{1l}'\tilde{\boldsymbol{\lambda}}_{1q}\right)^2\right]^{1/2}\left\{N^{-2}\sum_{l=1}^{N}\sum_{q=1}^{N}\left[\sum_{i=1}^{N}\varkappa_{1il}(\theta)\varkappa_{1iq}(\theta)\right]^2\right\}^{1/2} \\
&\leq \left(N^{-1}\sum_{l=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_{1l}\right\|^2\right)\left\{N^{-2}\sum_{l=1}^{N}\sum_{q=1}^{N}\left[\sum_{i=1}^{N}\varkappa_{1il}(\theta)\varkappa_{1iq}(\theta)\right]^2\right\}^{1/2}.
\end{aligned}
$$

Since

$$
\mathrm{E}\left\{\left[\sum_{i=1}^{N}\varkappa_{1il}(\theta)\varkappa_{1iq}(\theta)\right]^2\right\} = \mathrm{E}\left[\sum_{i=1}^{N}\sum_{u=1}^{N}\varkappa_{1il}(\theta)\varkappa_{1iq}(\theta)\varkappa_{1ul}(\theta)\varkappa_{1uq}(\theta)\right]^2 \leq N^2\max_{i,l}\mathrm{E}\left|\varkappa_{1il}(\theta)\right|^4
$$

and

$$\mathrm{E}\left|\varkappa_{1il}\left(\theta\right)\right|^{4}=T^{-2}\mathrm{E}\left|T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta\right)e_{it}e_{lt}-\mathrm{E}\left[\mathbb{I}_{1t}\left(\theta\right)e_{it}e_{lt}\right]\right|^{4}\leq T^{-2}M$$

by Assumption C3(d), then

$$\sum_{i=1}^{N}\tilde{\varkappa}_{1i\cdot}\left(\theta\right)\leq O_{p}\left(1\right)\sqrt{\frac{N^{2}}{T^{2}}}=O_{p}\left(\frac{N}{T}\right)$$

and $N^{-1}\sum_{i=1}^{N}\tilde{\varkappa}_{1i\cdot}\left(\theta\right)=O_{p}\left(T^{-1}\right)$. Regarding $\tilde{\varphi}_{i\cdot}$, we have

$$\begin{aligned}
\tilde{\varphi}_{i\cdot} &= N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\varphi_{il}\right\|^{2}\\
&= N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\left\{T^{-1}\sum_{t=1}^{T}\left[\mathbb{I}_{1t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1i}^{0\prime}\mathbf{f}_{t}^{0}+\mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{2i}^{0\prime}\mathbf{f}_{t}^{0}\right]e_{lt}\right\}\right\|^{2}\\
&\leq \left\{N^{-2}\sum_{l=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right]\right\|^{2}\right\}+\left\{N^{-2}\sum_{l=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{2i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right]\right\|^{2}\right\}\\
&\leq \left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{1i}^{0}\right\|^{2}\left(N^{-1}\sum_{l=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_{1l}\right\|^{2}\right)\\
&\quad+\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{2i}^{0}\right\|^{2}\left(N^{-1}\sum_{l=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_{1l}\right\|^{2}\right)\\
&\leq \left\{\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{1i}^{0}\right\|^{2}+\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{2i}^{0}\right\|^{2}\right\}O_{p}\left(1\right)
\end{aligned}$$

and

$$N^{-1}\sum_{i=1}^{N}\tilde{\varphi}_{i\cdot}=\left\{\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}O_{p}\left(1\right)+\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}O_{p}\left(1\right)\right\}O_{p}\left(1\right)=O_{p}\left(T^{-1}\right)$$

by Assumptions C2 and C4. In a similar way, it is proved that $\tilde{\varphi}_{\cdot i}=O_{p}\left(T^{-1}\right)$. As for $\tilde{\vartheta}_{i\cdot}$, under Assumption I,

$$\begin{aligned}
\tilde{\vartheta}_{i\cdot} &= N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\vartheta_{il}\right\|^{2}\\
&= N^{-2}\left(\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\vartheta_{il}\right)^{\prime}\left(\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\vartheta_{il}\right)\\
&= N^{-1}\left\{\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}\mathbf{f}_{t}^{0\prime}\boldsymbol{\delta}_{l}^{0}\right]\right\}^{\prime}N^{-1}\left\{\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}\mathbf{f}_{t}^{0\prime}\boldsymbol{\delta}_{l}^{0}\right]\right\}\\
&= N^{-1}\left\{\sum_{l=1}^{N^{\alpha^{0}}}\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}\mathbf{f}_{t}^{0\prime}\boldsymbol{\delta}_{l}^{0}\right]+\sum_{l=N^{\alpha^{0}}+1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}\mathbf{f}_{t}^{0\prime}\boldsymbol{\delta}_{l}^{0}\right]\right\}^{\prime}\\
&\quad\times N^{-1}\left\{\sum_{l=1}^{N^{\alpha^{0}}}\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}\mathbf{f}_{t}^{0\prime}\boldsymbol{\delta}_{l}^{0}\right]+\sum_{l=N^{\alpha^{0}}+1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}\mathbf{f}_{t}^{0\prime}\boldsymbol{\delta}_{l}^{0}\right]\right\}\\
&= N^{-1}O_{p}\left(N^{\alpha^{0}}\right)\cdot N^{-1}O_{p}\left(N^{\alpha^{0}}\right)\\
&= O_{p}\left(N^{2\alpha^{0}-2}\right).
\end{aligned}$$

In a similar way, it can be proved that $\tilde{\vartheta}_{\cdot i} = O_p\left(N^{2\alpha^0-2}\right)$ under Assumption I. Finally, under Assumptions I, C1 and C2,

$$\begin{aligned}
\tilde{\psi}_{i\cdot} &= N^{-2}\left\|\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\psi_{il}\right\|^2 \\
&= N^{-2}\left(\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\psi_{il}\right)'\left(\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\psi_{il}\right) \\
&= N^{-1}\left\{\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^0\right]\right\}' N^{-1}\left\{\sum_{l=1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^0\right]\right\} \\
&= N^{-1}\left\{\sum_{l=1}^{N^{\alpha^0}}\tilde{\boldsymbol{\lambda}}_{1l}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^0\right]+\sum_{l=N^{\alpha^0}+1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^0\right]\right\}' \\
&\quad\times N^{-1}\left\{\sum_{l=1}^{N^{\alpha^0}}\tilde{\boldsymbol{\lambda}}_{1l}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^0\right]+\sum_{l=N^{\alpha^0}+1}^{N}\tilde{\boldsymbol{\lambda}}_{1l}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\boldsymbol{\delta}_l^0\right]\right\} \\
&= N^{-1}O_p\left(N^{\alpha^0}\right)N^{-1}O_p\left(N^{\alpha^0}\right) \\
&= O_p\left(N^{2\alpha^0-2}\right).
\end{aligned}$$

Combining all above results, we have

$$N^{-1}\sum_{i=1}^{N}\left\|\tilde{\mathbf{V}}_1\left(\tilde{\boldsymbol{\lambda}}_{1i}-\tilde{\mathbf{H}}_1'\boldsymbol{\lambda}_{1i}^0\right)\right\|^2 = O_p\left(N^{-1}\right)+O_p\left(T^{-1}\right)+O_p\left(N^{2\alpha^0-2}\right),$$

which completes the proof of the theorem. ■

**Proof of Theorem 3.2.** As shown in Theorem 3.1, by Assumption I the regime indicator $\mathbb{I}_{jt}(\theta)$ is identified, for $j=1,2$: we can then split the sample according to the value of $\mathbb{I}_{jt}(\theta)$. We consider the case $j=1$: the case $j=2$ is analogous and omitted. As defined in Section 3.4, $\hat{\mathbf{V}}_1(\theta)$ is the $R^0\times R^0$ diagonal matrix of the first $R^0$ largest eigenvalues of $\hat{\boldsymbol{\Sigma}}_{1\mathbf{x}}(\theta)$ in (7) in decreasing order: the equality $\hat{\boldsymbol{\Sigma}}_{1\mathbf{x}}(\theta)\hat{\boldsymbol{\Lambda}}_1(\theta)=\hat{\boldsymbol{\Lambda}}_1(\theta)\hat{\mathbf{V}}_1(\theta)$ holds by the definitions of eigenvectors and eigenvalues. From the normalisation $N^{-1}\hat{\boldsymbol{\Lambda}}_1(\theta)'\hat{\boldsymbol{\Lambda}}_1(\theta)=\mathbf{I}_{R^0}$, it follows that $N^{-1}\sum_{i=1}^N\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)\right\|^2=O_p(1)$ for all $\theta$. By Lemma A.3 in Bai (2003), $\hat{\mathbf{V}}_1(\theta)\xrightarrow{p}\mathbf{V}_1(\theta)$ where $\mathbf{V}_1(\theta)$ is a positive definite matrix for all $\theta$, and $\left\|\hat{\mathbf{V}}_1(\theta)\right\|=O_p(1)$: we then focus on $\left\|\hat{\mathbf{V}}_1(\theta)\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]\right\|^2$. Theorem 3.2 relies on the identity

$$\begin{aligned}
\hat{\mathbf{V}}_1(\theta)\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right] &= N^{-1}\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\sigma_{1il}(\theta)+N^{-1}\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\varkappa_{1il}(\theta) \\
&\quad+N^{-1}\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\varphi_{1il}(\theta)+N^{-1}\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\varphi_{1li}(\theta),
\end{aligned}$$

where $\varkappa_{1il}(\theta)$, $\varphi_{1il}(\theta)$ and $\varphi_{1li}(\theta)$ are defined in (18). The matrices $\hat{\mathbf{H}}_{11}(\theta)$ and $\hat{\mathbf{H}}_{21}(\theta)$ both depend on $N$ and $T$: this dependence is implicitly suppressed to keep notation simple. Notice that

$$\left\|\hat{\mathbf{H}}_{11}(\theta)\right\|\leq\left\|\frac{\mathbf{F}_1^0\left(\theta^0\right)\mathbf{F}_1^0(\theta)'}{T}\right\|\left\|\frac{\boldsymbol{\Lambda}_1^{0\prime}\boldsymbol{\Lambda}_1^0}{N}\right\|^{1/2}\left\|\frac{\hat{\boldsymbol{\Lambda}}_1(\theta)'\hat{\boldsymbol{\Lambda}}_1(\theta)}{N}\right\|^{1/2}\left\|\hat{\mathbf{V}}_1(\theta)^{-1}\right\|=O_p(1)$$

by Assumptions C1 and C2. In an analogous way, it can be shown that $\left\|\hat{\mathbf{H}}_{21}(\theta)\right\|=O_p(1)$. By Loève's inequality

$$N^{-1}\sum_{i=1}^{N}\left\|\hat{\mathbf{V}}_1(\theta)\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]\right\|^2\leq 4N^{-1}\sum_{i=1}^{N}\left[\hat{\sigma}_{1i\cdot}(\theta)+\hat{\varkappa}_{1i\cdot}(\theta)+\hat{\varphi}_{1i\cdot}(\theta)+\hat{\varphi}_{1\cdot i}(\theta)\right],$$

where

$$\begin{aligned}
\hat{\sigma}_{1i\cdot}(\theta) &= N^{-2}\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\sigma_{1il}(\theta)\right\|^2, &\qquad \hat{\varkappa}_{1i\cdot}(\theta) &= N^{-2}\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\varkappa_{1il}(\theta)\right\|^2, \\
\hat{\varphi}_{1i\cdot}(\theta) &= N^{-2}\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\varphi_{1il}(\theta)\right\|^2, &\qquad \hat{\varphi}_{1\cdot i}(\theta) &= N^{-2}\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\varphi_{1li}(\theta)\right\|^2.
\end{aligned}$$

Starting from $\hat{\sigma}_{1i\cdot}(\theta)$,

$$\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\,\sigma_{1il}(\theta)\right\|^{2}\leq\left[\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1l}(\theta)\right\|^{2}\right]\left[\sum_{l=1}^{N}\sigma_{1il}^{2}(\theta)\right],$$

and

$$N^{-1}\sum_{i=1}^{N}\hat{\sigma}_{1i\cdot}(\theta)\leq N^{-1}\left[N^{-1}\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1l}(\theta)\right\|^{2}\right]\left[N^{-1}\sum_{i=1}^{N}\sum_{l=1}^{N}\sigma_{1il}^{2}(\theta)\right]=O_{p}\left(N^{-1}\right)$$

by Lemma A.1(a). As for $\hat{\varkappa}_{1i\cdot}(\theta)$,

$$\begin{aligned}
\sum_{i=1}^{N}\hat{\varkappa}_{1i\cdot}(\theta) &= N^{-2}\sum_{i=1}^{N}\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\,\varkappa_{1il}(\theta)\right\|^{2}\\
&= N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\sum_{q=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)'\,\hat{\boldsymbol{\lambda}}_{1q}(\theta)\,\varkappa_{1il}(\theta)\,\varkappa_{1iq}(\theta)\\
&\leq \left\{N^{-2}\sum_{l=1}^{N}\sum_{q=1}^{N}\left[\hat{\boldsymbol{\lambda}}_{1l}(\theta)'\,\hat{\boldsymbol{\lambda}}_{1q}(\theta)\right]^{2}\right\}^{1/2}\left\{N^{-2}\sum_{l=1}^{N}\sum_{q=1}^{N}\left[\sum_{i=1}^{N}\varkappa_{1il}(\theta)\,\varkappa_{1iq}(\theta)\right]^{2}\right\}^{1/2}\\
&\leq \left[N^{-1}\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1l}(\theta)\right\|^{2}\right]\left\{N^{-2}\sum_{l=1}^{N}\sum_{q=1}^{N}\left[\sum_{i=1}^{N}\varkappa_{1il}(\theta)\,\varkappa_{1iq}(\theta)\right]^{2}\right\}^{1/2}.
\end{aligned}$$

Since

$$\mathrm{E}\left\{\left[\sum_{i=1}^{N}\varkappa_{1il}(\theta)\,\varkappa_{1iq}(\theta)\right]^{2}\right\}=\mathrm{E}\left[\sum_{i=1}^{N}\sum_{u=1}^{N}\varkappa_{1il}(\theta)\,\varkappa_{1iq}(\theta)\,\varkappa_{1ul}(\theta)\,\varkappa_{1uq}(\theta)\right]\leq N^{2}\max_{i,l}\mathrm{E}\left|\varkappa_{1il}(\theta)\right|^{4}$$

and

$$\mathrm{E}\left|\varkappa_{1il}(\theta)\right|^{4}=T^{-2}\mathrm{E}\left|T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,e_{it}e_{lt}-\mathrm{E}\left[\mathbb{I}_{1t}(\theta)\,e_{it}e_{lt}\right]\right|^{4}\leq T^{-2}M$$

by Assumption C3(d), then

$$\sum_{i=1}^{N}\hat{\varkappa}_{1i\cdot}(\theta)\leq O_{p}(1)\sqrt{\frac{N^{2}}{T^{2}}}=O_{p}\left(\frac{N}{T}\right)$$

and $N^{-1}\sum_{i=1}^{N}\hat{\varkappa}_{1i\cdot}(\theta)=O_{p}\left(T^{-1}\right)$. Regarding $\hat{\varphi}_{1i\cdot}(\theta)$,

$$\begin{aligned}
\hat{\varphi}_{1i\cdot}(\theta) &= N^{-2}\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\,\varphi_{1il}(\theta)\right\|^{2}\\
&= N^{-2}\left\|\sum_{l=1}^{N}\hat{\boldsymbol{\lambda}}_{1l}(\theta)\left\{T^{-1}\sum_{t=1}^{T}\left\{\mathbb{I}_{1t}(\theta)\left[\mathbb{I}_{1t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{1i}^{0\prime}\mathbf{f}_{t}^{0}+\mathbb{I}_{2t}\left(\theta^{0}\right)\boldsymbol{\lambda}_{2i}^{0\prime}\mathbf{f}_{t}^{0}\right]e_{lt}\right\}\right\}\right\|^{2}\\
&\leq N^{-2}\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1l}(\theta)\,\boldsymbol{\lambda}_{1i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{1t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right]\right\|^{2}+N^{-2}\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1l}(\theta)\,\boldsymbol{\lambda}_{2i}^{0\prime}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{2t}\left(\theta^{0}\right)\mathbf{f}_{t}^{0}e_{lt}\right]\right\|^{2}\\
&\leq \left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{1i}^{0}\right\|^{2}\left(N^{-1}\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1l}(\theta)\right\|^{2}\right)\\
&\quad+\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{2i}^{0}\right\|^{2}\left(N^{-1}\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1l}(\theta)\right\|^{2}\right)\\
&\leq \left\{\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{1i}^{0}\right\|^{2}+\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}\left\|\boldsymbol{\lambda}_{2i}^{0}\right\|^{2}\right\}O_{p}(1)
\end{aligned}$$

and

$$N^{-1}\sum_{i=1}^{N}\hat{\varphi}_{1i\cdot}(\theta)=\left\{\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}O_{p}(1)+\left\{N^{-1}\sum_{l=1}^{N}\left[T^{-2}\left\|\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbf{f}_{t}^{0}e_{lt}\right\|^{2}\right]\right\}O_{p}(1)\right\}O_{p}(1)=O_{p}\left(T^{-1}\right)$$

by Assumptions C2 and C4. In an analogous way, it can be proved that

$$N^{-1}\sum_{i=1}^{N}\hat{\varphi}_{1\cdot i}(\theta)=O_{p}\left(T^{-1}\right).$$

Combining all results above, we have

$$N^{-1}\sum_{i=1}^{N}\left\|\hat{\mathbf{V}}_1(\theta)\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]\right\|^2=O_p\left(N^{-1}\right)+O_p\left(T^{-1}\right).$$

This completes the proof of the theorem. ■

**Proof of Theorem 3.3.** In order to prove the theorem, it is sufficient to prove that

$$\lim_{N,T\to\infty}\mathrm{P}\left[S_{\mathbf{F}\boldsymbol{\Lambda}}(\theta)\le S_{\mathbf{F}\boldsymbol{\Lambda}}\left(\theta^0\right)\right]=0,\quad\forall\theta\neq\theta^0,$$

where $S_{\mathbf{F}\boldsymbol{\Lambda}}(\theta)$ is defined in (8). Consider the identity

$$\begin{aligned}S_{\mathbf{F}\boldsymbol{\Lambda}}(\theta)-S_{\mathbf{F}\boldsymbol{\Lambda}}\left(\theta^0\right)&=S_{\mathbf{F}\boldsymbol{\Lambda}}(\theta)-S_{\mathbf{F}}\left[\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta),\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}(\theta)+\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{12}(\theta),\theta\right]\\&\quad+S_{\mathbf{F}}\left[\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta),\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}(\theta)+\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{12}(\theta),\theta\right]-S_{\mathbf{F}}\left[\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}\left(\theta^0\right),\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}\left(\theta^0\right),\theta^0\right]\\&\quad+S_{\mathbf{F}}\left[\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}\left(\theta^0\right),\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}\left(\theta^0\right),\theta^0\right]-S_{\mathbf{F}\boldsymbol{\Lambda}}\left(\theta^0\right),\end{aligned}$$

where $S_{\mathbf{F}}(\boldsymbol{\Lambda},\theta)$ is defined in (4). By Lemma A.2, $S_{\mathbf{F}\boldsymbol{\Lambda}}(\theta)-S_{\mathbf{F}}\left[\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta),\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}(\theta)+\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{12}(\theta),\theta\right]=O_p\left(C_{NT}^{-1}\right)$ for any $\theta$, including $\theta=\theta^0$. Since $\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}\left(\theta^0\right)$ and $\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}\left(\theta^0\right)$ span the same column space as $\boldsymbol{\Lambda}_1^0$ and $\boldsymbol{\Lambda}_2^0$, respectievly, we have

$$S_{\mathbf{F}}\left[\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}\left(\theta^0\right),\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}\left(\theta^0\right),\theta^0\right]=S_{\mathbf{F}}\left(\boldsymbol{\Lambda}_1^0,\boldsymbol{\Lambda}_2^0,\theta^0\right)$$

and $S_{\mathbf{F}}\left[\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta),\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{22}(\theta)+\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{12}(\theta),\theta\right]-S_{\mathbf{F}}\left(\boldsymbol{\Lambda}_1^0,\boldsymbol{\Lambda}_2^0,\theta^0\right)$ has a positive limit by Lemma A.3. This completes the proof of the theorem. ■

