Nonstationary Heterogeneous Panels with Multiple Structural Changes*

Badi H. Baltagi† Qu Feng ‡ Wei Wang§

25 Nov 2024

Abstract

Nonstationary panels have been widely used in empirical studies in economics, especially in macroeconomics and finance. This paper considers multiple structural changes in nonstationary heterogeneous panels with common factors. Kapetanios, Pesaran, Yamagata (2011) showed that unobserved nonstationary factors can be proxied by cross-sectional averages of observable data. This means that unobserved error factors can be treated as additional regressors, and different break points in slopes and error factor loadings can be considered as multiple breaks in linear regression models with panel data. **We** generalize the least squares approach by Bai and Perron (1998) to nonstationary panels and show that the break points in both slopes and error factor loadings can be consistently estimated for two important cases involving i) nonstationary factors and ii) nonstationary regressors considered by Phillips and Moon (1999). Monte Carlo simulations are conducted to study the performance of the main results in finite samples. **We illustrate our methods with an empirical example finding a significant change in the effect of international R&D spillovers on domestic total factor productivity in OECD countries in 1992, and we attribute it to the accelerated globalization starting from the early 1990s.**

Keywords: Nonstationary Panels, Multiple Structural Changes, Heterogeneity, Common Factors, Common Correlated Effects.

JEL Classification: C13, C23, C33, C38

*Earlier versions of this paper were presented at Monash University, the International Association of Applied Econometrics (IAAE 2023) Annual Conference in Oslo, The International Panel Data Conference (IPDC 2023) in Amsterdam, the Asian Meeting of the Econometric Society (2022) in Tokyo; Asian Meeting of the Econometric Society (2022) in China; SETA (2022) in Seoul. Constructive comments from the audience at these conferences are highly acknowledged. We would also like to thank two anonymous referees **as well as the editor Yuya Sasaki** for their extensive comments and suggestions. We dedicate this paper in honor of M. Hashem Pesaran for his many contributions to econometrics and panel data econometrics.

†Department of Economics and Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020, USA, Email Address: bbaltagi@syr.edu.

‡Economics Division, Nanyang Technological University, Singapore. Email Address: qfeng@ntu.edu.sg.

§School of Statistics and Mathematics, Shandong University of Finance and Economics, China. Email Address: wangwei_0115@outlook.com.

1 Introduction

Nonstationary panel data models allowing for cross-sectional dependence using a factor structure in the errors continue to be the focus of a lot of theoretical as well as empirical studies in econometrics. Hsiao (2018) provides a very detailed and insightful review of the main modeling and estimation approaches in the factor-augmented panel data literature. Feng and Kao (2020) also give a textbook treatment of this subject focusing on three main approaches for the factor-augmented panel data models. They include Pesaran's (2006) common correlated effects (CCE), Bai's (2009) iterated principal components (IPC), and the likelihood approaches proposed by Bai and Li (2014). More recently, the transformed approach developed by Hsiao, Shi, Zhou (2021) shows very good properties in dealing with error factors in panel data models.

This paper contributes to the literature of nonstationary panels with common factors by allowing for structural breaks in the slopes. It is motivated by Bai and Kao (2006) who consider a panel cointegration model with stationary factors, which are allowed to be correlated with the regressors. $\sqrt{n}T$-consistent fully modified (2sFM) estimators of the slope parameters are derived. In a panel cointegration model with nonstationary factors **considered by Bai, Kao and Ng (2009), factors are treated as parameters, and the dependent variable cointegrates with the regressors and factors. The IPC approach is applied to deal with unobserved factors, as in Bai (2009), and the $\sqrt{n}T$-consistent continuously updated bias-corrected (CupBC) and continuously updated fully modified (CupFM) estimators of the slope parameters are proposed.** Recently, Huang, Jin, and Su (2020) and Huang, Jin, Phillips, Su (2021) introduce the heterogeneity**,** modeled as a latent group structure in the slope parameters of the panel cointegration model with nonstationary factors, thus adding two features of heterogeneity and cross-section dependence in the nonstationary panel literature. A penalized principal component estimation, which is an iterative procedure between penalized regression and principal component analysis (PCA), is proposed to consistently estimate group membership and the slope parameters. Different from the homogeneous panel literature considered above, Kapetanios, Pesaran, and Yamagata (2011, KPY hereafter) estimate a model of *heterogeneous* panels with nonstationary factors. They find that the CCE approach proposed by Pesaran (2006) is still valid for $I(1)$ factors. In addition, Holly, Pesaran and Yamagata (2010) apply these methods to examine empirical features of the US housing markets.[1]

Following Huang et al. (2021) and Dong et al. (2021), this paper adds heterogeneity to the literature by considering multiple structural changes in the nonstationary panels with common factors. Specifically, we consider multiple breaks in the slopes and the error factor loadings in the heterogeneous panels with nonstationary regressors and factors. As such, this paper enriches the literature of nonstationary panels by accommodating two additional empirical features of multiple structural changes and cross-sectional dependence. As in Pesaran (2006) and KPY, unobserved nonstationary factors can be proxied by the cross-sectional averages of observable data. Thus, unobserved error factors can be treated as additional regressors, and different break points in slopes and error factor loadings can be considered as multiple breaks in linear regression models with panel data. Therefore, we generalize the least squares approach by Bai and Perron (1998) to nonstationary panels and show that the break points in both slopes and error factor loadings can be consistently estimated. In addition, different from KPY, we also consider the case of nonstationary regressors after the CCE transformation. This model can be considered as an extension of Phillips and Moon (1999, Section 5) to the case of allowing for an error factor structure and multiple breaks in slopes. Similarly, a T-consistent estimator of the heterogeneous slope parameters is obtained.

Estimation of structural breaks in panels has attracted a lot of attention since Bai's (2010) panel mean-shift model. Kim (2011) considers a common break in a deterministic trend model for large panels with nonstationary or stationary errors. Baltagi, Feng and Kao (2016, 2019, BFK hereafter) extend Pesaran's (2006) heterogeneous panels to the cases of common breaks in slopes with exogenous and endogenous regressors. **Baltagi, Kao and Wang (2015) apply Bai's (2009) IPC approach to deal with interactive fixed effects in the errors of a heterogeneous stationary panel with a common break in the slopes.** Baltagi, Kao and Liu (2017) look at the estimation of a break point in homogeneous nonstationary panels with only one regressor and no error factor structure. These models mainly focus on the case of a single common break. Li, Qian and Su (2017), Qian and Su (2016) propose the adaptive group fused LASSO (AGFL) in panels with multi-

[1] Dong, Gao and Peng (2021) propose a general model of nonstationary panels by considering varying-coefficient slopes and factor loadings.

ple breaks in slopes, with and without interactive effects, respectively.[2] Lumsdaine, Okui and Wang (2023) consider the estimation of panel group structure models with structural breaks. Kaddoura and Westerlund (2023) study panel data models with multiple structural breaks when T is fixed.[3]

Recently, Karavias, Narayan and Westerlund (2023) consider a single break in stationary homogeneous panels with interactive effects, and Ditzen, Karavias and Westerlund (2023) extend the analysis to the case of multiple breaks. Unlike these two papers, we focus on nonstationary heterogeneous panels and nonstationary factors with multiple breaks. In addition, multiple breaks in factor loadings are also considered in our paper. Thus, our model can be applied to empirical research using aggregate level data over a long period, e.g., the international R&D spillover model.

This paper is also related to the literature on structural instability in factor models since Stock and Watson (2009), and extensively studied by Breitung and Eickmeier (2011), Chen, Dolado and Gonzalo (2014), Yamamoto and Tanaka (2015), and Cheng, Liao and Schorfheide (2016). Recent advancements in this direction also include Baltagi, Kao and Wang (2017), Bai, Han and Shi (2020), and Duan, Bai and Han (2023), Baltagi, Kao and Wang (2021), Ma and Tu (2023).

The paper is organized as follows. Section 2 introduces the model of nonstationary panels with common factors and multiple structural changes in slopes and error factor loadings. Section 3 presents the main ideas for estimation. Asymptotic properties of the estimators are derived in Section 4. In Section 5, we consider the case of additional nonstationary components in regressors. Monte Carlo simulations are conducted in Section 6 and Section 7 displays an empirical application to international R&D spillovers. Section 8 provides concluding remarks. The mathematical proofs are relegated to the Appendix.

Notation: For any matrix or vector A, the Frobenius norm of A is defined as

[2] There have been important work on estimating and testing for multiple structural changes in the time series literature, including Bai (1997), Bai and Perron (1998, 2003), Qu and Perron (2007), Kejriwal and Perron (2008), Maheu and Song (2018), Oka and Perron (2018), Bergamelli et al. (2019), Pang et al. (2021), to name a few.

[3] **The multi-break homogeneous panel data model with fixed T considered by Kaddoura and Westerlund (2023) could be very useful in empirical studies. When T is fixed, the difference between stationary and nonstationary data is irrelevant for the proofs. Different from their model, we take a different approach by considering long panel with nonstationary data. Thus, we connect our paper with the nonstationary panels literature. When T is large, nonstationary data is treated differently from stationary data in the proofs. Consequently, the technical framework used is different, including assumptions, convergence rates and proofs.**

$\|A\| = \sqrt{tr(AA')}$. $(N, T) \to \infty$ denotes N and T tend to infinity simultaneously. $[\cdot]$ is the greatest integer function. Stochastic processes such as Brownian motion $W(r)$ on $[0, 1]$ are written as W, integrals such as $\int_c^d W(r)dr$ as $\int_c^d W$ and stochastic integrals $\int_c^d W(r)dW(r)$ as $\int_c^d WdW$. B_ω denotes the Brownian motion with covariance matrix Σ_ω. "$\Rightarrow$" denotes weak convergence.