**Proof of Corollary 3.1.** Corollary 3.1 easily follows from Theorem 3.3 and the proof is omitted. ■

**Proof of Lemma A.1.** Consider $j=1$ ($j=2$ is analogous and omitted). As for (a), let $\rho_{1il}(\theta)=\sigma_{1il}(\theta)\Big/\left[\sigma_{1ii}(\theta)\sigma_{1ll}(\theta)\right]^{1/2}$ such that $|\rho_{1il}(\theta)|\le 1$: since $|\sigma_{1ll}(\theta)|\le M$ for all $l$ by Assumption C3(c), then

$$\begin{aligned}N^{-1}\sum_{i=1}^{N}\sum_{l=1}^{N}\sigma_{1il}^2(\theta)&=N^{-1}\sum_{i=1}^{N}\sum_{l=1}^{N}\sigma_{1ii}(\theta)\sigma_{1ll}(\theta)\rho_{1il}^2(\theta)\\&\le MN^{-1}\sum_{i=1}^{N}\sum_{l=1}^{N}\left|\sigma_{1ii}(\theta)\sigma_{1ll}(\theta)\right|^{1/2}\left|\rho_{1il}(\theta)\right|\\&=MN^{-1}\sum_{i=1}^{N}\sum_{l=1}^{N}\left|\sigma_{1il}(\theta)\right|\le M^2\end{aligned}$$

by Assumption C3(c). In order to prove (b), for $j=1$ (the proof for $j=2$ is analogous) it is sufficient to prove that $\mathrm{E}\left|\mathbb{I}_{1t}(\theta)x_{it}\right|^4\le M$ for all $(\theta,i,t)$: we then have

$$\begin{aligned}\mathrm{E}\left|\mathbb{I}_{1t}(\theta)x_{it}\right|^4&=\mathrm{E}\left|\mathbb{I}_{1t}(\theta)\left[\mathbb{I}_{1t}\left(\theta^0\right)\boldsymbol{\lambda}_{1i}^{0\prime}\mathbf{f}_t^0+\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\lambda}_{2i}^{0\prime}\mathbf{f}_t^0+e_{it}\right]\right|^4\\&\le\mathrm{E}\left\{\left[\mathbb{I}_{1t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\boldsymbol{\lambda}_{1i}^{0\prime}\mathbf{f}_t^0\right]^4\right\}+\mathrm{E}\left\{\left[\mathbb{I}_{1t}(\theta)\mathbb{I}_{2t}\left(\theta^0\right)\boldsymbol{\lambda}_{2i}^{0\prime}\mathbf{f}_t^0\right]^4\right\}+\mathrm{E}\left|\mathbb{I}_{1t}(\theta)e_{it}\right|^4\\&\le\bar{\lambda}^4\mathrm{E}\left\|\mathbb{I}_{1t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0\right\|^4+\bar{\lambda}^4\mathrm{E}\left\|\mathbb{I}_{1t}(\theta)\mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}_t^0\right\|^4+\mathrm{E}\left|\mathbb{I}_{1t}(\theta)e_{it}\right|^4\\&\le M\end{aligned}$$

by Assumptions C1, C2 and C3(a). As for (c), set $j=1$ (the proof for $j=2$ is analogous and omitted) and consider

$$\begin{aligned}\mathrm{E}\left\|T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)e_{it}\boldsymbol{\lambda}_{1i}^0\right\|^2&=T^{-1}\sum_{t=1}^{T}\sum_{v=1}^{T}\mathrm{E}\left[\mathbb{I}_{1t}(\theta)\mathbb{I}_{1v}(\theta)e_{it}e_{iv}\right]\boldsymbol{\lambda}_{1i}^{0\prime}\boldsymbol{\lambda}_{1i}^0\\&\le\bar{\lambda}^2T^{-1}\sum_{t=1}^{T}\sum_{v=1}^{T}\left|\tau_{jtv}\right|\le\bar{\lambda}^2M\end{aligned}$$

by Assumptions C2 and C3(b). ■

**Proof of Lemma A.2.** Given $\hat{\mathbf{\Lambda}}_j(\theta)$ defined in (5), for $j=1,2$, define

$$\begin{aligned} &\mathbf{P}_{\hat{\mathbf{\Lambda}}_j}(\theta) = \hat{\mathbf{\Lambda}}_j(\theta)\left[\hat{\mathbf{\Lambda}}_j(\theta)'\hat{\mathbf{\Lambda}}_j(\theta)\right]^{-1}\hat{\mathbf{\Lambda}}_j(\theta)', \\ &\mathbf{P}_{\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}}(\theta) = \left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right] \\ &\qquad \times\left\{\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]'\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]\right\}^{-1}, \qquad j,m=1,2,\; j\neq m, \\ &\qquad \times\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]' \end{aligned} \tag{19}$$

so that

$$S_{\mathbf{F\Lambda}}(\theta) = (NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\left\{\mathbf{I}_N - \left[\mathbb{I}_{1t}(\theta)\mathbf{P}_{\hat{\mathbf{\Lambda}}_1}(\theta)+\mathbb{I}_{2t}(\theta)\mathbf{P}_{\hat{\mathbf{\Lambda}}_2}(\theta)\right]\right\}\mathbf{x}_t$$

and

$$\begin{aligned} &S_{\mathbf{F}}\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta), \mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}(\theta)+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}(\theta), \theta\right] \\ = &(NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\left\{\mathbf{I}_N - \left[\mathbb{I}_{1t}(\theta)\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)+\mathbb{I}_{2t}(\theta)\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\right]\right\}\mathbf{x}_t, \end{aligned}$$

where $S_{\mathbf{F}}(\mathbf{\Lambda},\theta)$ and $S_{\mathbf{F\Lambda}}(\theta)$ are defined in (4) and (8), respectively: it follows that

$$\begin{aligned} &S_{\mathbf{F\Lambda}}(\theta) - S_{\mathbf{F}}\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta), \mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}(\theta)+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}(\theta), \theta\right] \\ = &(NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\mathbb{I}_{1t}(\theta)\left[\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)-\mathbf{P}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]\mathbf{x}_t + (NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\mathbb{I}_{2t}(\theta)\left[\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)-\mathbf{P}_{\hat{\mathbf{\Lambda}}_2}(\theta)\right]\mathbf{x}_t. \end{aligned}$$

Let

$$\begin{aligned} &\mathbf{D}_{\hat{\mathbf{\Lambda}}_j}(\theta) = N^{-1}\hat{\mathbf{\Lambda}}_j(\theta)'\hat{\mathbf{\Lambda}}_j(\theta), \\ &\mathbf{D}_{\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}}(\theta) = N^{-1}\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]'\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right], \qquad j=1,2, \end{aligned}$$

so that for $j=1,2$ and $j\neq m$,

$$\begin{aligned} \mathbf{P}_{\hat{\mathbf{\Lambda}}_j}(\theta) - \mathbf{P}_{\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}}(\theta) = \; & N^{-1}\hat{\mathbf{\Lambda}}_j(\theta)\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_j}(\theta)\right]^{-1}\hat{\mathbf{\Lambda}}_j(\theta)' \\ & -N^{-1}\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]\left[\mathbf{D}_{\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}}(\theta)\right]^{-1}\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]' \\ = \; & N^{-1}\left[\hat{\mathbf{\Lambda}}_j(\theta)-\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)-\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_j}(\theta)\right]^{-1}\left[\hat{\mathbf{\Lambda}}_j(\theta)-\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)-\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]' \\ & +N^{-1}\left[\hat{\mathbf{\Lambda}}_j(\theta)-\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)-\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_j}(\theta)\right]^{-1}\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]' \\ & +N^{-1}\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_j}(\theta)\right]^{-1}\left[\hat{\mathbf{\Lambda}}_j(\theta)-\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)-\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]' \\ & +N^{-1}\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]\left\{\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_j}(\theta)\right]^{-1}-\left[\mathbf{D}_{\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}}(\theta)\right]^{-1}\right\} \\ & \times\left[\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}(\theta)+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}(\theta)\right]'. \end{aligned}$$

We consider the case $j=1$: the case $j=2$ is analogous and omitted. We have

$$\begin{aligned} &(NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\mathbb{I}_{1t}(\theta)\left[\mathbf{P}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]\mathbf{x}_t \\ = &(NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\mathbb{I}_{1t}(\theta)N^{-1}\left[\hat{\mathbf{\Lambda}}_1(\theta)-\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)-\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}\left[\hat{\mathbf{\Lambda}}_1(\theta)-\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)-\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]'\mathbf{x}_t \\ &+(NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\mathbb{I}_{1t}(\theta)N^{-1}\left[\hat{\mathbf{\Lambda}}_1(\theta)-\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)-\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]'\mathbf{x}_t \\ &+(NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\mathbb{I}_{1t}(\theta)N^{-1}\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}\left[\hat{\mathbf{\Lambda}}_1(\theta)-\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)-\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]'\mathbf{x}_t \\ &+(NT)^{-1}\sum_{t=1}^T \mathbf{x}_t'\mathbb{I}_{1t}(\theta)N^{-1}\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]\left\{\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\}\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]'\mathbf{x}_t \\ = &a_1(\theta)+a_2(\theta)+a_3(\theta)+a_4(\theta). \end{aligned}$$

Starting from $a_1(\theta)$,

$$\begin{aligned}
a_1(\theta) &= N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left\{\begin{array}{l}\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right]'\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}\left[\hat{\boldsymbol{\lambda}}_{1l}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1l}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2l}^{0}\right]\\ \times\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)x_{it}x_{lt}\right]\end{array}\right\}\\
&\le \left\{N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left\{\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right]'\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}\left[\hat{\boldsymbol{\lambda}}_{1l}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1l}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2l}^{0}\right]\right\}^2\right\}^{1/2}\\
&\quad\times\left\{N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)x_{it}x_{lt}\right]^2\right\}^{1/2}\\
&\le \left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2\right]\left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}\right\|O_p(1)\\
&= \left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2\right]O_p(1)
\end{aligned}$$

by Lemma A.1(b) and the fact that $\left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}\right\| = O_p(1)$, which is proved below: from Theorem 3.2 it follows that $a_1(\theta) = O_p\left(C_{NT}^{-2}\right)$ for all $\theta$. As for $a_2(\theta)$,

$$\begin{aligned}
a_2(\theta) &= N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left\{\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right]'\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}\left[\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1l}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2l}^{0}\right]\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)x_{it}x_{lt}\right]\right\}\\
&\le \left\{N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2\left\|\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1l}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2l}^{0}\right\|^2\left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}\right\|^2\right\}^{1/2}\\
&\quad\times\left\{N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)x_{it}x_{lt}\right]^2\right\}^{1/2}\\
&\le \left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2\right]^{1/2}\left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}\right\|\left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2\right]^{1/2}O_p(1)\\
&= \left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2\right]^{1/2}O_p(1),
\end{aligned}$$

and $a_2(\theta) = O_p\left(C_{NT}^{-1}\right)$ for all $\theta$. In an analogous way it is proved that $a_3(\theta) = O_p\left(C_{NT}^{-1}\right)$ for all $\theta$. Finally,

$$\begin{aligned}
a_4(\theta) &= N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left\{\begin{array}{l}\left[\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right]'\left\{\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\}\left[\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1l}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2l}^{0}\right]\\ \times\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)x_{it}x_{lt}\right]\end{array}\right\}\\
&\le \left\{N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left\|\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2\left\|\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1l}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2l}^{0}\right\|^2\left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\|^2\right\}^{1/2}\\
&\quad\times\left\{N^{-2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)x_{it}x_{lt}\right]^2\right\}^{1/2}\\
&\le \left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\|\left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^{0}+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^{0}\right\|^2+\right]O_p(1)\\
&\le \left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\|\left\{\left\|\hat{\mathbf{H}}_{11}(\theta)\right\|^2\left[N^{-1}\sum_{i=1}^{N}\left\|\boldsymbol{\lambda}_{1i}^{0}\right\|^2\right]+\left\|\hat{\mathbf{H}}_{21}(\theta)\right\|^2\left[N^{-1}\sum_{i=1}^{N}\left\|\boldsymbol{\lambda}_{2i}^{0}\right\|^2\right]\right\}O_p(1)\\
&= \left\|\left[\mathbf{D}_{\hat{\boldsymbol{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\boldsymbol{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\boldsymbol{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\|O_p(1)
\end{aligned}$$

where $O_p(1)$ comes from Lemma A.1(b) and Assumptions C1 and C2. Now, $\left\|\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right\|=O_p\left(C_{NT}^{-1}\right)$ for all $\theta$: this is because

$$\begin{aligned}
\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta) &= \frac{\hat{\mathbf{\Lambda}}_1(\theta)'\hat{\mathbf{\Lambda}}_1(\theta)}{N}-\frac{\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]'\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta)\right]}{N}\\
&= N^{-1}\sum_{i=1}^{N}\left\{\hat{\boldsymbol{\lambda}}_{1i}(\theta)\hat{\boldsymbol{\lambda}}_{1i}(\theta)'-\left[\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]\left[\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]'\right\}\\
&= N^{-1}\sum_{i=1}^{N}\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]'\\
&\quad +N^{-1}\sum_{i=1}^{N}\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]\left[\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]'\\
&\quad +N^{-1}\sum_{i=1}^{N}\left[\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]\left[\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right]'
\end{aligned}$$

so that

$$\begin{aligned}
\left\|\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right\| &\leq N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right\|^2\\
&\quad +2\left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right\|^2\right]^{1/2}\\
&\quad \times\left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0+\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right\|^2\right]^{1/2}\\
&= N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right\|^2\\
&\quad +2\left[N^{-1}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_{1i}(\theta)-\hat{\mathbf{H}}_{11}(\theta)'\boldsymbol{\lambda}_{1i}^0-\hat{\mathbf{H}}_{21}(\theta)'\boldsymbol{\lambda}_{2i}^0\right\|^2\right]^{1/2}O_p(1)
\end{aligned}$$

and the result follows. In general,

$$\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}=\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]\left[\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}$$

and

$$\left\|\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\|\leq\left\|\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right\|\left\|\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}\right\|\left\|\left[\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\|.$$

The matrix $\mathbf{\Lambda}_j^{0\prime}\mathbf{\Lambda}_j^0/N$ converges to a positive definite matrix by Assumption C2, for $j=1,2$, and the rank of $\hat{\mathbf{H}}_{11}(\theta)$ is equal to $R^0$ for all $\theta$: since the rank of $\hat{\mathbf{H}}_{21}(\theta)$ is equal to $R^0$ for $\theta\neq\theta^0$, and $\hat{\mathbf{H}}_{21}\left(\theta^0\right)=\mathbf{0}_{R^0}$, this implies that $\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)$ converges to a positive definite matrix. Since $\left\|\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right\|=O_p\left(C_{NT}^{-1}\right)$, $\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)$ also converges to a positive definite matrix: this implies that $\left\|\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}\right\|=O_p(1)$: therefore,

$$\begin{aligned}
\left\|\left[\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)\right]^{-1}-\left[\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]^{-1}\right\| &= \left\|\mathbf{D}_{\hat{\mathbf{\Lambda}}_1}(\theta)-\mathbf{D}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right\|O_p(1)\\
&= O_p\left(C_{NT}^{-1}\right)
\end{aligned}$$

and $a_4(\theta)=O_p\left(C_{NT}^{-1}\right)$ for all $\theta$. Combining all above results, we have

$$a_1(\theta)+a_2(\theta)+a_3(\theta)+a_4(\theta)=O_p\left(C_{NT}^{-2}\right)+O_p\left(C_{NT}^{-1}\right)+O_p\left(C_{NT}^{-1}\right)+O_p\left(C_{NT}^{-1}\right)=O_p\left(C_{NT}^{-1}\right):$$

this completes the proof of the lemma. ■

**Proof of Lemma A.3.** Let

$$\mathbf{P}_{\mathbf{\Lambda}_j^0}=\mathbf{\Lambda}_j^0\left(\mathbf{\Lambda}_j^{0\prime}\mathbf{\Lambda}_j^0\right)^{-1}\mathbf{\Lambda}_j^{0\prime},\quad j=1,2,$$

and recall $\mathbf{P}_{\mathbf{\Lambda}_j^0\hat{\mathbf{H}}_{jj}+\mathbf{\Lambda}_m^0\hat{\mathbf{H}}_{mj}}(\theta)$ as defined in (19). Write

$$
\begin{aligned}
& S_{\mathbf{F}}\left[\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}(\theta)+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}(\theta),\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}(\theta)+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}(\theta),\theta\right]-S_{\mathbf{F}}\left(\mathbf{\Lambda}_1^0,\mathbf{\Lambda}_2^0,\theta^0\right) \\
= \;& (NT)^{-1}\sum_{t=1}^{T}\mathbf{x}_t'\left\{\left[\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{P}_{\mathbf{\Lambda}_1^0}-\mathbb{I}_{1t}(\theta)\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]+\left[\mathbb{I}_{2t}\left(\theta^0\right)\mathbf{P}_{\mathbf{\Lambda}_2^0}-\mathbb{I}_{2t}(\theta)\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\right]\right\}\mathbf{x}_t \\
= \;& (NT)^{-1}\sum_{t=1}^{T}\left\{\begin{array}{l}\left[\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{\Lambda}_1^0\mathbf{f}_t^0+\mathbb{I}_{2t}\left(\theta^0\right)\mathbf{\Lambda}_2^0\mathbf{f}_t^0+\mathbf{e}_t\right]' \\ \times\left\{\left[\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{P}_{\mathbf{\Lambda}_1^0}-\mathbb{I}_{1t}(\theta)\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]+\left[\mathbb{I}_{2t}\left(\theta^0\right)\mathbf{P}_{\mathbf{\Lambda}_2^0}-\mathbb{I}_{2t}(\theta)\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\right]\right\} \\ \times\left[\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{\Lambda}_1^0\mathbf{f}_t^0+\mathbb{I}_{2t}\left(\theta^0\right)\mathbf{\Lambda}_2^0\mathbf{f}_t^0+\mathbf{e}_t\right]\end{array}\right\} \\
= \;& b_1(\theta)+b_2(\theta)+b_3(\theta),
\end{aligned}
$$

where

$$
\begin{aligned}
b_1(\theta) = \;& (NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\left[\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_1^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_1^0}\mathbf{\Lambda}_1^0\mathbf{f}_t^0-\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_1^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\mathbf{\Lambda}_1^0\mathbf{f}_t^0\right] \\
& +(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\left[\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_1^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_1^0}\mathbf{\Lambda}_1^0\mathbf{f}_t^0-\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_1^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\mathbf{\Lambda}_1^0\mathbf{f}_t^0\right] \\
& +(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\mathbb{I}_{2t}\left(\theta^0\right)\left[\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_2^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_2^0}\mathbf{\Lambda}_2^0\mathbf{f}_t^0-\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_2^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\mathbf{\Lambda}_2^0\mathbf{f}_t^0\right] \\
& +(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}(\theta)\mathbb{I}_{2t}\left(\theta^0\right)\left[\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_2^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_2^0}\mathbf{\Lambda}_2^0\mathbf{f}_t^0-\mathbf{f}_t^{0\prime}\mathbf{\Lambda}_2^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\mathbf{\Lambda}_2^0\mathbf{f}_t^0\right] \\
= \;& b_{11}(\theta)+b_{12}(\theta)+b_{13}(\theta)+b_{14}(\theta),
\end{aligned}
$$

$$
\begin{aligned}
b_2(\theta) = \;& 2(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\left[\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_1^0}\mathbf{\Lambda}_1^0\mathbf{f}_t^0-\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\mathbf{\Lambda}_1^0\mathbf{f}_t^0\right] \\
& +2(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\left[\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_1^0}\mathbf{\Lambda}_1^0\mathbf{f}_t^0-\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\mathbf{\Lambda}_1^0\mathbf{f}_t^0\right] \\
& +2(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\mathbb{I}_{2t}\left(\theta^0\right)\left[\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_2^0}\mathbf{\Lambda}_2^0\mathbf{f}_t^0-\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\mathbf{\Lambda}_2^0\mathbf{f}_t^0\right] \\
& +2(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}(\theta)\mathbb{I}_{2t}\left(\theta^0\right)\left[\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_2^0}\mathbf{\Lambda}_2^0\mathbf{f}_t^0-\mathbf{e}_t'\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\mathbf{\Lambda}_2^0\mathbf{f}_t^0\right] \\
= \;& b_{21}(\theta)+b_{22}(\theta)+b_{23}(\theta)+b_{24}(\theta),
\end{aligned}
$$

and

$$
\begin{aligned}
b_3(\theta) = \;& (NT)^{-1}\sum_{t=1}^{T}\mathbf{e}_t'\left[\mathbb{I}_{1t}\left(\theta^0\right)-\mathbb{I}_{1t}(\theta)\right]\mathbf{P}_{\mathbf{\Lambda}_1^0}\mathbf{e}_t \\
& +(NT)^{-1}\sum_{t=1}^{T}\mathbf{e}_t'\left[\mathbb{I}_{2t}\left(\theta^0\right)-\mathbb{I}_{2t}(\theta)\right]\mathbf{P}_{\mathbf{\Lambda}_2^0}\mathbf{e}_t \\
& +(NT)^{-1}\sum_{t=1}^{T}\mathbf{e}_t'\mathbb{I}_{1t}(\theta)\left[\mathbf{P}_{\mathbf{\Lambda}_1^0}-\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]\mathbf{e}_t \\
& +(NT)^{-1}\sum_{t=1}^{T}\mathbf{e}_t'\mathbb{I}_{2t}(\theta)\left[\mathbf{P}_{\mathbf{\Lambda}_2^0}-\mathbf{P}_{\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{12}}(\theta)\right]\mathbf{e}_t \\
= \;& b_{31}(\theta)+b_{32}(\theta)+b_{33}(\theta)+b_{34}(\theta).
\end{aligned}
$$