2 Model

By extending Pesaran's (2006) influential framework to the nonstationary case, KPY (2011) consider the following heterogeneous panel regression with nonstationary factors:

$$y_{it} = x_{it}'\beta_i + \gamma_i' f_t + \varepsilon_{it},\ i = 1, ..., N;\ \ t = 1, ..., T, \tag{1}$$

where x_{it} is a $p \times 1$ vector of explanatory variables with heterogeneous slopes β_i, ε_{it} is the idiosyncratic error, independent of x_{it}, and γ_i is the corresponding loading vector.[4] The $q \times 1$ vector of unobserved factors f_t follow $I(1)$ processes,

$$f_t = f_{t-1} + \varphi_t, \tag{2}$$

φ_t is the idiosyncratic error. x_{it} follow an $I(1)$ processes under the Assumption of commonly correlated effects,

$$x_{it} = \Gamma_i' f_t + v_{it}, \tag{3}$$

where Γ_i is an $q \times p$ factor loading matrix. v_{it} is a $p \times 1$ vector of disturbances. Thus, y_{it} is also nonstationary. KPY show that the CCE approach is robust to nonstationary factors. v_{it} is assumed to be $I(0)$ as in KPY, in what we call Case 1 in this and the next section. Case 2 assumes v_{it} to be $I(1)$ and this is studied in Section 5.

This paper generalizes KPY's model (1) above by considering multiple breaks in β_i:

$$y_{it} = x_{it}'\beta_i(\mathcal{K}_0) + \gamma_i' f_t + \varepsilon_{it}, i = 1, .., N; t = 1, ..., T. \tag{4}$$

Common breaks in the slopes $\beta_i(\mathcal{K}_0)$ could arise due to technological progress or major policy shifts in a long time horizon. Assume there are m_0 breaks in the slope pa-

[4]The fixed effects model can be considered as a special case when the first component of x_{it} is 1 and the other components of the slope parameters β_i are homogeneous. We examine the performance of the break estimators in a fixed effects model in the Monte Carlo experiments.

rameters.[5] As in Bai and Perron (1998), $\mathcal{K}_0$ denotes an m_0-partition $(K_{0,1}, ..., K_{0,m_0})$, and the value of the slopes $\beta_i(\mathcal{K}_0)$ vary across $m_0 + 1$ different regimes,[6] i.e.,

$$\beta_i(\mathcal{K}_0) = \begin{cases} \beta_{i1}, & t = 1, ..., K_{0,1}, \\ \vdots \\ \beta_{i,m_0+1}, & t = K_{0,m_0} + 1, ..., T. \end{cases}$$

This model generalizes the analysis of stationary panels with a single break in slopes by BFK (2016, 2019) to nonstationary panels with multiple structural breaks. Thus, additional technical challenges are involved in the derivations of asymptotic properties of estimators with nonstationary data in the case of multiple breaks.

Similarly, factor loadings γ_i could also suffer from structural changes, often seen in the macroeconomic literature (Stock and Watson, 2009). Assume there are m_1 breaks in the error factor loadings with an m_1-partition $\mathcal{K}_1 = (K_{1,1}, ..., K_{1,m_1})$,

$$\gamma_i(\mathcal{K}_1) = \begin{cases} \gamma_{i1}, & t = 1, ..., K_{1,1}, \\ \vdots \\ \gamma_{i,m_1+1}, & t = K_{1,m_1} + 1, ..., T. \end{cases}$$

The model becomes

$$y_{it} = x_{it}'\beta_i(\mathcal{K}_0) + \gamma_i(\mathcal{K}_1)'f_t + \varepsilon_{it},\ i = 1, .., N;\ t = 1, ..., T. \tag{5}$$

In addition, the nonstationary f_t and x_{it} follow processes (2) and (3). Model (5) includes model (4) above as a special case. We suppress the superscript 0 in the true values of $\mathcal{K}_0$ and $\mathcal{K}_1$ for now. Breaks $\mathcal{K}_1$ in error factor loadings are allowed to have overlaps with breaks $\mathcal{K}_0$ in the slopes. Different from breaks $\mathcal{K}_0$ in slopes to model the changes in long-run structural relationship between y and x, breaks $\mathcal{K}_1$ in error loadings γ_i can be considered equivalent to the instability of the variance of errors $\gamma_i' f_t + \varepsilon_{it}$ in (4), or changes in the error factor variance with constant loadings.

In the special case of $m_0 = 2, m_1 = 1$, of model (5), we assume $K_{0,1} < K_{0,2} < K_{1,1}$, without loss of generality. Thus, three breaks $K_{0,1}, K_{0,2}, K_{1,1}$ split the sample into 4 regimes:

$$y_{it} = \begin{cases} x_{it}'\beta_{i1} + \gamma_{i1}' f_t + \varepsilon_{it}, & t = 1, ..., K_{0,1} \\ x_{it}'\beta_{i2} + \gamma_{i1}' f_t + \varepsilon_{it}, & t = K_{0,1} + 1, ..., K_{0,2} \\ x_{it}'\beta_{i3} + \gamma_{i1}' f_t + \varepsilon_{it} & t = K_{0,2} + 1, ..., K_{1,1} \\ x_{it}'\beta_{i3} + \gamma_{i2}' f_t + \varepsilon_{it}, & t = K_{1,1} + 1, ..., T, \end{cases} \tag{6}$$

[5] To accommodate the case of partial structural changes in the slopes considered by Bai and Perron (1998), $w_{it}'\alpha_i$ can be added to the right-hand side of (4) to denote the regressors and their corresponding slopes that are constant over time.

[6] In this paper, we assume common breaks for the individual series in the panel. Kim (2014) and Smith (2024) studied the case of heterogeneous breaks in the panel. However, to handle the unobserved error factor structure in the model, we follow KPY's CCE approach, which is not applicable to heterogeneous breaks.

each of which can be considered the same as KPY. This is also the case when there are multiple breaks in slopes and error factor loadings, i.e., $m_0 > 1$, $m_1 > 1$. We follow KPY and use the CCE approach to deal with the unobserved nonstationary factors f_t. In this model, the parameters to be estimated include the slopes $\beta_i(\mathcal{K}_0)$ and the break points $\mathcal{K}_0$, $\mathcal{K}_1$.

Like estimating break point $\mathcal{K}_0$ in slopes, estimating $\mathcal{K}_1$ in factor loadings is equally important. As pointed out in the growing literature since Stock and Watson (2009), the structural instability in the factor structure could have implications for the accuracy of forecasting and number of estimated factors. In our model (5), ignoring the break $\mathcal{K}_1$ in γ_i could bias the estimates of the factor loadings in empirical studies, e.g., US housing markets by Holly, Pesaran and Yamagata (2010). In addition, when the focus is on ε_{it}, e.g., testing for remaining cross-sectional dependence in ε_{it} (Juodis and Reese, 2022), estimating $\mathcal{K}_1$ is necessary for obtaining a consistent estimate of ε_{it}.

Compared with Bai, Kao and Ng's (2009) model of panel cointegration with nonstationary factors, our model (5) adds two new empirical features: heterogeneous slopes and structural breaks in slopes and factor loadings. Structural breaks here can be regarded as a different way of modeling parameter heterogeneity from the latent group structure considered by Huang et al. (2021). Besides, we apply the CCE approach to deal with unobserved factors, instead of the IPC approach used in the two papers above. In addition, different from BFK's (2016, 2019) models of a common structural break in heterogeneous panels with exogenous and endogenous regressors, this paper focuses on *multiple* breaks and *nonstationary* factors and regressors. In line with Bai, Kao and Ng (2009), f_t are treated as additional explanatory variables, instead of an error component in (5). Thus, $\mathcal{K}_0$ and $\mathcal{K}_1$ are considered as multiple breaks in a linear regression and are estimated by least squares as proposed by Bai and Perron (1998).

As in the literature on nonstationary panels with factors, the major challenge in estimating our model (5) lies in the unobserved factors. In this paper, we adopt the CCE approach proposed by Pesaran (2006) and examined by KPY in the case of nonstationary factors. To simplify the analysis, we follow Stock and Watson's (2016, p.429) idea of using the cross-sectional averages of x_{it}, $\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N} x_{it}$, instead of those of y_{it} and x_{it}, to proxy for f_t in this paper.[7] The cross-sectional average of x_{it}

[7] Karavias et al. (2023) use this proxy for f_t. BFK (2019) focus on estimating a single break

in (3),

$$\bar{x}_t = \bar{\Gamma}' f_t + \bar{v}_t,\ \bar{\Gamma} = \frac{1}{N}\sum_{i=1}^{N}\Gamma_i \text{ and } \bar{v}_t = \frac{1}{N}\sum_{i=1}^{N} v_{it}.$$

When $\bar{\Gamma}$ is of full rank ($q \leq p$), like OLS,

$$f_t = (\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}(\bar{x}_t - \bar{v}_t). \tag{7}$$

Since $\bar{v}_t \to 0$ in probability as $N \to \infty$, it is also asymptotically valid to use $\bar{x}_t$ as observable proxies for nonstationary f_t,

$$f_t - (\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}\bar{x}_t \xrightarrow{p} 0 \text{ as } N \to \infty. \tag{8}$$

Hence, the idea of CCE is being used for nonstationary factors in each regime.[8]

Using (7) for f_t, (5) can be written as

$$\begin{aligned} y_{it} &= x_{it}'\beta_i(\mathcal{K}_0) + f_t'\gamma_i(\mathcal{K}_1) + \varepsilon_{it} \\ &= x_{it}'\beta_i(\mathcal{K}_0) + [(\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}\,(\bar{x}_t - \bar{v}_t)]'\gamma_i(\mathcal{K}_1) + \varepsilon_{it} \\ &= x_{it}'\beta_i(\mathcal{K}_0) + \bar{x}_t'\gamma_i^*(\mathcal{K}_1) + \varepsilon_{it}^*, \end{aligned} \tag{9}$$

where $\underset{p\times 1}{\gamma_i^*(\mathcal{K}_1)} = \bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\underset{q\times 1}{\gamma_i(\mathcal{K}_1)}$ and $\varepsilon_{it}^* = \varepsilon_{it} - \bar{v}_t'\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_i(\mathcal{K}_1)$. Thus, by proxying f_t with observables, equation (9) can be regarded as a panel data regression with multiple common breaks $\mathcal{K}_0, \mathcal{K}_1$ in slopes β_i and γ_i^*. In the special case of no breaks $\mathcal{K}_1$ in the loadings of model (4), $\gamma_i^*(\mathcal{K}_1)$ in equation (9) becomes $\gamma_i^* = \bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\ \gamma_i$. In this paper, we consider the general model (5) and use least squares proposed by Bai and Perron (1998) to estimate break points $(\mathcal{K}_0, \mathcal{K}_1)$, slopes $\beta_i(\mathcal{K}_0)$ and their cross-sectional averages.