Consider $b_1(\theta)$ first. We have

$$
\begin{aligned}
b_{11}(\theta) = \;& \operatorname{tr}\left\{N^{-1}\left[\mathbf{\Lambda}_1^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_1^0}\mathbf{\Lambda}_1^0-\mathbf{\Lambda}_1^{0\prime}\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\mathbf{\Lambda}_1^0\right]\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\right]\right\} \\
= \;& \operatorname{tr}\left\{\left\{N^{-1}\mathbf{\Lambda}_1^{0\prime}\left[\mathbf{P}_{\mathbf{\Lambda}_1^0}-\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]\mathbf{\Lambda}_1^0\right\}\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^0\mathbf{f}_t^{0\prime}\right]\right\} \\
\xrightarrow{p} \;& \operatorname{tr}\left\{\left\{\operatorname{p}\lim_{N\to\infty}\left\{N^{-1}\mathbf{\Lambda}_1^{0\prime}\left[\mathbf{P}_{\mathbf{\Lambda}_1^0}-\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]\mathbf{\Lambda}_1^0\right\}\right\}\mathbf{\Sigma}_{1\mathbf{f}}^0\left(\theta,\theta^0\right)\right\} \\
= \;& \operatorname{tr}\left[\mathbf{B}_{11}(\theta)\cdot\mathbf{\Sigma}_{1\mathbf{f}}^0\left(\theta,\theta^0\right)\right],
\end{aligned}
$$

where $\mathbf{B}_{11}(\theta)=\operatorname{p}\lim_{N\to\infty}\left\{N^{-1}\mathbf{\Lambda}_1^{0\prime}\left[\mathbf{P}_{\mathbf{\Lambda}_1^0}-\mathbf{P}_{\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{21}}(\theta)\right]\mathbf{\Lambda}_1^0\right\}$. Now $\mathbf{B}_{11}(\theta)$ is different from zero by Assumption C2 and it is also positive semi-definite. The matrix $\mathbf{\Sigma}_{1\mathbf{f}}^0\left(\theta,\theta^0\right)$ is positive definite by Assumption C1. It then follows that

$\text{p}\lim_{N,T\to\infty} b_{11}(\theta) = \text{tr}\left[\mathbf{B}_{11}(\theta) \cdot \mathbf{\Sigma}^0_{1\mathbf{f}}\left(\theta,\theta^0\right)\right] > 0$. Consider now

$$\begin{aligned}
b_{12}(\theta) &= \text{tr}\left\{N^{-1}\left[\mathbf{\Lambda}^{0\prime}_1\mathbf{P}_{\mathbf{\Lambda}^0_1}\mathbf{\Lambda}^0_1 - \mathbf{\Lambda}^{0\prime}_1\mathbf{P}_{\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{12}}(\theta)\,\mathbf{\Lambda}^0_1\right]\left[T^{-1}\sum_{t=1}^{T}\mathbb{I}_{2t}(\theta)\,\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_t\mathbf{f}^{0\prime}_t\right]\right\} \\
&= \text{tr}\left\{\left\{N^{-1}\mathbf{\Lambda}^{0\prime}_1\left[\mathbf{P}_{\mathbf{\Lambda}^0_1} - \mathbf{P}_{\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{12}}(\theta)\right]\mathbf{\Lambda}^0_1\right\}\left\{T^{-1}\sum_{t=1}^{T}\left[1-\mathbb{I}_{1t}(\theta)\right]\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_t\mathbf{f}^{0\prime}_t\right\}\right\} \\
&= \text{tr}\left\{\left\{N^{-1}\mathbf{\Lambda}^{0\prime}_1\left[\mathbf{P}_{\mathbf{\Lambda}^0_1} - \mathbf{P}_{\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{12}}(\theta)\right]\mathbf{\Lambda}^0_1\right\}\left\{T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^0\right)\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_t\mathbf{f}^{0\prime}_t - T^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}^0_t\mathbf{f}^{0\prime}_t\right\}\right\} \\
&\xrightarrow{p} \text{tr}\left\{\text{p}\lim_{N\to\infty}\left\{N^{-1}\mathbf{\Lambda}^{0\prime}_1\left[\mathbf{P}_{\mathbf{\Lambda}^0_1} - \mathbf{P}_{\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{12}}(\theta)\right]\mathbf{\Lambda}^0_1\right\}\left[\mathbf{\Sigma}^0_{1\mathbf{f}}\left(\theta^0,\theta^0\right) - \mathbf{\Sigma}^0_{1\mathbf{f}}\left(\theta,\theta^0\right)\right]\right\} \\
&= \text{tr}\left\{\mathbf{B}_{12}(\theta)\left[\mathbf{\Sigma}^0_{1\mathbf{f}}\left(\theta^0,\theta^0\right) - \mathbf{\Sigma}^0_{1\mathbf{f}}\left(\theta,\theta^0\right)\right]\right\},
\end{aligned}$$

where $\mathbf{B}_{12}(\theta) = \text{p}\lim_{N\to\infty}\left\{N^{-1}\mathbf{\Lambda}^{0\prime}_1\left[\mathbf{P}_{\mathbf{\Lambda}^0_1} - \mathbf{P}_{\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{12}}(\theta)\right]\mathbf{\Lambda}^0_1\right\}$: taking into account Assumption C1, it follows that $\text{p}\lim_{N,T\to\infty} b_{12}(\theta) > 0$. In an analogous way it can be proved that $\text{p}\lim_{N,T\to\infty} b_{13}(\theta) > 0$ and $\text{p}\lim_{N,T\to\infty} b_{14}(\theta) > 0$. Then

$$\text{p}\lim_{N\to\infty} b_1(\theta) = \text{p}\lim_{N\to\infty} b_{11}(\theta) + \text{p}\lim_{N\to\infty} b_{12}(\theta) + \text{p}\lim_{N\to\infty} b_{13}(\theta) + \text{p}\lim_{N\to\infty} b_{14}(\theta) > 0, \quad \forall\theta \neq \theta^0.$$

Consider now $b_2(\theta)$. We have

$$b_{21}(\theta) = 2(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{e}'_t\mathbf{P}_{\mathbf{\Lambda}^0_1}\mathbf{\Lambda}^0_1\mathbf{f}^0_t - 2(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{e}'_t\mathbf{P}_{\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{21}}(\theta)\,\mathbf{\Lambda}^0_1\mathbf{f}^0_t.$$

By Lemma A.1(c) and Assumption C1,

$$\begin{aligned}
\left|(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{e}'_t\mathbf{P}_{\mathbf{\Lambda}^0_1}\mathbf{\Lambda}^0_1\mathbf{f}^0_t\right| &= \left|(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{1t}\left(\theta^0\right)e_{it}\mathbf{\lambda}^{0\prime}_{1i}\mathbf{f}^0_t\right| \\
&\leq \left[T^{-1}\sum_{t=1}^{T}\left\|\mathbb{I}_{1t}(\theta)\,\mathbf{f}^0_t\right\|^2\right]^{1/2} N^{-1/2}\left[N^{-1}\sum_{i=1}^{N}\left\|T^{-1/2}\sum_{t=1}^{T}\mathbb{I}_{1t}\left(\theta^0\right)e_{it}\mathbf{\lambda}^0_{1i}\right\|^2\right]^{1/2} \\
&= O_p\left(\frac{1}{\sqrt{N}}\right).
\end{aligned}$$

Further,

$$(NT)^{-1}\sum_{t=1}^{T}\mathbb{I}_{1t}(\theta)\,\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{e}'_t\mathbf{P}_{\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{21}}(\theta)\,\mathbf{\Lambda}^0_1\mathbf{f}^0_t = O_p\left(\frac{1}{\sqrt{N}}\right).$$

Therefore, $b_{21}(\theta) = O_p\left(1\big/\sqrt{N}\right)$. In an analogous way, it can be proved that $b_{22}(\theta) = O_p\left(1\big/\sqrt{N}\right)$, $b_{23}(\theta) = O_p\left(1\big/\sqrt{N}\right)$, $b_{24}(\theta) = O_p\left(1\big/\sqrt{N}\right)$. Therefore, $b_2(\theta) = O_p\left(1\big/\sqrt{N}\right) \xrightarrow{p} 0$ as $N\to\infty$.

Finally, consider $b_3(\theta)$. We have, $b_{31}(\theta) = o_p(1)$ and $b_{32}(\theta) = o_p(1)$. Further, $\left[\mathbf{P}_{\mathbf{\Lambda}^0_1} - \mathbf{P}_{\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{11}+\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{21}}(\theta)\right]$ and $\left[\mathbf{P}_{\mathbf{\Lambda}^0_2} - \mathbf{P}_{\mathbf{\Lambda}^0_2\hat{\mathbf{H}}_{22}+\mathbf{\Lambda}^0_1\hat{\mathbf{H}}_{12}}(\theta)\right]$ are positive semi-definite matrices, which implies that $b_{33}(\theta) \geq 0$ and $b_{34}(\theta) \geq 0$: this implies that $\text{p}\lim_{N,T\to\infty} b_3(\theta) \geq 0$. This completes the proof of the lemma. ■

## A.2 Proofs of Results in Section 3.5

Let

$$g_{it}^{0}\left(\theta_1,\theta_2\right)=\left|\mathbb{I}_{2t}\left(\theta_2\right)-\mathbb{I}_{2t}\left(\theta_1\right)\right|\left\|\mathbf{f}_t^0 e_{it}\right\|,\quad i=1,\ldots,N,\quad t=1,\ldots,T,$$

$$q_{t}^{0}\left(\theta_1,\theta_2\right)=\left|\mathbb{I}_{2t}\left(\theta_2\right)-\mathbb{I}_{2t}\left(\theta_1\right)\right|\left\|\mathbf{f}_t^0\right\|,\quad t=1,\ldots,T,$$

$$w_{it}^{0}\left(\theta\right)=\left|\mathbb{I}_{2t}\left(\theta\right)-\mathbb{I}_{2t}\left(\theta^0\right)\right|\left(\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\right)^2,\quad i=1,\ldots,N,\quad t=1,\ldots,T,$$

$$w^{0}\left(\alpha^0,\theta\right)=\frac{1}{N^{\alpha^0}}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}w_{it}^{0}\left(\theta\right),$$

$$\mathbf{h}^{0}\left(\alpha^0,\theta\right)=\frac{1}{N^{\alpha^0}}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{I}_{2t}\left(\theta\right)\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0 e_{it}.$$

**Lemma A.4** *There exists a $C_1<\infty$ such that for all $\theta_L\leq\theta_1\leq\theta_2\leq\theta_U$ and $s\leq 4$,*

$$\mathrm{E}\left\{\left[g_{it}^{0}\left(\theta_1,\theta_2\right)\right]^s\right\}\leq C_1\left|\theta_2-\theta_1\right|,\quad i=1,\ldots,N, \tag{20}$$

*and*

$$\mathrm{E}\left\{\left[q_{t}^{0}\left(\theta_1,\theta_2\right)\right]^s\right\}\leq C_1\left|\theta_2-\theta_1\right|. \tag{21}$$

**Lemma A.5** *There exists a $K<\infty$ such that for all $\theta_L\leq\theta_1\leq\theta_2\leq\theta_U$,*

$$\mathrm{E}\left[\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\left[q_{t}^{0}\left(\theta_1,\theta_2\right)\right]^2-\mathrm{E}\left\{\left[q_{t}^{0}\left(\theta_1,\theta_2\right)\right]^2\right\}\right\}\right|^2\right]\leq K\left|\theta_2-\theta_1\right|.$$

**Lemma A.6** *There exist constants $B>0$ and $0<d<\infty$ such that for all $\eta>0$ and $\varepsilon>0$, there exists a $\bar{v}<\infty$ such that for all $N$ and $T$,*

$$\Pr\left[\inf_{\frac{\bar{v}}{N^{\left(2\alpha^0-1\right)}T}\leq\left|\theta-\theta^0\right|\leq B}\frac{w^{0}\left(\alpha^0,\theta\right)}{\left|\theta-\theta^0\right|}<\left(1-\eta\right)d\right]\leq\varepsilon.$$

**Lemma A.7** *For all $\eta>0$ and $\varepsilon>0$, there exists some $\bar{v}<\infty$ such that for any $B<\infty$,*

$$\Pr\left[\sup_{\frac{\bar{v}}{N^{\left(2\alpha^0-1\right)}T}\leq\left|\theta-\theta^0\right|\leq B}\frac{\left\|\mathbf{h}^{0}\left(\alpha^0,\theta\right)-\mathbf{h}^{0}\left(\alpha^0,\theta^0\right)\right\|}{\left|\theta-\theta^0\right|}>\eta\right]\leq\varepsilon.$$

**Proof of Theorem 3.4.** Let $B$ and $d$ be defined as in Lemma A.6. Pick $\eta>0$ small enough so that

$$\left(1-\eta\right)d-2\eta>0. \tag{22}$$

Let $\mathbb{E}_{NT}$ be the joint event that $\left|\hat{\theta}-\theta^0\right|\leq B$, $\left\|\hat{\boldsymbol{\lambda}}_{ji}'\hat{\mathbf{f}}_t-\boldsymbol{\lambda}_{ji}^{0\prime}\mathbf{f}_t^0\right\|$ is small enough so that (25) below is satisfied, for $j=1,2$, $i=1,\ldots,N$, $t=1,\ldots,T$, and

$$\inf_{\frac{\bar{v}}{N^{\left(2\alpha^0-1\right)}T}\leq\left|\theta-\theta^0\right|\leq B}\frac{w^{0}\left(\alpha^0,\theta\right)}{\left|\theta-\theta^0\right|}\geq\left(1-\eta\right)d, \tag{23}$$

and

$$\sup_{\frac{\bar{v}}{N^{\left(2\alpha^0-1\right)}T}\leq\left|\theta-\theta^0\right|\leq B}\frac{\left\|\mathbf{h}^{0}\left(\alpha^0,\theta\right)-\mathbf{h}^{0}\left(\alpha^0,\theta^0\right)\right\|}{\left|\theta-\theta^0\right|}\leq\eta. \tag{24}$$