Remark 1: Breitung and Eickmeier (2011) point out that the structural breaks in the factor loadings can be captured by inflating the number of factors in the PCA estimation. However, the inflated number of factors may fail the rank condition required by the CCE approach above. This implies that using the cross-sectional averages does not necessarily capture the inflated number of factors. As shown in the next section, our estimator of $\mathcal{K}_0$ and $\beta_i(\mathcal{K}_0)$ can be robust to the breaks $\mathcal{K}_1$ in

point in heterogeneous slopes using the cross-sectional average (y_{it}, x_{it}) to proxy for f_t and treat the error factor structure as nuisance parameters. This paper also estimates break points in error factor loadings $\mathcal{K}_1$ along with $\mathcal{K}_0$. To simplify the analysis, we use the cross-sectional average x_{it} to proxy for f_t. In additional Monte Carlo simulations, we use the cross-sectional average (y_{it}, x_{it}) to proxy for f_t and similar results are obtained.

[8]As in KPY, when the rank condition holds, there is no need to estimate the number of error factors.

error factor structure in a simultaneous estimation approach. Identifying the breaks $\mathcal{K}_1$ can be separately achieved if the rank condition is satisfied with inflated number of factors.[9]

3 Estimation

To simplify notation, let $\underset{2p\times 1}{z_{it}} = (x_{it}', \bar{x}_t')'$, $\underset{2p\times 1}{\delta_i}(\mathcal{K}_0, \mathcal{K}_1) = (\beta_i(\mathcal{K}_0)', \gamma_i^*(\mathcal{K}_1)')'$. Thus, equation (9) above can be written as

$$y_{it} = z_{it}'\delta_i(\mathcal{K}_0, \mathcal{K}_1) + \varepsilon_{it}^*. \tag{10}$$

We rearrange the $m_0 + m_1$ breaks $\mathcal{K}_0, \mathcal{K}_1$ in time line as $\{\mathcal{K}^0\} = \{\mathcal{K}_0, \mathcal{K}_1\} = \{k_1^0, k_2^0, ..., k_m^0\}$ with $m = m_0 + m_1$. Superscript 0 denotes for true values of breaks. After reparameterization, model (10) can be considered as a panel data regression with multiple structural changes in slopes:

$$y_{it} = z_{it}'\delta_{ij} + \varepsilon_{it}^*, t = k_{j-1}^0 + 1, ..., k_j^0, \tag{11}$$

where $j = 1, ..., m+1$, and $k_0^0 = 0$, $k_{m+1}^0 = T$.

Remark 2: Equation (11) can be considered as a panel data version of the multiple structural change model considered by Bai and Perron (1998) using non-stationary data. It also extends the stationary panel data model with one common break in BFK (2016) to the case of multiple common breaks with nonstationary data.

Remark 3: The intuition on identifying break points in this literature apply here as well. First, as pointed out by Bai (1997) and Bai and Perron (1998), the key information to identify the break points in time series regressions depend on the break magnitude and the variance of the regressors relative to the variance of the errors. Second, in panels with mean shifts or (trend) stationary regressors, Bai (2010), Kim (2011) and BFK (2016) show that the break magnitude increases with N under the common break assumption. Thus the break point can be consistently estimated in panels as $(N, T) \to \infty$. Third, Baltagi, Kao and Liu (2017), Pang Du and Chong (2021) show that using nonstationary regressors, the variance of the regressors increases with T, implying that it is easier to identify break points in regressions using nonstationary regressors than stationary regressors.

[9] In this case, we can use partitioned regression to consistently estimate $\mathcal{K}_0$ and $\beta_{it}(\mathcal{K}_0)$ first when the rank condition is satisfied with a small number of factors. After $\hat{\mathcal{K}}_0$ and $\hat{\beta}_{it}(\hat{\mathcal{K}}_0)$ are obtained, PCA or other methods can be applied to identify the factor structure and the breaks in loadings in errors $f_t'\gamma_{it}(\mathcal{K}_1) + \varepsilon_{it}$ estimated by $y_{it} - x_{it}'\hat{\beta}_{it}(\hat{\mathcal{K}}_0)$.

Define $Y_i = (y_{i1}, \cdots, y_{iT})'$, $\underset{2p(m+1)\times 1}{\delta_i} = (\delta'_{i1}, ..., \delta'_{i,m+1})'$, $\underset{T\times[2p(m+1)]}{\underline{Z}_i(\mathcal{K}^0)} = diag(Z_{i1}, ..., Z_{i,m+1})$ with $Z_{ij} = (z_{i,k^0_{j-1}+1}, ..., z_{i,k^0_j})'$, $j = 1, ..., m+1$ and $\varepsilon^*_i = (\varepsilon^*_{i1}, \cdots, \varepsilon^*_{iT})'$. Thus, equation (11) can be written in matrix form: for $i = 1, ..., N$,

$$Y_i = \underline{Z}_i(\mathcal{K}^0)\delta_i + \varepsilon^*_i \tag{12}$$

For possible breaks $\mathcal{K} = m$-partition $(k_1, ..., k_m)$, the OLS estimator of δ_i is $\hat{\delta}_i(\mathcal{K}) = \left[\underline{Z}_i(\mathcal{K})'\underline{Z}_i(\mathcal{K})\right]^{-1} \underline{Z}_i(\mathcal{K})'Y_i$, and the corresponding sum of squared residuals is

$$SSR_i(\mathcal{K}) = \left[Y_i - \underline{Z}_i(\mathcal{K})\hat{\delta}_i(\mathcal{K})\right]' \left[Y_i - \underline{Z}_i(\mathcal{K})\hat{\delta}_i(\mathcal{K})\right].$$

Thus, the OLS estimator of $\mathcal{K}^0 = (k^0_1, ..., k^0_m)$ is defined as

$$\hat{\mathcal{K}} = (\hat{k}_1, ..., \hat{k}_m) = \arg\min_{(k_1,...,k_m)} \frac{1}{NT}\sum_{i=1}^{N} SSR_i(\mathcal{K}). \tag{13}$$

Due to the computation complexity $O_p(T^m)$ of the grid search algorithm, obtaining $(\hat{k}_1, ..., \hat{k}_m)$ by solving (13) is generally very time consuming when $m \geq 3$ and T is large. In practice, we recommend the dynamic programming algorithm proposed by Bai and Perron (2003).[10]

In this paper, we assume that m is known. This assumption can be relaxed by following the idea of sequential estimation based on parameter-consistancy tests by Bai and Perron (1998). Alternatively, m can be determined by an information criterion approach with a penalty factor related to m as in Boldea et al. (2020) who consider a fixed effects panel data model with multiple breaks.

Next, we consider the estimation of $\beta_i(\mathcal{K}_0)$. Denote $\underset{T\times p}{X_i} = (x_{i1}, \cdots, x_{iT})'$ and $\underset{T\times p}{\bar{X}} = (\bar{x}_1, \cdots, \bar{x}_T)'$. Stacking the time dimension of equation (9) in matrix form gives

$$Y_i = \begin{pmatrix} x'_{i1}\beta_{i1} \\ \vdots \\ x'_{iT}\beta_{i,m_0+1} \end{pmatrix} + \begin{pmatrix} \bar{x}'_1\gamma^*_{i1} \\ \vdots \\ \bar{x}'_T\gamma^*_{i,m_1+1} \end{pmatrix} + \varepsilon^*_i.$$

Reparameterizing $\underset{T\times[(m_0+1)p]}{\underline{X}_i(\mathcal{K}_0)} = diag\left(X_{i1}, X_{i2}, \cdots, X_{i,m_0+1}\right)$ with $\underset{K_{0,1}\times p}{X_{i1}} = (x_{i1}, ..., x_{i,K_{0,1}})'$, $\underset{(K_{0,2}-K_{0,1})\times p}{X_{i2}} = (x_{i,K_{0,1}+1}, ..., x_{i,K_{0,2}})', \cdots, \underset{(T-K_{0,m_0})\times p}{X_{i,m_0+1}} = (x_{i,K_{0,m_0}+1}, ..., x_{iT})'$ and $\underset{[(m_0+1)p]\times 1}{b_i} = (\beta'_{i1}, \cdots, \beta'_{i,m_0+1})'$ gives

$$Y_i = \underset{T\times[(m_0+1)p]}{\underline{X}_i(\mathcal{K}_0)}\underset{[(m_0+1)p]\times 1}{b_i} + \underset{T\times[(m_1+1)p]}{\overline{X}(\mathcal{K}_1)}\underset{[(m_1+1)p]\times 1}{g_i} + \varepsilon^*_i, \tag{14}$$

[10] In the simulations and empirical application, our selection range of breaks is $0.1T < k_1 < \cdots < k_m < 0.9T$. To avoid the singularity problem, we also impose the restriction of $\min_j k_j - k_{j-1} > p$ for $j = 2, \cdots m$.

where $\underset{T\times[(m_1+1)p]}{\overline{X}(\mathcal{K}_1)} = diag\left((\bar{x}_1', ..., \bar{x}_{K_{1,1}}')', \cdots, (\bar{x}_{K_{1,m_1}+1}', ..., \bar{x}_T')'\right)$ and $\underset{[(m_1+1)p]\times 1}{g_i} = (\gamma_{i1}^{*\prime}, \cdots, \gamma_{i,m_1+1}^{*\prime})'$.