Fix $\varepsilon>0$ and pick $\bar{v}$, $\bar{N}$ and $\bar{T}$ so that $\Pr\left(\mathbb{E}_{NT}\right)\geq 1-\varepsilon$ for all $N\geq\bar{N}$ and $T\geq\bar{T}$, which is possible under Corollary 3.1, and Lemmas A.6 and A.7. Given $S\left(\boldsymbol{\Lambda},\mathbf{F},\theta\right)$ defined in (2), let

$$S\left(\alpha^0,\boldsymbol{\Lambda},\mathbf{F},\theta\right)=\frac{1}{N^{\alpha^0}T}\sum_{t=1}^{T}\left[\mathbf{x}_t-\boldsymbol{\Lambda}_1\mathbf{f}_t-\boldsymbol{\Delta}\mathbf{f}_{2t}\left(\theta\right)\right]'\left[\mathbf{x}_t-\boldsymbol{\Lambda}_1\mathbf{f}_t-\boldsymbol{\Delta}\mathbf{f}_{2t}\left(\theta\right)\right],$$

where $\mathbf{f}_{2t}(\theta) = \mathbb{I}_{2t}(\theta)\mathbf{f}_t$ and $\boldsymbol{\Delta} = \boldsymbol{\Lambda}_2 - \boldsymbol{\Lambda}_1$. Since $S(\alpha^0, \boldsymbol{\Lambda}, \mathbf{F}, \theta)$ is continuous in $(\boldsymbol{\Lambda}, \mathbf{F})$, for small enough $\left\|\hat{\boldsymbol{\lambda}}_{ji}'\hat{\mathbf{f}}_t - \boldsymbol{\lambda}_{ji}^{0\prime}\mathbf{f}_t^0\right\|$, for $j = 1, 2$, $i = 1, \ldots, N$, $t = 1, \ldots, T$, it follows that

$$\begin{aligned}
S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta\right) - S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta^0\right) &= \frac{1}{N^{\alpha^0}T}\sum_{t=1}^{T}\left[\mathbf{x}_t - \hat{\boldsymbol{\Lambda}}_1\hat{\mathbf{f}}_t - \hat{\boldsymbol{\Delta}}\hat{\mathbf{f}}_{2t}(\theta)\right]'\left[\mathbf{x}_t - \hat{\boldsymbol{\Lambda}}_1\hat{\mathbf{f}}_t - \hat{\boldsymbol{\Delta}}\hat{\mathbf{f}}_{2t}(\theta)\right] \\
&\quad - \frac{1}{N^{\alpha^0}T}\sum_{t=1}^{T}\left[\mathbf{x}_t - \hat{\boldsymbol{\Lambda}}_1\hat{\mathbf{f}}_t - \hat{\boldsymbol{\Delta}}\hat{\mathbf{f}}_{2t}(\theta^0)\right]'\left[\mathbf{x}_t - \hat{\boldsymbol{\Lambda}}_1\hat{\mathbf{f}}_t - \hat{\boldsymbol{\Delta}}\hat{\mathbf{f}}_{2t}(\theta^0)\right] \\
&= D\left\{\begin{array}{l} \frac{1}{N^{\alpha^0}T}\sum_{t=1}^{T}\left[\mathbf{x}_t - \boldsymbol{\Lambda}_1^0\mathbf{f}_t^0 - \boldsymbol{\Delta}^0\mathbf{f}_{2t}^0(\theta)\right]'\left[\mathbf{x}_t - \boldsymbol{\Lambda}_1^0\mathbf{f}_t^0 - \boldsymbol{\Delta}^0\mathbf{f}_{2t}^0(\theta)\right] \\ - \frac{1}{N^{\alpha^0}T}\sum_{t=1}^{T}\left[\mathbf{x}_t - \boldsymbol{\Lambda}_1^0\mathbf{f}_t^0 - \boldsymbol{\Delta}^0\mathbf{f}_{2t}^0(\theta^0)\right]'\left[\mathbf{x}_t - \boldsymbol{\Lambda}_1^0\mathbf{f}_t^0 - \boldsymbol{\Delta}^0\mathbf{f}_{2t}^0(\theta^0)\right] \end{array}\right\} \\
&= D\left[S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta) - S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta^0)\right],
\end{aligned} \tag{25}$$

for some $D > 0$, where $\hat{\mathbf{f}}_{2t}(\theta) = \mathbb{I}_{2t}(\theta)\hat{\mathbf{f}}_t$, $\hat{\boldsymbol{\Delta}} = \hat{\boldsymbol{\Lambda}}_2 - \hat{\boldsymbol{\Lambda}}_1$, $\mathbf{f}_{2t}^0(\theta) = \mathbb{I}_{2t}(\theta)\mathbf{f}_t^0$ and $\boldsymbol{\Delta}^0 = \boldsymbol{\Lambda}_2^0 - \boldsymbol{\Lambda}_1^0$: the sign of $S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta\right) - S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta^0\right)$ is then equal to the sign of $S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta) - S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta^0)$. We have

$$\begin{aligned}
S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta) - S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta^0) &= \frac{1}{N^{\alpha^0}T}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\Delta}^{0\prime}\boldsymbol{\Delta}^0\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right] \\
&\quad - 2\frac{1}{N^{\alpha^0}T}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\Delta}^{0\prime}\mathbf{e}_t \\
&= S_1(\alpha^0, \theta) + S_2(\alpha^0, \theta)
\end{aligned}$$

and

$$\begin{aligned}
\frac{S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta) - S(\alpha^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \theta^0)}{|\theta - \theta^0|} &= \frac{1}{N^{\alpha^0}T|\theta - \theta^0|}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\Delta}^{0\prime}\boldsymbol{\Delta}^0\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right] \\
&\quad - 2\frac{1}{N^{\alpha^0}T|\theta - \theta^0|}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\Delta}^{0\prime}\mathbf{e}_t \\
&= \frac{S_1(\alpha^0, \theta)}{|\theta - \theta^0|} + \frac{S_2(\alpha^0, \theta)}{|\theta - \theta^0|}.
\end{aligned} \tag{26}$$

Suppose $\theta \in \left[\theta^0 + \bar{v}N^{-(2\alpha^0 - 1)}T^{-1}, \theta^0 + B\right]$ and that event $\mathbb{E}_{NT}$ holds. It follows that

$$\begin{aligned}
\frac{S_1(\alpha^0, \theta)}{\theta - \theta^0} &= \frac{1}{N^{\alpha^0}T(\theta - \theta^0)}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\delta}_i^0\boldsymbol{\delta}_i^{0\prime}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right] \\
&= \frac{1}{N^{\alpha^0}T(\theta - \theta^0)}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\mathbb{I}_{2t}(\theta) - \mathbb{I}_{2t}(\theta^0)\right|\left(\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0\right)^2 \\
&= \frac{w^0(\alpha^0, \theta)}{\theta - \theta^0},
\end{aligned} \tag{27}$$

and

$$\begin{aligned}
\frac{S_2(\alpha^0, \theta)}{\theta - \theta^0} &= -2\frac{1}{N^{\alpha^0}T|\theta - \theta^0|}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\Delta}^{0\prime}\mathbf{e}_t \\
&= -2\frac{1}{N^{\alpha^0}T(\theta - \theta^0)}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\delta}_i^0 e_{it} \\
&\geq -2\frac{1}{\theta - \theta^0}\left\|\frac{1}{N^{\alpha^0}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\mathbf{f}_{2t}^0(\theta) - \mathbf{f}_{2t}^0(\theta^0)\right]'\boldsymbol{\delta}_i^0 e_{it}\right\| \\
&= -2\frac{\left\|\mathbf{h}^0(\alpha^0, \theta) - \mathbf{h}^0(\alpha^0, \theta^0)\right\|}{\theta - \theta^0},
\end{aligned} \tag{28}$$

By (22) through (28) it the follows that for some $D > 0$,

$$\frac{S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta\right) - S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta^0\right)}{\theta - \theta^0} \geq D\left[\frac{w^0(\alpha^0, \theta)}{\theta - \theta^0} - 2\frac{\left\|\mathbf{h}^0(\alpha^0, \theta) - \mathbf{h}^0(\alpha^0, \theta^0)\right\|}{\theta - \theta^0}\right] \geq D\left[(1 - \eta)d - 2\eta\right] \geq 0.$$

Given the event $\mathbb{E}_{NT}$, if $\theta \in \left[\theta^0 + \bar{v}N^{-(2\alpha^0 - 1)}T^{-1}, \theta^0 + B\right]$ then $S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta\right) - S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta^0\right) > 0$. In a similar way, it can be shown that if $\theta \in \left[\theta^0 - B, \theta^0 - \bar{v}N^{-(2\alpha^0 - 1)}T^{-1}\right]$ then $S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta\right) - S\left(\alpha^0, \hat{\boldsymbol{\Lambda}}, \hat{\mathbf{F}}, \theta^0\right) > 0$. As

$S\left(\alpha^0, \hat{\mathbf{\Lambda}}, \hat{\mathbf{F}}, \hat{\theta}\right) - S\left(\alpha^0, \hat{\mathbf{\Lambda}}, \hat{\mathbf{F}}, \theta^0\right) \leq 0$, if $\mathbb{E}_{NT}$ occurs then $\left|\hat{\theta} - \theta^0\right| \leq \bar{v} N^{-(2\alpha^0-1)} T^{-1}$: since $\Pr\left(\mathbb{E}_{NT}\right) \geq 1-\varepsilon$ for $N \geq \bar{N}$ and $T \geq \bar{T}$, then $\Pr\left(\left|\hat{\theta}-\theta^0\right| > \bar{v} N^{-(2\alpha^0-1)} T^{-1}\right) \leq \varepsilon$ for $N \geq \bar{N}$ and $T \geq \bar{T}$: this is sufficient to show that $N^{(2\alpha^0-1)} T\left(\hat{\theta}-\theta^0\right) = O_p(1)$. The convergence rate of the estimator for the loadings follows from (11). ■

**Proof of Corollary 3.2.** Corollary 3.2 easily follows from Theorem 3.4 and the proof is omitted. ■

**Proof of Lemma A.4.** We show (20): the proof of (21) is analogous. Given a random matrix $\mathbf{A}$,

$$\frac{\partial}{\partial\theta} \mathrm{E}\left[\mathbf{A} \mathbb{I}_{2t}(\theta)\right] = \mathrm{E}\left(\mathbf{A} \left| z_t = \theta\right.\right) f_{\mathcal{Z}}(\theta). \tag{29}$$

Under Assumption CR(b)

$$\begin{aligned}
\frac{\partial}{\partial\theta} \mathrm{E}\left[\left\|\mathbf{f}_t^0 e_{it}\right\|^s \mathbb{I}_{2t}(\theta)\right] &= \mathrm{E}\left(\left\|\mathbf{f}_t^0 e_{it}\right\|^s \left| z_t = \theta\right.\right) f_{\mathcal{Z}}(\theta) \\
&\leq \left[\mathrm{E}\left(\left\|\mathbf{f}_t^0 e_{it}\right\|^4 \left| z_t = \theta\right.\right)\right]^{s/4} f_{\mathcal{Z}}(\theta) \\
&\leq C^{s/4} \bar{f} \leq C_1,
\end{aligned}$$

where $C_1 = \max[1, C]\bar{f}$. Notice that $\mathbb{I}_{2t}(\theta_1) - \mathbb{I}_{2t}(\theta_2)$ is either equal to one or to zero: by a first-order Taylor expansion it follows that

$$\mathrm{E}\left\{\left[g_{it}^0(\theta_1, \theta_2)\right]^s\right\} = \mathrm{E}\left[\left\|\mathbf{f}_t^0 e_{it}\right\|^s \mathbb{I}_{2t}(\theta_1)\right] - \mathrm{E}\left[\left\|\mathbf{f}_t^0 e_{it}\right\|^s \mathbb{I}_{2t}(\theta_2)\right] \leq C_1 \left|\theta_2 - \theta_1\right|.$$

■

**Proof of Lemma A.5.** Lemma 3.4 in Peligrad (1982) shows that under Assumption CR(a) there exists a $K' < \infty$ such that, taking into account (21) in Lemma A.4,

$$\begin{aligned}
\mathrm{E}\left[\left|\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left\{\left[q_t^0(\theta_1, \theta_2)\right]^2 - \mathrm{E}\left\{\left[q_t^0(\theta_1, \theta_2)\right]^2\right\}\right\}\right|^2\right] &\leq K' \mathrm{E}\left\{\left\{\left[q_t^0(\theta_1, \theta_2)\right]^2 - \mathrm{E}\left\{\left[q_t^0(\theta_1, \theta_2)\right]^2\right\}\right\}^2\right\} \\
&\leq 2K' \mathrm{E}\left\{\left[q_t^0(\theta_1, \theta_2)\right]^4\right\} \\
&\leq 2K' C_1 \left|\theta_2 - \theta_1\right| :
\end{aligned}$$

setting $K = 2K'C_1$ completes the proof of the lemma. ■

**Proof of Lemma A.6.** For $\theta \leq \theta^0$,

$$\begin{aligned}
\mathrm{E}\left[w^0\left(\alpha^0, \theta\right)\right] &= \frac{1}{N^{\alpha^0}} \sum_{i=1}^{N} \mathrm{E}\left[w_{it}^0(\theta)\right] \\
&= \frac{1}{N^{\alpha^0}} \sum_{i=1}^{N^{\alpha^0}} \mathrm{E}\left[w_{it}^0(\theta)\right] + \frac{1}{N^{\alpha^0}} \sum_{i=N^{\alpha^0}+1}^{N} \mathrm{E}\left[w_{it}^0(\theta)\right] \\
&= \frac{1}{N^{\alpha^0}} \sum_{i=1}^{N^{\alpha^0}} \boldsymbol{\delta}_i^{0\prime} \left[\mathbf{\Sigma}_{2\mathbf{f}}^0(\theta, \theta) - \mathbf{\Sigma}_{2\mathbf{f}}^0\left(\theta^0, \theta^0\right)\right] \boldsymbol{\delta}_i^0 + \frac{1}{N^{\alpha^0}} \sum_{i=N^{\alpha^0}+1}^{N} \boldsymbol{\delta}_i^{0\prime} \left[\mathbf{\Sigma}_{2\mathbf{f}}^0(\theta, \theta) - \mathbf{\Sigma}_{2\mathbf{f}}^0\left(\theta^0, \theta^0\right)\right] \boldsymbol{\delta}_i^0,
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \mathrm{E}\left[w^0\left(\alpha^0, \theta\right)\right]}{\partial\theta} &= \frac{1}{N^{\alpha^0}} \sum_{i=1}^{N} \boldsymbol{\delta}_i^{0\prime} \mathbf{D}_{\mathbf{f}}^0(\theta) f_{\mathcal{Z}}(\theta) \boldsymbol{\delta}_i^0 \\
&= \frac{1}{N^{\alpha^0}} \sum_{i=1}^{N^{\alpha^0}} \boldsymbol{\delta}_i^{0\prime} \mathbf{D}_{\mathbf{f}}^0(\theta) f_{\mathcal{Z}}(\theta) \boldsymbol{\delta}_i^0 + \frac{1}{N^{\alpha^0}} \sum_{i=N^{\alpha^0}+1}^{N} \boldsymbol{\delta}_i^{0\prime} \mathbf{D}_{\mathbf{f}}^0(\theta) f_{\mathcal{Z}}(\theta) \boldsymbol{\delta}_i^0
\end{aligned}$$

by (29) (the sign is reversed if $\theta > \theta^0$). By Assumptions CR(c) and CR(d), $\partial \mathrm{E}\left[w^0\left(\alpha^0, \theta\right)\right]/\partial\theta$ is continuous at $\theta = \theta^0$, and $\partial \mathrm{E}\left[w^0\left(\alpha^0, \theta^0\right)\right]/\partial\theta > 0$, respectively: there then exists a $B$ small enough such that

$$d = \min_{\left|\theta - \theta^0\right| \leq B} \frac{\partial \mathrm{E}\left[w^0\left(\alpha^0, \theta\right)\right]}{\partial\theta} > 0.$$

The first-order Taylor expansion of $\mathrm{E}\left[w^0\left(\alpha^0,\theta\right)\right]$ about $\theta=\theta^0$ results in

$$\inf_{\left|\theta-\theta^0\right|\leq B}\mathrm{E}\left[w^0\left(\alpha^0,\theta\right)\right]\geq d\left|\theta-\theta^0\right|, \tag{30}$$

since $\mathrm{E}\left[w^0\left(\alpha^0,\theta^0\right)\right]=0$. Notice that

$$\begin{aligned}\mathrm{E}\left\{\left|w^0\left(\alpha^0,\theta\right)-\mathrm{E}\left[w^0\left(\alpha^0,\theta\right)\right]\right|^2\right\} &= \mathrm{E}\left\{\left|\frac{1}{N^{\alpha^0}}\frac{1}{T}\sum_{i=1}^{N}\sum_{T=1}^{T}\left\{w_{it}^0\left(\theta\right)-\mathrm{E}\left[w_{it}^0\left(\theta\right)\right]\right\}\right|^2\right\}\\ &\leq \frac{C_2}{N^{2\alpha^0}}\sum_{i=1}^{N}\mathrm{E}\left\{\left|\frac{1}{T}\sum_{T=1}^{T}\left\{w_{it}^0\left(\theta\right)-\mathrm{E}\left[w_{it}^0\left(\theta\right)\right]\right\}\right|^2\right\}\end{aligned}$$

for some $C_2<\infty$, and

$$\begin{aligned}\mathrm{E}\left\{\left|\frac{1}{T}\sum_{T=1}^{T}\left\{w_{it}^0\left(\theta\right)-\mathrm{E}\left[w_{it}^0\left(\theta\right)\right]\right\}\right|^2\right\} &\leq \left\|\boldsymbol{\delta}_i^0\right\|^4T^{-1}\mathrm{E}\left\{\left|\frac{1}{\sqrt{T}}\sum_{T=1}^{T}\left\{q_t^0\left(\theta,\theta^0\right)-\mathrm{E}\left[q_t^0\left(\theta,\theta^0\right)\right]\right\}\right|^2\right\}, \quad i=1,\ldots,N,\\ &\leq \left\|\boldsymbol{\delta}_i^0\right\|^4T^{-1}K\left|\theta-\theta^0\right|\end{aligned}$$

by Lemma A.5: since

$$\left\|\boldsymbol{\delta}_i^0\right\|=\left\|\boldsymbol{\lambda}_{2i}^0-\boldsymbol{\lambda}_{1i}^0\right\|\leq\left\|\boldsymbol{\lambda}_{1i}^0\right\|+\left\|\boldsymbol{\lambda}_{2i}^0\right\|\leq 2\bar{\lambda},\quad i=1,\ldots,N, \tag{31}$$

by Assumption C2, it follows that

$$\mathrm{E}\left\{\left|w^0\left(\alpha^0,\theta\right)-\mathrm{E}\left[w^0\left(\alpha^0,\theta\right)\right]\right|^2\right\}\leq\frac{C_2 16\bar{\lambda}^4}{N^{\left(2\alpha^0-1\right)}T}K\left|\theta-\theta^0\right|. \tag{32}$$

For any $\eta$ and $\varepsilon$, set

$$b=\frac{1-\eta/2}{1-\eta}>1 \tag{33}$$

and

$$\bar{v}=\frac{8C_2 16\bar{\lambda}^4K}{\eta^2d^2\left(1-1/b\right)^2\varepsilon}. \tag{34}$$

Assume $N$ and $T$ large enough so that $\bar{v}\Big/\left[N^{\left(2\alpha^0-1\right)}T\right]\leq B$, otherwise the lemma is trivially satisfied. For $l_N=1,\ldots,N+1$ and $l_T=1,\ldots,T+1$, set $\theta_{l_Nl_T}=\theta^0+\bar{v}b^{l_N-1}b^{l_T-1}\Big/\left[N^{\left(2\alpha^0-1\right)}T\right]$, where $N$ and $T$ are integers such that $\theta_{NT}-\theta^0=\bar{v}b^{N-1}b^{T-1}\Big/\left[N^{\left(2\alpha^0-1\right)}T\right]\leq B$, $\theta_{N+1,T}-\theta^0>B$ and $\theta_{N,T+1}-\theta^0>B$ (since $\bar{v}\Big/\left[N^{\left(2\alpha^0-1\right)}T\right]\leq B$ then $NT\geq 1$). By Markov's inequality, (30), (32) and (34),

$$\begin{aligned}\Pr\left\{\sup_{\substack{1\leq l_N\leq N,\\ 1\leq l_T\leq T}}\left|\frac{w^0\left(\alpha^0,\theta_{l_Nl_T}\right)}{\mathrm{E}\left[w^0\left(\alpha^0,\theta_{l_Nl_T}\right)\right]}-1\right|>\frac{\eta}{2}\right\} &\leq \left(\frac{2}{\eta}\right)^2\sum_{l_N=1}^{N}\sum_{l_T=1}^{T}\frac{\mathrm{E}\left\{\left|w^0\left(\alpha^0,\theta_{l_Nl_T}\right)-\mathrm{E}\left[w^0\left(\alpha^0,\theta_{l_Nl_T}\right)\right]\right|^2\right\}}{\left|\mathrm{E}\left[w^0\left(\alpha^0,\theta_{l_Nl_T}\right)\right]\right|^2}\\ &\leq \frac{4}{\eta^2}\sum_{l_N=1}^{N}\sum_{l_T=1}^{T}\frac{C_2N^{-\left(2\alpha^0-1\right)}T^{-1}16\bar{\lambda}^4K\left(\theta_{l_Nl_T}-\theta^0\right)}{d^2\left(\theta_{l_Nl_T}-\theta^0\right)^2}\\ &\leq \frac{4}{\eta^2}\frac{C_2 16\bar{\lambda}^4K}{d^2\bar{v}}\left(\sum_{l_N=0}^{\infty}\frac{1}{b^{l_N}}\right)\left(\sum_{l_T=0}^{\infty}\frac{1}{b^{l_T}}\right)\\ &= \frac{4}{\eta^2}\frac{C_2 16\bar{\lambda}^4K}{d^2\bar{v}}\frac{1}{\left(1-1/b\right)^2}\leq\frac{\varepsilon}{2}:\end{aligned}$$

it follows that for all $1\leq l_N\leq N$ and $1\leq l_T\leq T$, and with probability greater than $1-\varepsilon/2$,

$$\left|\frac{w^0\left(\alpha^0,\theta_{l_Nl_T}\right)}{\mathrm{E}\left[w^0\left(\alpha^0,\theta_{l_Nl_T}\right)\right]}-1\right|\leq\frac{\eta}{2}. \tag{35}$$

Using (33), for any $\theta$ such that $\bar{v} \Big/ \left[N^{\left(2\alpha^0-1\right)}T\right] \leq \left(\theta-\theta^0\right) \leq B$, there exists some $l_N \leq N$ and $l_T \leq T$ such that $\theta_{l_N l_T} < \theta < \min\left\{\theta_{l_{N+1},l_T}, \theta_{l_N,l_{T+1}}\right\}$ and on the event (35)

$$\frac{w^0\left(\alpha^0,\theta\right)}{\left(\theta-\theta^0\right)} \geq \frac{w^0\left(\alpha^0,\theta_{l_N l_T}\right)}{\mathrm{E}\left[w^0\left(\alpha^0,\theta_{l_N l_T}\right)\right]} \frac{\mathrm{E}\left[w^0\left(\alpha^0,\theta_{l_N l_T}\right)\right]}{\left[\min\left\{\theta_{l_{N+1},l_T}, \theta_{l_N,l_{T+1}}\right\}-\theta^0\right]} \geq \left(1-\frac{\eta}{2}\right)\frac{d\left(\theta_{l_N l_T}-\theta^0\right)}{\left[\min\left\{\theta_{l_{N+1},l_T}, \theta_{l_N,l_{T+1}}\right\}-\theta^0\right]} = \left(1-\eta\right)d$$

where we set $\left(\theta_{l_N l_T}-\theta^0\right) \Big/ \left[\min\left\{\theta_{l_{N+1},l_T}, \theta_{l_N,l_{T+1}}\right\}-\theta^0\right] = 1/b$: this event has probability greater than $1-\varepsilon/2$ and then

$$\Pr\left[\inf_{\frac{\bar{v}}{N^{\left(2\alpha^0-1\right)}T} \leq \left(\theta-\theta^0\right) \leq B} \frac{w^0\left(\alpha^0,\theta\right)}{\left(\theta-\theta^0\right)} < \left(1-\eta\right)d\right] \leq \frac{\varepsilon}{2},$$

holds. Taking the infimum over $-\bar{v} \Big/ \left[N^{\left(2\alpha^0-1\right)}T\right] \geq \left(\theta-\theta^0\right) \geq -B$ allows to prove a similar inequality using the same argument: this completes the proof of the lemma. ■

**Proof of Lemma A.7.** Given some $C_3 < \infty$ to be determined later, fix $\eta > 0$ and set

$$\bar{v} = \frac{8}{\left(0.5\right)^2\left(0.5\right)^2}\frac{C_1 C_3 \bar{\lambda}^2}{\eta^2 \varepsilon}. \tag{36}$$

For $l_N = 1,\ldots,N$ and $l_T = 1,\ldots,T$, set $\theta_{l_N l_T}-\theta^0 = \bar{v}2^{l_N-1}2^{l_T-1} \Big/ \left[N^{\left(2\alpha^0-1\right)}T\right] \leq B$. Markov's inequality, (20) in Lemma A.4, (31) and (36) ensure that

$$\begin{aligned}
\Pr\left[\sup_{\substack{1\leq l_N\leq N,\\ 1\leq l_T\leq T}} \frac{\left\|\mathbf{h}^0\left(\alpha^0,\theta_{l_N l_T}\right)-\mathbf{h}^0\left(\alpha^0,\theta^0\right)\right\|}{\left(\theta_{l_N l_T}-\theta^0\right)} > \eta\right] &\leq \frac{1}{\eta^2}\sum_{l_N=1}^{N}\sum_{l_T=1}^{T}\frac{\mathrm{E}\left[\left\|\mathbf{h}^0\left(\alpha^0,\theta_{l_N l_T}\right)-\mathbf{h}^0\left(\alpha^0,\theta^0\right)\right\|^2\right]}{\left(\theta_{l_N l_T}-\theta^0\right)^2}\\
&\leq \frac{1}{\eta^2}\sum_{l_N=1}^{N}\sum_{l_T=1}^{T}\frac{\mathrm{E}\left\{\left\|\frac{1}{N^{\alpha^0}}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\mathbb{I}_{2t}\left(\theta_{l_N l_T}\right)-\mathbb{I}_{2t}\left(\theta^0\right)\right]\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0 e_{it}\right\|^2\right\}}{\left(\theta_{l_N l_T}-\theta^0\right)^2}\\
&\leq \frac{C_3}{\eta^2}\frac{1}{N^{2\alpha^0}}\frac{1}{T^2}\sum_{l_N=1}^{N}\sum_{l_T=1}^{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\mathrm{E}\left\{\left\|\left[\mathbb{I}_{2t}\left(\theta_{l_N l_T}\right)-\mathbb{I}_{2t}\left(\theta^0\right)\right]\boldsymbol{\delta}_i^{0\prime}\mathbf{f}_t^0 e_{it}\right\|^2\right\}}{\left(\theta_{l_N l_T}-\theta^0\right)^2}\\
&\leq \frac{C_3}{\eta^2}\frac{1}{N^{2\alpha^0}}\frac{1}{T}\sum_{l_N=1}^{N}\sum_{l_T=1}^{T}\sum_{i=1}^{N}\frac{\left\|\boldsymbol{\delta}_i^0\right\|^2\mathrm{E}\left\{\left[\mathbb{I}_{2t}\left(\theta_{l_N l_T}\right)-\mathbb{I}_{2t}\left(\theta^0\right)\right]\left\|\mathbf{f}_t^0 e_{it}\right\|^2\right\}}{\left(\theta_{l_N l_T}-\theta^0\right)^2}\\
&\leq \frac{C_3}{\eta^2}\frac{1}{N^{\left(2\alpha^0-1\right)}}\frac{1}{T}\sum_{l_N=1}^{N}\sum_{l_T=1}^{T}\left[4\bar{\lambda}^2\frac{C_1\left(\theta_{l_N l_T}-\theta^0\right)}{\left(\theta_{l_N l_T}-\theta^0\right)^2}\right]\\
&= 4\frac{C_1 C_3\bar{\lambda}^2}{\eta^2\bar{v}}\left[\sum_{l_N=1}^{N}\frac{1}{\left(2^{l_N-1}\right)}\right]\left[\sum_{l_T=1}^{T}\frac{1}{\left(2^{l_T-1}\right)}\right]\\
&\leq \frac{4}{\left(0.5\right)^2\left(0.5\right)^2}\frac{C_1 C_3\bar{\lambda}^2}{\eta^2\bar{v}} \leq \frac{\varepsilon}{2}.
\end{aligned}$$

It follows that for all $1 \leq l_N \leq N$ and $1 \leq l_T \leq T$, and with probability greater than $1-\varepsilon/2$,

$$\frac{\left\|\mathbf{h}^0\left(\alpha^0,\theta_{l_N l_T}\right)-\mathbf{h}^0\left(\alpha^0,\theta^0\right)\right\|}{\left(\theta_{l_N l_T}-\theta^0\right)} \leq \eta,$$

which implies that

$$\Pr\left[\sup_{\frac{\bar{v}}{N^{\left(2\alpha^0-1\right)}T} \leq \left(\theta-\theta^0\right) \leq B} \frac{\left\|\mathbf{h}^0\left(\alpha^0,\theta\right)-\mathbf{h}^0\left(\alpha^0,\theta^0\right)\right\|}{\left(\theta-\theta^0\right)} > \eta\right] \leq \frac{\varepsilon}{2}.$$

Taking the infimum over $-\bar{v} \Big/ \left[N^{\left(2\alpha^0-1\right)}T\right] \geq \left(\theta-\theta^0\right) \geq -B$ allows to prove a similar inequality using the same argument, which completes the proof. ■

## A.3 Proof of the Result in Section 4

Given the loss function in (12) and for any fixed $R \geq 1$, let $\hat{\mathbf{\Lambda}}_j^R(\theta) = \left[\hat{\boldsymbol{\lambda}}_{j1}^R(\theta), \ldots, \hat{\boldsymbol{\lambda}}_{jN}^R(\theta)\right]'$ be the $N \times R$ matrix of estimated loadings for fixed $\theta$, for $j = 1, 2$. Let $\hat{\mathbf{V}}_j^R(\theta)$ be the $R \times R$ diagonal matrix of the first $R$ largest eigenvalues of $\hat{\mathbf{\Sigma}}_{j\mathbf{x}}(\theta)$ in (7) in decreasing order, for $j = 1, 2$. Define the $R^0 \times R$ rotation matrix

$$\hat{\mathbf{H}}_{jj}^R\left(\theta^0\right) = \frac{\mathbf{F}_j^0\left(\theta^0\right)\mathbf{F}_j^0\left(\theta^0\right)'}{T}\frac{\mathbf{\Lambda}_j^{0\prime}\hat{\mathbf{\Lambda}}_j^R\left(\theta^0\right)}{N}\hat{\mathbf{V}}_j^R\left(\theta^0\right)^{-1}, \quad j = 1, 2, \tag{37}$$

where $\mathbf{F}_j^0(\theta)$ is defined in Section 3.4.

**Lemma A.8** *For any fixed $R \geq 1$, there exists a $R^0 \times R$ matrix $\hat{\mathbf{H}}_{jj}^R\left(\theta^0\right)$ as defined in* (37)*, with* $\operatorname{rank}\left[\hat{\mathbf{H}}_{jj}^R\left(\theta^0\right)\right] = \min\left\{R^0, R\right\}$*, and $C_{NT} = \min\left\{\sqrt{N}, \sqrt{T}\right\}$, such that*

$$C_{NT}^2\left[\frac{1}{N}\sum_{i=1}^N\left\|\hat{\boldsymbol{\lambda}}_{ji}^R\left(\theta^0\right) - \hat{\mathbf{H}}_{jj}^R\left(\theta^0\right)'\boldsymbol{\lambda}_{ji}^0\right\|^2\right] = O_p(1), \quad j = 1, 2.$$

**Lemma A.9** *Let $\hat{\theta}^{\bar{R}}$ be the estimator for $\theta^0$ obtained from the loss function in* (12) *for any a priori chosen number of factors $R = \bar{R}$ such that $\bar{R} \geq R^0$. Then under assumptions I, C1-C4 and CR,*

$$N^{(2\alpha^0-1)}T\left(\hat{\theta}^{\bar{R}} - \theta^0\right) = O_p(1).$$

**Proof of Theorem 4.1.** Consider

$$\mathbf{x}_t = \mathbb{I}_{1t}\left(\theta^0\right)\mathbf{\Lambda}_1^0\mathbf{f}_t^0 + \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{\Lambda}_2^0\mathbf{f}_t^0 + \mathbf{e}_t = \mathbf{\Lambda}^0\left[\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^{0\prime}, \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}_t^{0\prime}\right]' + \mathbf{e}_t,$$

where $\mathbf{\Lambda}^0 = \left(\mathbf{\Lambda}_1^0, \mathbf{\Lambda}_2^0\right) = \left[\left(\boldsymbol{\lambda}_{11}^0, \ldots, \boldsymbol{\lambda}_{1N}^0\right)', \left(\boldsymbol{\lambda}_{21}^0, \ldots, \boldsymbol{\lambda}_{2N}^0\right)'\right] = \left(\boldsymbol{\lambda}_1^0, \ldots, \boldsymbol{\lambda}_N^0\right)'$ is a $R^0 \times 2R^0$ matrix, with $\boldsymbol{\lambda}_i^0 = \left(\boldsymbol{\lambda}_{1i}^{0\prime}, \boldsymbol{\lambda}_{2i}^{0\prime}\right)'$ a $2R^0 \times 1$ vector, and $\left[\mathbb{I}_{1t}\left(\theta^0\right)\mathbf{f}_t^{0\prime}, \mathbb{I}_{2t}\left(\theta^0\right)\mathbf{f}_t^{0\prime}\right]'$ is a $2R^0 \times 1$ vector. Given the loss function in (12), let $\hat{\mathbf{f}}_t^R(\theta)$ be the $R \times 1$ vector of estimated factors for fixed $\theta$, for $t = 1, \ldots, T$. Further, let $\hat{\mathbf{H}}_{jj}^{R+}\left(\theta^0\right)$ be the generalized inverse of $\hat{\mathbf{H}}_{jj}^R\left(\theta^0\right)$ in (37) such that $\hat{\mathbf{H}}_{jj}^R\left(\theta^0\right)\hat{\mathbf{H}}_{jj}^{R+}\left(\theta^0\right) = \mathbf{I}_R$, for $j = 1, 2$. Lemmas A.8 and A.9 imply that

$$C_{NT}^2\left[\frac{1}{N}\sum_{i=1}^N\left\|\hat{\boldsymbol{\lambda}}_{ji}^R\left(\hat{\theta}^{\bar{R}}\right) - \hat{\mathbf{H}}_{jj}^R\left(\theta^0\right)'\boldsymbol{\lambda}_{ji}^0\right\|^2\right] = O_p(1), \quad j = 1, 2,$$

so that

$$C_{NT}^2\left[\frac{1}{T}\sum_{t=1}^T\left\|\hat{\mathbf{f}}_t^R\left(\hat{\theta}^{\bar{R}}\right) - \left[\mathbb{I}_{1t}\left(\theta^0\right)\hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right) + \mathbb{I}_{2t}\left(\theta^0\right)\hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right)\right]\mathbf{f}_t^0\right\|^2\right] = O_p(1)$$

or

$$C_{NT}^2\left\{\frac{1}{T}\sum_{t=1}^T\left\|\begin{bmatrix}\mathbb{I}_{1t}\left(\hat{\theta}^{\bar{R}}\right)\hat{\mathbf{f}}_t^R\left(\hat{\theta}^{\bar{R}}\right)\\ \mathbb{I}_{2t}\left(\hat{\theta}^{\bar{R}}\right)\hat{\mathbf{f}}_t^R\left(\hat{\theta}^{\bar{R}}\right)\end{bmatrix} - \begin{bmatrix}\mathbb{I}_{1t}\left(\theta^0\right)\hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right)\mathbf{f}_t^0\\ \mathbb{I}_{2t}\left(\theta^0\right)\hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right)\mathbf{f}_t^0\end{bmatrix}\right\|^2\right\} = O_p(1),$$

which is analogous to Theorem 1 and Corollary 2 in Bai and Ng (2002): this is sufficient to complete the proof of the theorem. ■

**Proof of Lemma A.8.** The proof of Lemma A.8 is similar to that of Theorem 3.2 and omitted. ■

**Proof of Lemma A.9.** Given the loss function in (12) and following similar steps as in the proof of Theorem 3.3, it can be shown that

$$\lim_{N,T\to\infty} \mathrm{P}\left\{S\left[\hat{\mathbf{\Lambda}}^R(\theta), \hat{\mathbf{F}}^R(\theta), \theta\right] \le S\left[\hat{\mathbf{\Lambda}}^R(\theta^0), \hat{\mathbf{F}}^R(\theta^0), \theta^0\right]\right\} = 0, \quad \forall\theta \neq \theta^0, \quad R^0 \le R \le R^{\max}.$$

In order to prove the lemma it is then sufficient to show that

$$S\left[\hat{\mathbf{\Lambda}}^R(\theta^0), \hat{\mathbf{F}}^R(\theta^0), \theta^0\right] - S\left[\hat{\mathbf{\Lambda}}(\theta^0), \hat{\mathbf{F}}(\theta^0), \theta^0\right] = O_p\left(C_{NT}^{-2}\right)$$

for any fixed $R$ such that $R^0 \le R \le R^{\max}$, where $S\left[\hat{\mathbf{\Lambda}}(\theta), \hat{\mathbf{F}}(\theta), \theta\right] = S\left[\hat{\mathbf{\Lambda}}^{R^0}(\theta), \hat{\mathbf{F}}^{R^0}(\theta), \theta\right]$. Notice that

$$\begin{aligned}
& \left|S\left[\hat{\mathbf{\Lambda}}^R(\theta^0), \hat{\mathbf{F}}^R(\theta^0), \theta^0\right] - S\left[\hat{\mathbf{\Lambda}}(\theta^0), \hat{\mathbf{F}}(\theta^0), \theta^0\right]\right| \\
\le\ & \left|S\left[\hat{\mathbf{\Lambda}}^R(\theta^0), \hat{\mathbf{F}}^R(\theta^0), \theta^0\right] - S\left(\mathbf{\Lambda}^0, \mathbf{F}^0, \theta^0\right)\right| + \left|S\left(\mathbf{\Lambda}^0, \mathbf{F}^0, \theta^0\right) - S\left[\hat{\mathbf{\Lambda}}(\theta^0), \hat{\mathbf{F}}(\theta^0), \theta^0\right]\right| \\
\le\ & 2 \max_{R^0 \le R \le R^{\max}} \left|S\left[\hat{\mathbf{\Lambda}}^R(\theta^0), \hat{\mathbf{F}}^R(\theta^0), \theta^0\right] - S\left(\mathbf{\Lambda}^0, \mathbf{F}^0, \theta^0\right)\right| :
\end{aligned}$$

it therefore is sufficient to show that

$$S\left[\hat{\mathbf{\Lambda}}^R(\theta^0), \hat{\mathbf{F}}^R(\theta^0), \theta^0\right] - S\left(\mathbf{\Lambda}^0, \mathbf{F}^0, \theta^0\right) = O_p\left(C_{NT}^{-2}\right)$$

for each $R$ such that $R^0 \le R \le R^{\max}$. We have

$$\begin{aligned}
\mathbf{x}_t &= \mathbb{I}_{1t}(\theta^0)\mathbf{\Lambda}_1^0\mathbf{f}_t^0 + \mathbb{I}_{2t}(\theta^0)\mathbf{\Lambda}_2^0\mathbf{f}_t^0 + \mathbf{e}_t \\
&= \mathbb{I}_{1t}(\theta^0)\mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}^R(\theta^0)\hat{\mathbf{H}}_{11}^{R+}(\theta^0)\mathbf{f}_t^0 + \mathbb{I}_{2t}(\theta^0)\mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}^R(\theta^0)\hat{\mathbf{H}}_{22}^{R+}(\theta^0)\mathbf{f}_t^0 + \mathbf{e}_t,
\end{aligned}$$

where $\hat{\mathbf{H}}_{jj}^{R+}(\theta^0)$ is defined in the proof of Theorem 4.1, for $j = 1, 2$. This implies

$$\begin{aligned}
\mathbf{x}_t &= \mathbb{I}_{1t}(\theta^0)\hat{\mathbf{\Lambda}}_1^R(\theta^0)\hat{\mathbf{H}}_{11}^{R+}(\theta^0)\mathbf{f}_t^0 + \mathbb{I}_{2t}(\theta^0)\hat{\mathbf{\Lambda}}_2^R(\theta^0)\hat{\mathbf{H}}_{22}^{R+}(\theta^0)\mathbf{f}_t^0 \\
&\quad + \mathbf{e}_t - \mathbb{I}_{1t}(\theta^0)\left[\hat{\mathbf{\Lambda}}_1^R(\theta^0) - \mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}^R(\theta^0)\right]\hat{\mathbf{H}}_{11}^{R+}(\theta^0)\mathbf{f}_t^0 - \mathbb{I}_{2t}(\theta^0)\left[\hat{\mathbf{\Lambda}}_2^R(\theta^0) - \mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}^R(\theta^0)\right]\hat{\mathbf{H}}_{22}^{R+}(\theta^0)\mathbf{f}_t^0 \\
&= \mathbb{I}_{1t}(\theta^0)\hat{\mathbf{\Lambda}}_1^R(\theta^0)\hat{\mathbf{H}}_{11}^{R+}(\theta^0)\mathbf{f}_t^0 + \mathbb{I}_{2t}(\theta^0)\hat{\mathbf{\Lambda}}_2^R(\theta^0)\hat{\mathbf{H}}_{22}^{R+}(\theta^0)\mathbf{f}_t^0 + \mathbf{u}_t,
\end{aligned}$$

where

$$\mathbf{u}_t = \mathbf{e}_t - \mathbb{I}_{1t}(\theta^0)\left[\hat{\mathbf{\Lambda}}_1^R(\theta^0) - \mathbf{\Lambda}_1^0\hat{\mathbf{H}}_{11}^R(\theta^0)\right]\hat{\mathbf{H}}_{11}^{R+}(\theta^0)\mathbf{f}_t^0 - \mathbb{I}_{2t}(\theta^0)\left[\hat{\mathbf{\Lambda}}_2^R(\theta^0) - \mathbf{\Lambda}_2^0\hat{\mathbf{H}}_{22}^R(\theta^0)\right]\hat{\mathbf{H}}_{22}^{R+}(\theta^0)\mathbf{f}_t^0.$$