In equation (14), we focus on the individual slopes b_i. Hence, we perform a partitioned regression that removes the second term $\underline{X}_i(\mathcal{K}_1)g_i$.[11] This partitioned regression on equation (14) yields:

$$\hat{b}_i = \hat{b}_i\left(\hat{\mathcal{K}}_0, \hat{\mathcal{K}}_1\right) = \left[\underline{X}_i(\hat{\mathcal{K}}_0)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} \underline{X}_i(\hat{\mathcal{K}}_0)\right]^{-1} \underline{X}_i\left(\hat{\mathcal{K}}_0\right)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} Y_i, \tag{15}$$

where $M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} = I - \underline{\overline{X}}(\hat{\mathcal{K}}_1)[\underline{\overline{X}}(\hat{\mathcal{K}}_1)'\underline{\overline{X}}(\hat{\mathcal{K}}_1)]^{-1}\underline{\overline{X}}(\hat{\mathcal{K}}_1)'$. Similarly, the mean of b_i can also be estimated consistently by the following mean-group estimator

$$\hat{b}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{b}_i = \frac{1}{N}\sum_{i=1}^{N}\left[\underline{X}_i(\hat{\mathcal{K}}_0)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} \underline{X}_i(\hat{\mathcal{K}}_0)\right]^{-1} \underline{X}_i(\hat{\mathcal{K}}_0)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} Y_i. \tag{16}$$

In the case of no breaks in error factors considered in equation (4), equation (14) reduces to

$$Y_i = \underset{T\times[(m_0+1)p]}{\underline{X}_i(\mathcal{K}_0)} \underset{[(m_0+1)p]\times 1}{b_i} + \underset{T\times p}{\bar{X}} \underset{p\times 1}{\gamma_i^*} + \varepsilon_i^*,$$

thus

$$\hat{b}_i = \hat{b}_i\left(\hat{\mathcal{K}}_0\right) = \left[\underline{X}_i(\hat{\mathcal{K}}_0)' M_{\bar{X}} \underline{X}_i(\hat{\mathcal{K}}_0)\right]^{-1} \underline{X}_i\left(\hat{\mathcal{K}}_0\right)' M_{\bar{X}} Y_i,$$

where $M_{\bar{X}} = I - \bar{X}\left(\bar{X}'\bar{X}\right)^{-1}\bar{X}'$ and the corresponding mean-group estimator becomes

$$\frac{1}{N}\sum_{i=1}^{N}\left[\underline{X}_i(\hat{\mathcal{K}}_0)' M_{\bar{X}} \underline{X}_i(\hat{\mathcal{K}}_0)\right]^{-1} \underline{X}_i\left(\hat{\mathcal{K}}_0\right)' M_{\bar{X}} Y_i.$$

The partitioned regression (15) suggests that the CCE transformed regressors $M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)$ become stationary after partialling out $I(1)$ f_t in the case of stationary v_{it}. This leads to $\sqrt{T}$-consistent $\hat{b}_i$ as shown in the next Section. By contrast, when v_{it} follows an $I(1)$ process, $M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)$ remains nonstationary. In this case, y_{it} and x_{it} are cointegrated after dealing with the unobserved factors in each regime, and T-consistency of $\hat{b}_i$ can be obtained. This is different from the setup in KPY. We will consider $I(0)$ v_{it} as Case 1 in Section 4, and $I(1)$ v_{it} as Case 2 in Section 5.

[11] If the structural breaks $\mathcal{K}_2$ exist in the loadings Γ_i, i.e., $x_{it} = \Gamma_{it}'(\mathcal{K}_2)f_t + v_{it}$ where $\Gamma_{it}(\mathcal{K}_2)$ is similarly defined as $\gamma_{it}(\mathcal{K}_1)$. It is still asymptotically valid to use $\bar{x}_t$ as proxies for the nonstationary f_t. We can simultaneously estimate the breaks $(\mathcal{K}_1, \mathcal{K}_2)$ as in Section 3, or we can ignore them if the focus is on estimating $\mathcal{K}_0$ and the slopes, and the rank condition still holds after using a bigger set of factors to represent breaks in factor loadings. For simplicity, we only consider the case of no breaks in Γ_i in this paper.

4 Main Results

4.1 Assumptions

The following Assumptions are needed for establishing the asymptotic properties of the breaks and slope estimators above.

Assumption 1 *For* $j = \{1, \cdots, m\}$, $k_j^0 = \left[\lambda_j^0 T\right]$ *with* $0 < \lambda_1^0 < \cdots < \lambda_m^0 < 1$.

Assumption 2 $Rank(\bar{\Gamma}) = q \leq p$.