Notice that

$$S\left(\mathbf{\Lambda}^0, \mathbf{F}^0, \theta^0\right) = (NT)^{-1}\sum_{t=1}^{T}\mathbf{e}_t'\mathbf{e}_t$$

and

$$
\begin{aligned}
& S\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \\
= \;& (NT)^{-1} \sum_{t=1}^{T} \mathbf{u}_t' \mathbf{u}_t \\
= \;& (NT)^{-1} \sum_{t=1}^{T} \mathbf{e}_t' \mathbf{e}_t \\
& -2(NT)^{-1} \sum_{t=1}^{T} \mathbf{f}_t^{0\prime} \left\{ \begin{array}{c} \mathbb{I}_{1t}\left(\theta^0\right) \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right)' \left[\hat{\mathbf{\Lambda}}_1^R\left(\theta^0\right) - \mathbf{\Lambda}_1^0 \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)\right]' \\ +\mathbb{I}_{2t}\left(\theta^0\right) \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right)' \left[\hat{\mathbf{\Lambda}}_2^R\left(\theta^0\right) - \mathbf{\Lambda}_2^0 \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)\right]' \end{array} \right\} \mathbf{e}_t \\
& +(NT)^{-1} \sum_{t=1}^{T} \mathbf{f}_t^{0\prime} \left\{ \begin{array}{c} \mathbb{I}_{1t}\left(\theta^0\right) \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right)' \left[\hat{\mathbf{\Lambda}}_1^R\left(\theta^0\right) - \mathbf{\Lambda}_1^0 \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)\right]' \left[\hat{\mathbf{\Lambda}}_1^R\left(\theta^0\right) - \mathbf{\Lambda}_1^0 \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)\right] \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right) \\ +\mathbb{I}_{2t}\left(\theta^0\right) \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right)' \left[\hat{\mathbf{\Lambda}}_2^R\left(\theta^0\right) - \mathbf{\Lambda}_2^0 \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)\right]' \left[\hat{\mathbf{\Lambda}}_2^R\left(\theta^0\right) - \mathbf{\Lambda}_2^0 \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)\right] \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right) \end{array} \right\} \mathbf{f}_t^0 \\
= \;& S\left(\mathbf{\Lambda}^0, \mathbf{F}^0, \theta^0\right) + S^{(1)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] + S^{(2)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right]
\end{aligned}
$$

so that

$$
\begin{aligned}
\left| S\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] - S\left(\mathbf{\Lambda}^0, \mathbf{F}^0, \theta^0\right) \right| &= \left| S^{(1)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] + S^{(2)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \right| \\
&\le \left| S^{(1)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \right| + \left| S^{(2)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \right|.
\end{aligned}
$$

For any $A \times A$ matrix $\mathbf{A}$, $|\mathrm{tr}(\mathbf{A})| \le A\|\mathbf{A}\|$, where $\mathbf{A}$ denotes the trace operator. It follows

$$
\begin{aligned}
& \left| S^{(1)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \right| \\
= \;& \left| \mathrm{tr}\left\{ S^{(1)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \right\} \right| \\
= \;& 2 \left| \mathrm{tr}\left\{ (NT)^{-1} \sum_{t=1}^{T} \mathbf{f}_t^{0\prime} \left\{ \begin{array}{c} \mathbb{I}_{1t}\left(\theta^0\right) \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right)' \left[\hat{\mathbf{\Lambda}}_1^R\left(\theta^0\right) - \mathbf{\Lambda}_1^0 \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)\right]' \\ +\mathbb{I}_{2t}\left(\theta^0\right) \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right)' \left[\hat{\mathbf{\Lambda}}_2^R\left(\theta^0\right) - \mathbf{\Lambda}_2^0 \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)\right]' \end{array} \right\} \mathbf{e}_t \right\} \right| \\
\le \;& 2R \left[ \begin{array}{c} \left\| \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right) \right\| \left\| \dfrac{\hat{\mathbf{\Lambda}}_1^R\left(\theta^0\right) - \mathbf{\Lambda}_1^0 \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)}{\sqrt{N}} \right\| \left\| \dfrac{1}{\sqrt{NT}} \sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right) \mathbf{e}_t \mathbf{f}_t^{0\prime} \right\| \\ + \left\| \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right) \right\| \left\| \dfrac{\hat{\mathbf{\Lambda}}_2^R\left(\theta^0\right) - \mathbf{\Lambda}_2^0 \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)}{\sqrt{N}} \right\| \left\| \dfrac{1}{\sqrt{NT}} \sum_{t=1}^{T} \mathbb{I}_{2t}\left(\theta^0\right) \mathbf{e}_t \mathbf{f}_t^{0\prime} \right\| \end{array} \right] \\
\le \;& 2R \left\{ \begin{array}{c} \left\| \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right) \right\| \left[ \dfrac{1}{N} \sum_{i=1}^{N} \left\| \hat{\boldsymbol{\lambda}}_{1i}^R\left(\theta^0\right) - \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)' \boldsymbol{\lambda}_{1i}^0 \right\|^2 \right]^{1/2} \dfrac{1}{\sqrt{T}} \left[ \dfrac{1}{N} \sum_{i=1}^{N} \left\| \dfrac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right) e_{it} \mathbf{f}_t^0 \right\|^2 \right]^{1/2} \\ + \left\| \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right) \right\| \left[ \dfrac{1}{N} \sum_{i=1}^{N} \left\| \hat{\boldsymbol{\lambda}}_{2i}^R\left(\theta^0\right) - \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)' \boldsymbol{\lambda}_{2i}^0 \right\|^2 \right]^{1/2} \dfrac{1}{\sqrt{T}} \left[ \dfrac{1}{N} \sum_{i=1}^{N} \left\| \dfrac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathbb{I}_{2t}\left(\theta^0\right) e_{it} \mathbf{f}_t^0 \right\|^2 \right]^{1/2} \end{array} \right\} \\
= \;& O_p\left(C_{NT}^{-1}\right) \frac{1}{\sqrt{T}} + O_p\left(C_{NT}^{-1}\right) \frac{1}{\sqrt{T}} = O_p\left(C_{NT}^{-2}\right)
\end{aligned}
$$

by Assumption C.4 and Lemma A.8. Further,

$$
\begin{aligned}
\left| S^{(2)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \right| &= S^{(2)}\left[\hat{\mathbf{\Lambda}}^R\left(\theta^0\right), \hat{\mathbf{F}}^R\left(\theta^0\right), \theta^0\right] \\
&\le \left[ \frac{1}{N} \sum_{i=1}^{N} \left\| \hat{\boldsymbol{\lambda}}_{1i}^R\left(\theta^0\right) - \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)' \boldsymbol{\lambda}_{1i}^0 \right\|^2 \right] \left\| \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right) \right\|^2 \left[ \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{1t}\left(\theta^0\right) \left\| \mathbf{f}_t^0 \right\|^2 \right] \\
&\quad + \left[ \frac{1}{N} \sum_{i=1}^{N} \left\| \hat{\boldsymbol{\lambda}}_{2i}^R\left(\theta^0\right) - \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)' \boldsymbol{\lambda}_{2i}^0 \right\|^2 \right] \left\| \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right) \right\|^2 \left[ \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{2t}\left(\theta^0\right) \left\| \mathbf{f}_t^0 \right\|^2 \right] \\
&\le \left( \frac{1}{T} \sum_{t=1}^{T} \left\| \mathbf{f}_t^0 \right\|^2 \right) \left\{ \begin{array}{c} \left[ \dfrac{1}{N} \sum_{i=1}^{N} \left\| \hat{\boldsymbol{\lambda}}_{1i}^R\left(\theta^0\right) - \hat{\mathbf{H}}_{11}^R\left(\theta^0\right)' \boldsymbol{\lambda}_{1i}^0 \right\|^2 \right] \left\| \hat{\mathbf{H}}_{11}^{R+}\left(\theta^0\right) \right\|^2 \\ + \left[ \dfrac{1}{N} \sum_{i=1}^{N} \left\| \hat{\boldsymbol{\lambda}}_{2i}^R\left(\theta^0\right) - \hat{\mathbf{H}}_{22}^R\left(\theta^0\right)' \boldsymbol{\lambda}_{2i}^0 \right\|^2 \right] \left\| \hat{\mathbf{H}}_{22}^{R+}\left(\theta^0\right) \right\|^2 \end{array} \right\} \\
&= O_p(1) \cdot \left[ O_p\left(C_{NT}^{-2}\right) \cdot O_p(1) + O_p\left(C_{NT}^{-2}\right) \cdot O_p(1) \right] = O_p\left(C_{NT}^{-2}\right),
\end{aligned}
$$

which completes the proof of the lemma. ■

# References


[1] Acharya, V. V., L. H. Pedersen, T. Philippon and M. Richardson (2010), "Measuring Systemic Risk," Working Paper, New York University.

[2] Adrian, T. and M. K. Brunnermeier (2011), "CoVaR," NBER Working Paper 17454.

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

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

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

[6] 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.

[7] Baker, S. R., N. Bloom and S. J. Davis (2013), "Measuring Economic Policy Uncertainty", Working Paper, Northwestern University, Stanford and University of Chicago.

[8] Billio, M., M. Getmansky, A. W. Lo and L. Pelizzon (2012), "Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors," *Journal of Financial Economics*, 104 (3), 535 − 559.

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

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

[11] Chan, K. S. (1993), "Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model", *The Annals of Statistics*, 21(1), 520 − 533.

[12] Chan, K. S. and R. S. Tsay (1998), "Limiting Properties of the Least Squares Estimator of a Continuous Threshold Autoregressive Model," *Biometrika*, 85 (2), 413 − 426.

[13] Chen, L. (2015), "Estimating the Common Break Date in Large Factor Models," *Economics Letters*, 131, 70 – 74.

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

[15] Cheng, X., Z. Liao and F. Schorfheide (2015), "Shrinkage Estimation of High-Dimensional Factor Models with Structural Instabilities," NBER Working Paper 19792.

[16] 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.

[17] 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.

[18] 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.

[19] Diebold, F. X. and K. Yilmaz (2014), "On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms," *Journal of Econometrics*, 182 (1), 119 – 134.

[20] Doornik, J. A. (2012), *Ox 7: An Object-orientated Matrix Programming Language*, London: Timberlake Consultants Press.

[21] Engle, R. and B. Kelly (2012), "Dynamic Equicorrelation," *Journal of Business and Economic Statistics*, 30 (2), 212 – 228.

[22] Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000), "The Generalized Dynamic Factor Model: Identification and Estimation," *Review of Economics and Statistics*, 82 (4), 540 – 554.

[23] Forni, M., M. Hallin, M. Lippi and L. Reichlin (2004), "The Generalized Dynamic Factor Model: Consistency and Rates," *Journal of Econometrics*, 119 (2), 231 – 255.

[24] Forni, M., M. Hallin, M. Lippi and P. Zaffaroni (2015), "Dynamic Factor Models with Infinite-Dimensional Factor Spaces: One-Sided Representations," *Journal of Econometrics*, 185 (2), 359 – 371.

[25] Forni, M. and M. Lippi (2001), "The Generalized Dynamic Factor Model: Representation Theory," *Econometric Theory*, 17 (6), 1113 − 1141.

[26] Han, X. and A. Inoue (2014), "Tests for Parameter Instability in Dynamic Factor Models," *Econometric Theory*, Forthcoming.

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

[28] Hansen, B. E. (2000), "Sample Splitting and Threshold Estimation", *Econometrica*, 68 (3), 575−603.

[29] Hansen, B. E. (2011), "Threshold Autoregression in Economics," *Statistics and Its Interface*, 4 (2), 123 − 127.

[30] Jurado, K., S. C. Ludvigson and S. Ng (2014), "Meauring Uncertainty," *American Economic Review*, 105 (3), 1177 − 1216.

[31] Kapetanios, G. (2000), "Small Sample Properties of the Conditional Least Squares Estimator in SETAR Models", *Economics Letters*, 69 (3), 267 − 276.

[32] Ludvigson, S. C. and S. Ng (2007), "The Empirical Risk-Return Relation: A Factor Analysis Approach," *Journal of Financial Economics*, 83 (1), 171 − 222.

[33] Massacci, D. (2014), "Multivariate Regime Switching Model with Flexible Threshold Variable," Working Paper, EIEF.

[34] Massacci, D. (2015), "Estimation and Model Selection in Unstable Large Dimensional Factor Models," Unpublished Manuscript.

[35] Ng, S. and J. H. Wright (2013), "Facts and Challenges from the Great Recession for Forecasting and Macroeconomic Modeling," *Journal of Economic Literature*, 51 (4), 1120 − 1154.

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

[37] Ross, S. (1976), "The Arbitrage Theory of Capital Pricing," *Journal of Economic Theory*, 13 (3), 341 − 360.

[38] 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.

[39] Timmermann, A. (2008), "Elusive Return Predictability," *International Journal of Forecasting*, 24 (1), 1 − 18.

[40] Tong, H. and K. S. Lim (1980), "Threshold Autoregression, Limit Cycles and Cyclical Data", *Journal of the Royal Statistical Society. Series B (Methodological)*, 42 (3), 245 − 292.

[41] Tsay, R. S. (1989), "Testing and Modelling Threshold Autoregressive Processes," *Journal of the American Statistical Association*, 84 (405), 231 − 240.

[42] Tsay, R. S. (1998), "Testing and Modelling Multivariate Threshold Models," *Journal of the American Statistical Association*, 93 (443), 1188 − 1202.

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

**Table 1: Bias and RMSE in the case of the Estimator for $\theta^0 = 2$**

This table presents results from the Monte Carlo analysis of the estimator for $\theta^0 = 2$. The DGP is

$$x_{it}^s = \mathbb{I}\left(z_t^s \leq \theta^0\right)\lambda_{1i}^0 f_t^{0s} + \mathbb{I}\left(z_t^s > \theta^0\right)\lambda_{2i}^0 f_t^{0s} + e_{it}^s, \quad \theta^0 = 2, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T,$$

$$\lambda_{1i}^0 \sim \mathcal{N}(1,1), \quad \lambda_{2i}^0 = \lambda_{1i}^0 + \delta_i^0 \quad \delta_i^0 > 0, i = 1, \ldots, \left[N^{\alpha^0}\right], \quad \delta_i^0 = 0, i = \left[N^{\alpha^0}\right] + 1, \ldots, N,$$

$$z_t^s = \mu_z(1-\rho_z) + \rho_z z_{t-1}^s + \left(1-\rho_z^2\right)^{1/2}\epsilon_{zt}^s, \quad \mu_z = \theta^0 - \left[\mathbf{\Phi}^{-1}\left(\pi^0\right)\right], \quad \rho_z \sim \mathcal{U}(0.05, 0.95), \quad z_{-50}^s = \mu_z, \quad \epsilon_{zt}^s \sim \text{IID}\mathcal{N}(0,1), \quad t = -49, \ldots, 0, \ldots, T,$$

$$f_t^{0s} = \rho_f f_{t-1}^{0s} + \left(1-\rho_f^2\right)^{1/2}\epsilon_{ft}^s, \quad f_{-50}^{0s} = 0, \quad \epsilon_{ft}^s \sim \text{IID}\mathcal{N}(0,1), \quad \rho_f \sim \mathcal{U}(0.05, 0.95) \quad t = -49, \ldots, 0, \ldots, T,$$

$$e_{it}^s = \rho_e e_{i,t-1}^s + \sigma_{ii}^{1/2}\left(1-\rho_e^2\right)^{1/2}\epsilon_{e_it}^s, \quad e_{i,-50}^s = 0, \quad \epsilon_{e_it}^s \sim \text{IID}\mathcal{N}(0,1), \quad \rho_e \sim \mathcal{U}(0.05, 0.95), \quad \sigma_{ii} \sim \chi(1), \quad i = 1, \ldots, N, \quad t = -49, \ldots, 0, \ldots, T.$$

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A:** $\alpha^0 = 0.60$ | | | | | | | | | | | | | | | | | | | |
| $N$ | | 25 | | | | | | 50 | | | | | | 100 | | | | | |
| $\delta_i^0 > 0$ | | 0.25 | | 1.00 | | 1.75 | | 0.25 | | 1.00 | | 1.75 | | 0.25 | | 1.00 | | 1.75 | |
| $T$ | $\pi^0$ | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** |
| | 0.15 | 0.9150 | 1.3405 | 0.2161 | 0.6513 | 0.0814 | 0.4063 | 0.9549 | 1.3669 | 0.2911 | 0.7399 | 0.0911 | 0.4241 | 0.9873 | 1.3811 | 0.3132 | 0.7629 | 0.0740 | 0.3756 |
| | 0.30 | 0.4246 | 1.0102 | 0.0241 | 0.2342 | 0.0052 | 0.1045 | 0.4425 | 1.0173 | 0.0272 | 0.2378 | 0.0024 | 0.0835 | 0.4930 | 1.0327 | 0.0324 | 0.2671 | 0.0006 | 0.0468 |
| 100 | 0.50 | -0.0392 | 0.8826 | -0.0055 | 0.0937 | -0.0005 | 0.0193 | -0.0120 | 0.8937 | 0.0014 | 0.1016 | -0.0010 | 0.0203 | 0.0297 | 0.8692 | -0.0026 | 0.0996 | -0.0013 | 0.0157 |
| | 0.70 | -0.4549 | 1.0219 | -0.0260 | 0.1921 | -0.0033 | 0.0532 | -0.4567 | 1.0257 | -0.0268 | 0.2036 | -0.0023 | 0.0306 | -0.4207 | 0.9927 | -0.0226 | 0.1948 | -0.0027 | 0.0411 |
| | 0.85 | -0.9428 | 1.3580 | -0.1939 | 0.6016 | -0.0525 | 0.3078 | -0.9622 | 1.3654 | -0.2207 | 0.6257 | -0.0531 | 0.2895 | -0.9242 | 1.3287 | -0.2126 | 0.6186 | -0.0456 | 0.2837 |
| | 0.15 | 0.8239 | 1.2799 | 0.0274 | 0.2364 | 0.0011 | 0.0667 | 0.8501 | 1.2834 | 0.0322 | 0.2300 | 0.0014 | 0.0423 | 0.8556 | 1.2813 | 0.0385 | 0.2688 | 0.0014 | 0.0719 |
| | 0.30 | 0.3460 | 0.9152 | 0.0019 | 0.0343 | 0.0000 | 0.0077 | 0.3511 | 0.9240 | 0.0000 | 0.0233 | 0.0000 | 0.0088 | 0.3772 | 0.9131 | -0.0004 | 0.0183 | -0.0005 | 0.0075 |
| 200 | 0.50 | -0.0101 | 0.7709 | -0.0004 | 0.0156 | 0.0000 | 0.0070 | -0.0162 | 0.7795 | -0.0013 | 0.0194 | -0.0002 | 0.0067 | -0.0095 | 0.7425 | -0.0007 | 0.0113 | -0.0002 | 0.0048 |
| | 0.70 | -0.3480 | 0.9164 | -0.0005 | 0.0201 | -0.0006 | 0.0077 | -0.3867 | 0.9373 | -0.0011 | 0.0197 | -0.0006 | 0.0086 | -0.3584 | 0.8887 | -0.0005 | 0.0138 | -0.0003 | 0.0052 |
| | 0.85 | -0.8343 | 1.2934 | -0.0238 | 0.1989 | -0.0023 | 0.0668 | -0.8822 | 1.3156 | -0.0214 | 0.1892 | -0.0038 | 0.0921 | -0.8581 | 1.2718 | -0.0206 | 0.1900 | -0.0035 | 0.0807 |
| | 0.15 | 0.5965 | 1.0811 | 0.0002 | 0.0113 | 0.0000 | 0.0037 | 0.6958 | 1.1732 | 0.0006 | 0.0365 | 0.0000 | 0.0042 | 0.6310 | 1.0975 | 0.0000 | 0.0037 | 0.0000 | 0.0037 |
| | 0.30 | 0.1623 | 0.6257 | -0.0002 | 0.0059 | 0.0000 | 0.0023 | 0.2365 | 0.7310 | -0.0001 | 0.0060 | 0.0000 | 0.0029 | 0.1782 | 0.6745 | -0.0001 | 0.0022 | -0.0001 | 0.0022 |
| 400 | 0.50 | -0.0029 | 0.4862 | 0.0000 | 0.0047 | -0.0001 | 0.0028 | 0.0050 | 0.5616 | 0.0000 | 0.0055 | -0.0002 | 0.0028 | -0.0277 | 0.4889 | 0.0000 | 0.0030 | 0.0000 | 0.0030 |
| | 0.70 | -0.1864 | 0.6528 | -0.0002 | 0.0056 | 0.0000 | 0.0031 | -0.2261 | 0.7122 | -0.0002 | 0.0050 | 0.0000 | 0.0028 | -0.2155 | 0.6645 | 0.0000 | 0.0018 | 0.0000 | 0.0018 |
| | 0.85 | -0.5789 | 1.0520 | -0.0001 | 0.0088 | -0.0001 | 0.0030 | -0.6548 | 1.1254 | -0.0004 | 0.0094 | -0.0002 | 0.0035 | -0.6531 | 1.1211 | -0.0001 | 0.0033 | -0.0001 | 0.0033 |

**Table 1-*Continued*: Bias and RMSE in the case of the Estimator for $\theta^0 = 2$**

| Panel B: $\alpha^0 = 1.00$ | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | | 25 | | | | | | 50 | | | | | | 100 | | | | | |
| $\delta_i^0 > 0$ | | 0.25 | | 1.00 | | 1.75 | | 0.25 | | 1.00 | | 1.75 | | 0.25 | | 1.00 | | 1.75 | |
| $T$ | $\pi^0$ | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** |
| | 0.15 | 0.8011 | 1.2508 | 0.1563 | 0.5674 | 0.0830 | 0.4309 | 0.8011 | 1.2434 | 0.1246 | 0.5030 | 0.0568 | 0.3490 | 0.8183 | 1.2430 | 0.1060 | 0.4566 | 0.0400 | 0.2700 |
| | 0.30 | 0.3181 | 0.8998 | 0.0108 | 0.1838 | 0.0016 | 0.1051 | 0.3106 | 0.8534 | 0.0043 | 0.1181 | -0.0003 | 0.0549 | 0.3023 | 0.8004 | 0.0014 | 0.0822 | -0.0006 | 0.0107 |
| 100 | 0.50 | -0.0412 | 0.7581 | -0.0046 | 0.0398 | -0.0032 | 0.0254 | -0.0100 | 0.6852 | -0.0020 | 0.0269 | -0.0015 | 0.0169 | 0.0090 | 0.6311 | -0.0021 | 0.0229 | -0.0012 | 0.0119 |
| | 0.70 | -0.3929 | 0.9111 | -0.0144 | 0.1067 | -0.0057 | 0.0520 | -0.3216 | 0.8272 | -0.0034 | 0.0302 | -0.0016 | 0.0169 | -0.2636 | 0.7506 | -0.0036 | 0.0435 | -0.0012 | 0.0154 |
| | 0.85 | -0.8776 | 1.3035 | -0.0854 | 0.3633 | -0.0263 | 0.2003 | -0.7910 | 1.2272 | -0.0583 | 0.3104 | -0.0154 | 0.1481 | -0.7191 | 1.1512 | -0.0411 | 0.2629 | -0.0137 | 0.1505 |
| | 0.15 | 0.6257 | 1.1115 | 0.0104 | 0.1720 | 0.0000 | 0.0497 | 0.5451 | 1.0161 | 0.0047 | 0.0912 | -0.0001 | 0.0064 | 0.4424 | 0.9198 | 0.0023 | 0.0794 | 0.0000 | 0.0206 |
| | 0.30 | 0.1849 | 0.6676 | -0.0014 | 0.0160 | -0.0005 | 0.0082 | 0.1352 | 0.5678 | -0.0009 | 0.0110 | -0.0003 | 0.0052 | 0.0754 | 0.4277 | -0.0007 | 0.0089 | -0.0004 | 0.0055 |
| 200 | 0.50 | -0.0223 | 0.4862 | -0.0013 | 0.0122 | -0.0006 | 0.0091 | -0.0240 | 0.3901 | -0.0006 | 0.0077 | -0.0003 | 0.0041 | -0.0212 | 0.2801 | -0.0005 | 0.0066 | -0.0002 | 0.0039 |
| | 0.70 | -0.2023 | 0.6483 | -0.0009 | 0.0125 | -0.0006 | 0.0073 | -0.1518 | 0.5498 | -0.0009 | 0.0096 | -0.0006 | 0.0065 | -0.0937 | 0.4185 | -0.0004 | 0.0064 | -0.0003 | 0.0040 |
| | 0.85 | -0.6109 | 1.0854 | -0.0059 | 0.0762 | -0.0014 | 0.0147 | -0.5410 | 1.0166 | -0.0030 | 0.0651 | -0.0008 | 0.0091 | -0.4433 | 0.8927 | -0.0023 | 0.0553 | -0.0004 | 0.0081 |
| | 0.15 | 0.2969 | 0.7634 | -0.0004 | 0.0065 | -0.0004 | 0.0045 | 0.2172 | 0.6574 | -0.0002 | 0.0037 | -0.0001 | 0.0029 | 0.1304 | 0.5093 | 0.0000 | 0.0037 | 0.0000 | 0.0032 |
| | 0.30 | 0.0272 | 0.3044 | -0.0004 | 0.0052 | -0.0002 | 0.0027 | 0.0108 | 0.1950 | -0.0001 | 0.0032 | -0.0002 | 0.0029 | 0.0041 | 0.1136 | -0.0001 | 0.0030 | 0.0000 | 0.0021 |
| 400 | 0.50 | -0.0106 | 0.1988 | -0.0004 | 0.0042 | -0.0002 | 0.0027 | -0.0011 | 0.0990 | -0.0001 | 0.0031 | 0.0000 | 0.0020 | -0.0034 | 0.0397 | -0.0001 | 0.0019 | 0.0000 | 0.0011 |
| | 0.70 | -0.0406 | 0.2552 | -0.0003 | 0.0049 | -0.0002 | 0.0033 | -0.0157 | 0.1655 | -0.0001 | 0.0028 | 0.0000 | 0.0022 | -0.0054 | 0.0567 | -0.0002 | 0.0026 | -0.0001 | 0.0024 |
| | 0.85 | -0.2411 | 0.6635 | -0.0005 | 0.0053 | -0.0003 | 0.0036 | -0.1465 | 0.5026 | -0.0002 | 0.0043 | -0.0001 | 0.0019 | -0.0743 | 0.3356 | -0.0003 | 0.0048 | -0.0002 | 0.0034 |

**Table 2: MSE in the case of the Estimator for $c_{it}^{0s}$**

This table presents results from the Monte Carlo analysis of the estimator for

$$c_{it}^{0s} = \mathbb{I}\left(z_t^s \leq \theta^0\right) \lambda_{1i}^0 f_t^{0s} + \mathbb{I}\left(z_t^s > \theta^0\right) \lambda_{2i}^0 f_t^{0s}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

The DGP is described in Table 1.

| **Panel A: $\alpha^0 = 0.60$** | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Unfeasible Estimator** | | | | | | | | | **Feasible Estimator** | | | | | | | | |
| $N$ | | 25 | | | 50 | | | 100 | | | 25 | | | 50 | | | 100 | | |
| $\delta_i^0 > 0$ | | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 |
| $T$ | $\pi^0$ | | | | | | | | | | | | | | | | | | |
| | 0.15 | 0.0587 | 0.0548 | 0.0513 | 0.0510 | 0.0515 | 0.0516 | 0.0361 | 0.0364 | 0.0367 | 0.0696 | 0.0603 | 0.0534 | 0.0605 | 0.0575 | 0.0534 | 0.0423 | 0.0406 | 0.0378 |
| | 0.30 | 0.0593 | 0.0560 | 0.0531 | 0.0515 | 0.0518 | 0.0519 | 0.0366 | 0.0369 | 0.0371 | 0.0701 | 0.0584 | 0.0536 | 0.0608 | 0.0543 | 0.0523 | 0.0425 | 0.0381 | 0.0372 |
| 100 | 0.50 | 0.0595 | 0.0570 | 0.0549 | 0.0513 | 0.0515 | 0.0515 | 0.0367 | 0.0369 | 0.0370 | 0.0705 | 0.0585 | 0.0551 | 0.0609 | 0.0530 | 0.0516 | 0.0426 | 0.0375 | 0.0371 |
| | 0.70 | 0.0594 | 0.0577 | 0.0563 | 0.0513 | 0.0513 | 0.0512 | 0.0366 | 0.0367 | 0.0367 | 0.0706 | 0.0597 | 0.0566 | 0.0607 | 0.0532 | 0.0515 | 0.0425 | 0.0379 | 0.0367 |
| | 0.85 | 0.0594 | 0.0584 | 0.0574 | 0.0509 | 0.0508 | 0.0506 | 0.0362 | 0.0362 | 0.0361 | 0.0706 | 0.0632 | 0.0587 | 0.0604 | 0.0558 | 0.0518 | 0.0422 | 0.0393 | 0.0369 |
| | 0.15 | 0.0462 | 0.0424 | 0.0391 | 0.0366 | 0.0372 | 0.0374 | 0.0220 | 0.0225 | 0.0228 | 0.0518 | 0.0432 | 0.0392 | 0.0415 | 0.0379 | 0.0374 | 0.0251 | 0.0229 | 0.0228 |
| | 0.30 | 0.0464 | 0.0433 | 0.0405 | 0.0368 | 0.0372 | 0.0374 | 0.0222 | 0.0225 | 0.0228 | 0.0521 | 0.0435 | 0.0405 | 0.0416 | 0.0374 | 0.0374 | 0.0252 | 0.0226 | 0.0228 |
| 200 | 0.50 | 0.0467 | 0.0445 | 0.0424 | 0.0367 | 0.0370 | 0.0371 | 0.0222 | 0.0224 | 0.0226 | 0.0523 | 0.0446 | 0.0425 | 0.0416 | 0.0372 | 0.0371 | 0.0251 | 0.0224 | 0.0226 |
| | 0.70 | 0.0470 | 0.0456 | 0.0443 | 0.0365 | 0.0367 | 0.0367 | 0.0221 | 0.0222 | 0.0223 | 0.0526 | 0.0457 | 0.0443 | 0.0415 | 0.0368 | 0.0367 | 0.0251 | 0.0223 | 0.0223 |
| | 0.85 | 0.0470 | 0.0462 | 0.0455 | 0.0363 | 0.0363 | 0.0363 | 0.0219 | 0.0219 | 0.0219 | 0.0527 | 0.0467 | 0.0455 | 0.0413 | 0.0369 | 0.0363 | 0.0250 | 0.0222 | 0.0220 |
| | 0.15 | 0.0402 | 0.0365 | 0.0333 | 0.0292 | 0.0298 | 0.0301 | 0.0149 | 0.0153 | 0.0157 | 0.0430 | 0.0366 | 0.0333 | 0.0318 | 0.0298 | 0.0301 | 0.0164 | 0.0157 | 0.0157 |
| | 0.30 | 0.0404 | 0.0374 | 0.0347 | 0.0292 | 0.0297 | 0.0299 | 0.0149 | 0.0152 | 0.0155 | 0.0428 | 0.0374 | 0.0347 | 0.0316 | 0.0297 | 0.0299 | 0.0163 | 0.0155 | 0.0155 |
| 400 | 0.50 | 0.0406 | 0.0384 | 0.0365 | 0.0292 | 0.0295 | 0.0297 | 0.0149 | 0.0151 | 0.0153 | 0.0428 | 0.0385 | 0.0365 | 0.0313 | 0.0296 | 0.0297 | 0.0160 | 0.0153 | 0.0153 |
| | 0.70 | 0.0409 | 0.0395 | 0.0383 | 0.0291 | 0.0292 | 0.0293 | 0.0148 | 0.0150 | 0.0151 | 0.0433 | 0.0396 | 0.0383 | 0.0314 | 0.0293 | 0.0293 | 0.0161 | 0.0151 | 0.0151 |
| | 0.85 | 0.0410 | 0.0403 | 0.0397 | 0.0290 | 0.0291 | 0.0291 | 0.0147 | 0.0148 | 0.0148 | 0.0438 | 0.0403 | 0.0397 | 0.0316 | 0.0291 | 0.0291 | 0.0163 | 0.0148 | 0.0148 |

| **Panel B: $\alpha^0 = 1.00$** | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Unfeasible Estimator** | | | | | | | | | **Feasible Estimator** | | | | | | | | |
| $N$ | | 25 | | | 50 | | | 100 | | | 25 | | | 50 | | | 100 | | |
| $\delta_i^0 > 0$ | | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 |
| $T$ | $\pi^0$ | | | | | | | | | | | | | | | | | | |
| | 0.15 | 0.0588 | 0.0571 | 0.0563 | 0.0503 | 0.0495 | 0.0490 | 0.0359 | 0.0361 | 0.0364 | 0.0695 | 0.0609 | 0.0582 | 0.0604 | 0.0519 | 0.0500 | 0.0428 | 0.0377 | 0.0369 |
| | 0.30 | 0.0594 | 0.0578 | 0.0570 | 0.0509 | 0.0501 | 0.0497 | 0.0364 | 0.0366 | 0.0367 | 0.0700 | 0.0591 | 0.0575 | 0.0606 | 0.0507 | 0.0498 | 0.0425 | 0.0368 | 0.0367 |
| 100 | 0.50 | 0.0594 | 0.0581 | 0.0574 | 0.0509 | 0.0502 | 0.0498 | 0.0366 | 0.0366 | 0.0367 | 0.0701 | 0.0586 | 0.0576 | 0.0605 | 0.0504 | 0.0499 | 0.0422 | 0.0367 | 0.0367 |
| | 0.70 | 0.0591 | 0.0580 | 0.0575 | 0.0509 | 0.0503 | 0.0500 | 0.0365 | 0.0364 | 0.0364 | 0.0703 | 0.0589 | 0.0577 | 0.0607 | 0.0505 | 0.0501 | 0.0423 | 0.0366 | 0.0364 |
| | 0.85 | 0.0590 | 0.0579 | 0.0575 | 0.0506 | 0.0500 | 0.0497 | 0.0361 | 0.0359 | 0.0358 | 0.0707 | 0.0601 | 0.0581 | 0.0609 | 0.0513 | 0.0500 | 0.0427 | 0.0365 | 0.0360 |
| | 0.15 | 0.0466 | 0.0453 | 0.0447 | 0.0360 | 0.0353 | 0.0349 | 0.0220 | 0.0223 | 0.0225 | 0.0520 | 0.0458 | 0.0448 | 0.0408 | 0.0354 | 0.0349 | 0.0248 | 0.0223 | 0.0225 |
| | 0.30 | 0.0467 | 0.0456 | 0.0450 | 0.0363 | 0.0357 | 0.0353 | 0.0221 | 0.0223 | 0.0225 | 0.0516 | 0.0457 | 0.0450 | 0.0401 | 0.0357 | 0.0353 | 0.0240 | 0.0224 | 0.0225 |
| 200 | 0.50 | 0.0469 | 0.0460 | 0.0455 | 0.0364 | 0.0359 | 0.0356 | 0.0221 | 0.0222 | 0.0224 | 0.0511 | 0.0460 | 0.0455 | 0.0394 | 0.0359 | 0.0356 | 0.0235 | 0.0223 | 0.0224 |
| | 0.70 | 0.0470 | 0.0463 | 0.0460 | 0.0363 | 0.0359 | 0.0357 | 0.0220 | 0.0221 | 0.0221 | 0.0516 | 0.0464 | 0.0460 | 0.0399 | 0.0359 | 0.0357 | 0.0238 | 0.0221 | 0.0221 |
| | 0.85 | 0.0469 | 0.0463 | 0.0461 | 0.0361 | 0.0358 | 0.0357 | 0.0218 | 0.0218 | 0.0218 | 0.0524 | 0.0465 | 0.0461 | 0.0410 | 0.0359 | 0.0357 | 0.0246 | 0.0218 | 0.0218 |
| | 0.15 | 0.0407 | 0.0396 | 0.0390 | 0.0287 | 0.0280 | 0.0277 | 0.0148 | 0.0152 | 0.0155 | 0.0429 | 0.0396 | 0.0390 | 0.0302 | 0.0281 | 0.0277 | 0.0155 | 0.0152 | 0.0155 |
| | 0.30 | 0.0408 | 0.0399 | 0.0394 | 0.0288 | 0.0282 | 0.0279 | 0.0148 | 0.0151 | 0.0154 | 0.0421 | 0.0399 | 0.0394 | 0.0294 | 0.0282 | 0.0279 | 0.0150 | 0.0151 | 0.0154 |
| 400 | 0.50 | 0.0409 | 0.0402 | 0.0398 | 0.0289 | 0.0285 | 0.0282 | 0.0148 | 0.0150 | 0.0152 | 0.0419 | 0.0402 | 0.0398 | 0.0293 | 0.0285 | 0.0282 | 0.0150 | 0.0150 | 0.0152 |
| | 0.70 | 0.0410 | 0.0405 | 0.0402 | 0.0288 | 0.0286 | 0.0284 | 0.0148 | 0.0149 | 0.0150 | 0.0421 | 0.0405 | 0.0402 | 0.0294 | 0.0286 | 0.0284 | 0.0149 | 0.0149 | 0.0150 |
| | 0.85 | 0.0410 | 0.0407 | 0.0405 | 0.0289 | 0.0287 | 0.0286 | 0.0147 | 0.0147 | 0.0147 | 0.0430 | 0.0407 | 0.0405 | 0.0302 | 0.0287 | 0.0286 | 0.0153 | 0.0147 | 0.0147 |

**Table 3: Model Selection Criteria, $R^0 = 2$, $\alpha^0 = 0.60$**

This table presents results from the Monte Carlo analysis of the model selection criteria in (13). The DGP is

$$x^s_{it} = \mathbb{I}\left(z^s_t \leq \theta^0\right)\left(\lambda^0_{11i} f^{0s}_{1t} + \lambda^0_{12i} f^{0s}_{2t}\right) + \mathbb{I}\left(z^s_t > \theta^0\right)\left(\lambda^0_{21i} f^{0s}_{1t} + \lambda^0_{22i} f^{0s}_{2t}\right) + e^s_{it}, \quad \theta^0 = 2, \quad i = 1,\ldots,N, \quad t = 1,\ldots,T,$$

$$\lambda^0_{11i} \sim \mathcal{N}(1,1), \quad \lambda^0_{12i} \sim \mathcal{N}(1,1), \quad \lambda^0_{21i} = \lambda^0_{11i} + \delta^0_i, \quad \lambda^0_{22i} = \lambda^0_{12i} + \delta^0_i, \quad \delta^0_i > 0, i = 1,\ldots,\left[N^{\alpha^0}\right], \quad \delta^0_i = 0, i = \left[N^{\alpha^0}\right] + 1,\ldots,N,$$

$$z^s_t = \mu_z(1-\rho_z) + \rho_z z^s_{t-1} + \left(1-\rho^2_z\right)^{1/2} \epsilon^s_{zt}, \quad \mu_z = \theta^0 - \left[\mathbf{\Phi}^{-1}\left(\pi^0\right)\right], \quad \rho_z \sim \mathcal{U}(0.05, 0.95), \quad z^s_{-50} = \mu_z, \quad \epsilon^s_{zt} \sim \text{IID}\mathcal{N}(0,1), \quad t = -49,\ldots,0,\ldots,T,$$

$$f^{0s}_{jt} = \rho_f f^{0s}_{j,t-1} + \left(1-\rho^2_f\right)^{1/2} \epsilon^s_{jft}, \quad f^{0s}_{j,-50} = 0, \quad \epsilon^s_{jft} \sim \text{IID}\mathcal{N}(0,1), \quad \rho_f \sim \mathcal{U}(0.05, 0.95) \quad t = -49,\ldots,0,\ldots,T, \quad j = 1,2,$$

$$e^s_{it} = \rho_e e^s_{i,t-1} + \sigma^{1/2}_{ii}\left(1-\rho^2_e\right)^{1/2} \epsilon^s_{e_it}, \quad e^s_{i,-50} = 0, \quad \epsilon^s_{e_it} \sim \text{IID}\mathcal{N}(0,1), \quad \rho_e \sim \mathcal{U}(0.05, 0.95), \quad \sigma_{ii} \sim \chi(1), \quad i = 1,\ldots,N, \quad t = -49,\ldots,0,\ldots,T.$$

**Panel A: $IC_{p1}(R,R)$**

| | | **Unfeasible Estimator** | | | | | | | | | **Feasible Estimator** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | | 25 | | | 50 | | | 100 | | | 25 | | | 50 | | | 100 | | |
| $\delta^0_i > 0$ | | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 |
| $T$ | $\pi^0$ | | | | | | | | | | | | | | | | | | |
| | 0.15 | 2.7675 | 2.7925 | 2.8080 | 2.7680 | 2.7630 | 2.7540 | 2.0150 | 2.0150 | 2.0150 | 2.7950 | 2.7810 | 2.8860 | 2.9625 | 2.8950 | 3.0250 | 2.0180 | 2.0215 | 2.1640 |
| | 0.30 | 2.7830 | 2.8010 | 2.8160 | 2.8865 | 2.8800 | 2.8725 | 2.0150 | 2.0150 | 2.0160 | 2.7875 | 2.7875 | 2.8505 | 2.9550 | 2.8965 | 2.9725 | 2.0175 | 2.0200 | 2.1245 |
| 100 | 0.50 | 2.7835 | 2.7995 | 2.8070 | 2.9285 | 2.9230 | 2.9175 | 2.0145 | 2.0145 | 2.0145 | 2.7955 | 2.7935 | 2.8125 | 2.9750 | 2.9325 | 2.9280 | 2.0175 | 2.0155 | 2.0270 |
| | 0.70 | 2.7750 | 2.7840 | 2.7875 | 2.8990 | 2.8980 | 2.8980 | 2.0165 | 2.0165 | 2.0165 | 2.7855 | 2.7665 | 2.8120 | 2.9705 | 2.9275 | 3.0145 | 2.0180 | 2.0160 | 2.1035 |
| | 0.85 | 2.7585 | 2.7660 | 2.7675 | 2.7695 | 2.7690 | 2.7690 | 2.0130 | 2.0130 | 2.0130 | 2.7815 | 2.7535 | 2.8625 | 2.9650 | 2.9060 | 3.1285 | 2.0185 | 2.0165 | 2.2280 |
| | 0.15 | 2.9050 | 2.9250 | 2.9375 | 2.9155 | 2.9035 | 2.8975 | 2.0105 | 2.0105 | 2.0105 | 2.9060 | 2.9115 | 2.9450 | 2.9750 | 2.9180 | 2.9795 | 2.0115 | 2.0125 | 2.1225 |
| | 0.30 | 2.9035 | 2.9205 | 2.9295 | 2.9510 | 2.9435 | 2.9375 | 2.0125 | 2.0130 | 2.0130 | 2.9045 | 2.9190 | 2.9315 | 2.9750 | 2.9470 | 2.9455 | 2.0120 | 2.0120 | 2.0245 |
| 200 | 0.50 | 2.9075 | 2.9195 | 2.9235 | 2.9755 | 2.9630 | 2.9600 | 2.0130 | 2.0135 | 2.0135 | 2.9060 | 2.9185 | 2.9235 | 2.9785 | 2.9655 | 2.9625 | 2.0140 | 2.0135 | 2.0135 |
| | 0.70 | 2.9085 | 2.9140 | 2.9190 | 2.9455 | 2.9400 | 2.9415 | 2.0145 | 2.0150 | 2.0150 | 2.9075 | 2.9125 | 2.9210 | 2.9805 | 2.9475 | 2.9505 | 2.0125 | 2.0155 | 2.0260 |
| | 0.85 | 2.8985 | 2.9040 | 2.9050 | 2.9085 | 2.9050 | 2.9035 | 2.0120 | 2.0120 | 2.0115 | 2.9035 | 2.8925 | 2.9175 | 2.9720 | 2.9320 | 3.0320 | 2.0120 | 2.0110 | 2.1730 |
| | 0.15 | 2.9700 | 2.9780 | 2.9840 | 2.9760 | 2.9600 | 2.9540 | 2.0040 | 2.0040 | 2.0040 | 2.9680 | 2.9780 | 2.9835 | 2.9930 | 2.9600 | 2.9570 | 2.0050 | 2.0040 | 2.0190 |
| | 0.30 | 2.9670 | 2.9745 | 2.9795 | 2.9815 | 2.9740 | 2.9705 | 2.0060 | 2.0060 | 2.0060 | 2.9645 | 2.9745 | 2.9795 | 2.9910 | 2.9735 | 2.9705 | 2.0060 | 2.0060 | 2.0060 |
| 400 | 0.50 | 2.9655 | 2.9710 | 2.9765 | 2.9960 | 2.9820 | 2.9770 | 2.0040 | 2.0040 | 2.0040 | 2.9645 | 2.9710 | 2.9765 | 2.9980 | 2.9815 | 2.9915 | 2.0045 | 2.0040 | 2.0040 |
| | 0.70 | 2.9660 | 2.9695 | 2.9710 | 2.9995 | 2.9975 | 2.9910 | 2.0045 | 2.0045 | 2.0045 | 2.9660 | 2.9695 | 2.9710 | 3.0045 | 2.9980 | 2.9915 | 2.0050 | 2.0045 | 2.0045 |
| | 0.85 | 2.9650 | 2.9680 | 2.9680 | 2.9850 | 2.9835 | 2.9825 | 2.0045 | 2.0045 | 2.0045 | 2.9645 | 2.9670 | 2.9680 | 3.0010 | 2.9835 | 2.9880 | 2.0050 | 2.0045 | 2.0190 |

**Table 3-*Continued*: Model Selection Criteria, $R^0 = 2$, $\alpha^0 = 0.60$**

**Panel B: $IC_{p2}(R, R)$**

| | | Unfeasible Estimator | | | | | | | | | Feasible Estimator | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | | 25 | | | 50 | | | 100 | | | 25 | | | 50 | | | 100 | | |
| $\delta_i^0 > 0$ | | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 |
| $T$ | $\pi^0$ | | | | | | | | | | | | | | | | | | |
| 100 | 0.15 | 2.5940 | 2.6155 | 2.6275 | 2.2545 | 2.2555 | 2.2555 | 2.0000 | 2.0000 | 2.0000 | 2.6130 | 2.5935 | 2.6735 | 2.3175 | 2.2825 | 2.3700 | 2.0000 | 2.0000 | 2.0150 |
| | 0.30 | 2.6030 | 2.6175 | 2.6305 | 2.3010 | 2.3035 | 2.3060 | 2.0000 | 2.0000 | 2.0000 | 2.6080 | 2.6015 | 2.6505 | 2.3155 | 2.2965 | 2.3545 | 2.0000 | 2.0000 | 2.0175 |
| | 0.50 | 2.6035 | 2.6170 | 2.6295 | 2.3210 | 2.3215 | 2.3215 | 2.0000 | 2.0000 | 2.0000 | 2.6090 | 2.6150 | 2.6340 | 2.3130 | 2.3160 | 2.3240 | 2.0000 | 2.0000 | 2.0010 |
| | 0.70 | 2.6000 | 2.6100 | 2.6165 | 2.2940 | 2.2955 | 2.2940 | 2.0000 | 2.0000 | 2.0000 | 2.6030 | 2.5855 | 2.6315 | 2.3155 | 2.2985 | 2.3510 | 2.0000 | 2.0000 | 2.0165 |
| | 0.85 | 2.5830 | 2.5860 | 2.5895 | 2.2485 | 2.2480 | 2.2490 | 2.0005 | 2.0005 | 2.0005 | 2.6020 | 2.5545 | 2.6380 | 2.3115 | 2.2905 | 2.4310 | 2.0000 | 2.0000 | 2.0390 |
| 200 | 0.15 | 2.8245 | 2.8475 | 2.8620 | 2.5140 | 2.5110 | 2.5105 | 2.0000 | 2.0000 | 2.0000 | 2.8275 | 2.8340 | 2.8705 | 2.5570 | 2.5095 | 2.5695 | 2.0005 | 2.0000 | 2.0360 |
| | 0.30 | 2.8240 | 2.8445 | 2.8610 | 2.5395 | 2.5375 | 2.5390 | 2.0005 | 2.0005 | 2.0005 | 2.8250 | 2.8445 | 2.8605 | 2.5570 | 2.5355 | 2.5465 | 2.0005 | 2.0005 | 2.0070 |
| | 0.50 | 2.8290 | 2.8375 | 2.8475 | 2.5550 | 2.5510 | 2.5505 | 2.0005 | 2.0005 | 2.0005 | 2.8235 | 2.8370 | 2.8475 | 2.5585 | 2.5530 | 2.5515 | 2.0005 | 2.0005 | 2.0005 |
| | 0.70 | 2.8270 | 2.8380 | 2.8425 | 2.5305 | 2.5290 | 2.5265 | 2.0005 | 2.0005 | 2.0005 | 2.8210 | 2.8345 | 2.8425 | 2.5570 | 2.5260 | 2.5265 | 2.0005 | 2.0005 | 2.0040 |
| | 0.85 | 2.8270 | 2.8320 | 2.8355 | 2.5045 | 2.5045 | 2.5040 | 2.0005 | 2.0005 | 2.0005 | 2.8240 | 2.8145 | 2.8480 | 2.5510 | 2.5075 | 2.5780 | 2.0005 | 2.0005 | 2.0545 |
| 400 | 0.15 | 2.9480 | 2.9575 | 2.9650 | 2.7355 | 2.7280 | 2.7270 | 2.0005 | 2.0005 | 2.0005 | 2.9450 | 2.9570 | 2.9655 | 2.7390 | 2.7285 | 2.7300 | 2.0000 | 2.0005 | 2.0065 |
| | 0.30 | 2.9470 | 2.9585 | 2.9640 | 2.7420 | 2.7360 | 2.7335 | 2.0000 | 2.0000 | 2.0000 | 2.9460 | 2.9585 | 2.9640 | 2.7375 | 2.7355 | 2.7340 | 2.0000 | 2.0000 | 2.0000 |
| | 0.50 | 2.9450 | 2.9530 | 2.9560 | 2.7385 | 2.7375 | 2.7385 | 2.0000 | 2.0000 | 2.0000 | 2.9440 | 2.9530 | 2.9560 | 2.7405 | 2.7375 | 2.7435 | 2.0000 | 2.0000 | 2.0000 |
| | 0.70 | 2.9455 | 2.9465 | 2.9510 | 2.7455 | 2.7450 | 2.7435 | 2.0000 | 2.0000 | 2.0000 | 2.9425 | 2.9470 | 2.9510 | 2.7360 | 2.7450 | 2.7435 | 2.0000 | 2.0000 | 2.0000 |
| | 0.85 | 2.9425 | 2.9460 | 2.9485 | 2.7295 | 2.7290 | 2.7285 | 2.0000 | 2.0000 | 2.0000 | 2.9420 | 2.9460 | 2.9490 | 2.7455 | 2.7315 | 2.7335 | 2.0000 | 2.0000 | 2.0080 |

**Panel C: $IC_{p3}(R, R)$**

| | | Unfeasible Estimator | | | | | | | | | Feasible Estimator | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | | 25 | | | 50 | | | 100 | | | 25 | | | 50 | | | 100 | | |
| $\delta_i^0 > 0$ | | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 | 0.25 | 1.00 | 1.75 |
| $T$ | $\pi^0$ | | | | | | | | | | | | | | | | | | |
| 100 | 0.15 | 3.0615 | 3.0905 | 3.1125 | 5.6800 | 5.6675 | 5.6560 | 6.4015 | 6.3960 | 6.3990 | 3.0880 | 3.1230 | 3.3290 | 5.9090 | 5.9460 | 6.0755 | 7.9105 | 7.9095 | 7.9080 |
| | 0.30 | 3.0810 | 3.1110 | 3.1300 | 5.7920 | 5.7890 | 5.7805 | 7.5255 | 7.5225 | 7.5245 | 3.0860 | 3.1205 | 3.2355 | 5.9045 | 5.8710 | 5.9170 | 7.9120 | 7.8680 | 7.8455 |
| | 0.50 | 3.0840 | 3.1030 | 3.1145 | 5.8240 | 5.8175 | 5.8170 | 7.8165 | 7.8160 | 7.8105 | 3.0845 | 3.1035 | 3.1325 | 5.9020 | 5.8490 | 5.8395 | 7.9240 | 7.8685 | 7.8510 |
| | 0.70 | 3.0745 | 3.0820 | 3.0905 | 5.7865 | 5.7855 | 5.7860 | 7.5635 | 7.5625 | 7.5625 | 3.0815 | 3.0880 | 3.1905 | 5.9010 | 5.9020 | 5.9425 | 7.9150 | 7.8645 | 7.8325 |
| | 0.85 | 3.0525 | 3.0550 | 3.0570 | 5.6895 | 5.6905 | 5.6900 | 6.4745 | 6.4735 | 6.4765 | 3.0815 | 3.1005 | 3.3110 | 5.9140 | 5.9515 | 6.0945 | 7.9185 | 7.9025 | 7.9055 |
| 200 | 0.15 | 2.9865 | 2.9920 | 2.9975 | 5.1190 | 5.0990 | 5.0760 | 3.1615 | 3.1515 | 3.1460 | 2.9850 | 2.9890 | 3.0240 | 5.1605 | 5.1120 | 5.1680 | 3.3265 | 3.3240 | 3.4500 |
| | 0.30 | 2.9835 | 2.9880 | 2.9930 | 5.1175 | 5.0980 | 5.0770 | 3.2590 | 3.2545 | 3.2460 | 2.9840 | 2.9870 | 2.9990 | 5.1500 | 5.1005 | 5.0820 | 3.3300 | 3.2765 | 3.2710 |
| | 0.50 | 2.9840 | 2.9865 | 2.9915 | 5.1095 | 5.1020 | 5.0905 | 3.2865 | 3.2795 | 3.2790 | 2.9830 | 2.9865 | 2.9915 | 5.1365 | 5.1050 | 5.0920 | 3.3305 | 3.2840 | 3.2815 |
| | 0.70 | 2.9850 | 2.9865 | 2.9885 | 5.1275 | 5.1190 | 5.1155 | 3.2775 | 3.2740 | 3.2725 | 2.9835 | 2.9865 | 2.9930 | 5.1465 | 5.1220 | 5.1220 | 3.3355 | 3.2915 | 3.2855 |
| | 0.85 | 2.9805 | 2.9820 | 2.9830 | 5.1305 | 5.1275 | 5.1275 | 3.1560 | 3.1555 | 3.1560 | 2.9835 | 2.9800 | 3.0100 | 5.1540 | 5.1655 | 5.2290 | 3.3400 | 3.3020 | 3.4610 |
| 400 | 0.15 | 2.9965 | 2.9975 | 2.9990 | 4.4935 | 4.4380 | 4.3905 | 2.4585 | 2.4595 | 2.4600 | 2.9960 | 2.9975 | 2.9990 | 4.5125 | 4.4405 | 4.4015 | 2.4760 | 2.4620 | 2.4790 |
| | 0.30 | 2.9955 | 2.9975 | 2.9985 | 4.4915 | 4.4425 | 4.4050 | 2.4605 | 2.4615 | 2.4625 | 2.9950 | 2.9975 | 2.9985 | 4.5090 | 4.4465 | 4.4055 | 2.4730 | 2.4620 | 2.4625 |
| | 0.50 | 2.9945 | 2.9975 | 2.9980 | 4.5025 | 4.4815 | 4.4600 | 2.4670 | 2.4660 | 2.4660 | 2.9945 | 2.9975 | 2.9980 | 4.5170 | 4.4810 | 4.4845 | 2.4715 | 2.4650 | 2.4660 |
| | 0.70 | 2.9950 | 2.9955 | 2.9960 | 4.5145 | 4.5010 | 4.4845 | 2.4675 | 2.4670 | 2.4660 | 2.9945 | 2.9955 | 2.9960 | 4.5295 | 4.4990 | 4.4845 | 2.4785 | 2.4685 | 2.4665 |
| | 0.85 | 2.9950 | 2.9955 | 2.9955 | 4.5215 | 4.5165 | 4.5135 | 2.4630 | 2.4620 | 2.4635 | 2.9950 | 2.9955 | 2.9955 | 4.5335 | 4.5215 | 4.5190 | 2.4750 | 2.4690 | 2.4860 |

**Table 4: Empirical Application, 1985 - 2014**

This table presents results from the empirical application of the model in (1). The vector $\mathbf{x}_t$ is made of the 147 updated monthly financial variables employed in Jurado *et al.* (2015): all variables have zero mean and unit variance. The threshold variable $z_t$ is the lagged index of economic policy uncertainty proposed in Baker *et al.* (2013). The model is estimated over the period $1985:01 - 2014:12$, a total of 360 observations. $\hat{\theta}$ is the point estimate of the threshold parameter $\theta^0$ and $\hat{\pi} = T^{-1} \sum_{t=1}^{T} \mathbb{I}\left(z_t \leq \hat{\theta}\right)$. The optimal number of factors $\hat{R}$ is estimated according to the selection criteria $IC_{p1}(R,R)$, $IC_{p2}(R,R)$ and $IC_{p3}(R,R)$ in (13). The connectedness measures $C_1\left(\hat{R}\right)$ and $C_2\left(\hat{R}\right)$ are as in (17).

| | | | |
|---|---|---|---|
| $\hat{\theta}$ | 131.413 | | |
| $\hat{\pi}$ | 0.783 | | |
| $1-\hat{\pi}$ | 0.217 | | |
| | $IC_{p1}(R,R)$ | $IC_{p2}(R,R)$ | $IC_{p3}(R,R)$ |
| $\hat{R}$ | 3 | 3 | 6 |
| $C_1\left(\hat{R}\right)$ | 0.678 | 0.678 | 0.736 |
| $C_2\left(\hat{R}\right)$ | 0.865 | 0.865 | 0.898 |