Assumption 3 *Factor loadings* $\gamma_i\left(\mathcal{K}_1\right)$ *and* Γ_i *are independent and identically distributed (IID) across* i*, and independent of* ε_{jt}, v_{jt} *and* f_t *for all* i, j, t*. Assume*

$$\gamma_i(K_1) = \gamma(K_1) + \eta_i = \eta_i + \begin{cases} \gamma_1, & 1 \leq t < K_{1,1} \\ \vdots & \vdots \\ \gamma_{m_1+1}, & K_{m_1,1} + 1 < t \leq T \end{cases} \quad with\ \eta_i \sim IID(0, \Sigma_\gamma)$$

and $vec(\Gamma_i) = vec(\Gamma) + \xi_i$, $\xi_i \sim IID(0, \Omega_\xi)$, $i = 1, ...N$, *where the means* $\gamma(\mathcal{K}_1)$, Γ *are non-zero and fixed and the variances* Ω_η, Ω_ξ *are finite.*

Assumption 4 *For* $i = 1, ..., N$, $b_i = b + v_{b,i}$, $v_{b,i} \sim IID(0, \Sigma_b)$, *where* $b = \left(\beta_1', \beta_2', \cdots, \beta_{m_0+1}'\right)'$, $v_{b,i} = (v_{\beta_1,i}', v_{\beta_2,i}', \cdots v_{\beta_{m_0+1},i}')'$ *and* $\Sigma_b = diag(\Sigma_{\beta_1}, \Sigma_{\beta_2}, \cdots, \Sigma_{\beta_{m_0+1}})$ *for* $i = 1, 2, ..., N$*, where* $\|b\| < \infty$, $\|\Sigma_b\| < \infty$*, and the random deviations* $v_{b,i}$ *are independent of* x_{it} *and* ε_{jt} *for all* i, j *and* t.

Assumption 5 *In* ***the*** *nonstationary factor process* $f = f_{t-1} + \varphi_t$, φ_t *is* ***a vector of*** $L_{2+\vartheta}$ ***bounded process for some*** $\vartheta > 0$***,such that*** $E[||\varphi_t||^{2+\vartheta}] < \infty$, *and stationary near epoque dependent process of size* $1/2$*, on some* α*-mixing process of size* $-(2+\vartheta)/\vartheta$ *and independent of* v_{jt} *and* ε_{jt} *for all* i, j, t.

Assumption 6 $\lambda_0^0 = 1/T$ *and for* $j = \{1, \cdots, m\}$, $\lambda_j^0 \in (0,1)$, *and* $\lambda_{m+1}^0 = 1$. ***Matrices*** $\frac{1}{T^2}\sum_{t=[\lambda_j^0 T]}^{[\lambda_{j+1}^0 T]} f_t f_t'$ *and* $\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=[\lambda_j^0 T]}^{[\lambda_{j+1}^0 T]} z_{it} z_{it}'$ *have minimum eigenvalues bounded away from zero in probability.*

Assumption 7 *(i) The disturbances* $\varepsilon_{it}, i = 1, ..., N$*, are cross-sectionally independent; (ii) For each series* i*,* ε_{it} *is independent of* $\varphi_{t'}$ *for all* t *and* t'*; (iii) errors* ε_{is} *and* v_{jt} *are independent for all* i, j, s, t*; (iv)* ε_{it} *is a stationary process with absolute summable autocovariances, such that* $\varepsilon_{it} = \sum_{l=0}^{\infty} a_{il}\zeta_{i,t-l}$*, where* $\{\zeta_{it}, t = 1, ..., T\}$ *are IID random variables with zero mean and have a finite fourth-order moments.*

Assume $0 < Var(\varepsilon_{it}) = \sum_{l=0}^{\infty} a_{il}^2 = \sigma_i^2 < \infty$. *(v) for the* $T \times 1$ *vector* $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{i,T})'$, $Var(\varepsilon_i) = \Sigma_{\varepsilon,i}$ *and* $0 < \|\Sigma_{\varepsilon,i}\| < \infty$.

Assumption 8 *(i) The disturbances* $v_{it}, i = 1, ..., N$, *are cross-sectionally independent; (ii) For each series* i, v_{it} *is independent of* $\varphi_{t'}$ *for all* t *and* t'*; (iii)* v_{it} *are linear stationary processes with zero mean and absolute summable autocovariances,* $v_{it} = \sum_{l=0}^{\infty} \Xi_{il} \varrho_{i,t-l}$, *where* $(\zeta_{it}, \varrho_{it}')'$ *are* $(p+1) \times 1$ *vectors of IID random variables with variance-covariance matrix* I_{p+1} *and has a finite fourth-order moments, and* $Var(v_{it}) = \sum_{l=0}^{\infty} \Xi_{il} \Xi_{il}' = \Sigma_{v,i}$, *and* $0 < \|\Sigma_{v,i}\| < \infty$. *(iv)* $\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \Sigma_{v,i}$ *is nonsingular.*

For $j = \{1, \cdots, m\}$, define $\phi_{N,j} = \sum_{i=1}^{N} (\delta_{i,j+1} - \delta_{ij})'(\delta_{i,j+1} - \delta_{ij})$ in equation (11) as the magnitude of common breaks in panels.

Assumption 9 $\phi_{N,j} \to \infty$, $\frac{T}{N}\phi_{N,j} \to \infty$, *as* $(N, T) \to \infty$ *for* $j = \{1, \cdots, m\}$.

Assumption 1 is common in the time series and panel data literature of structural changes, e.g., Bai (1997), Bai and Perron (1998), Bai (2010), BFK (2016, 2019). It rules out the case that true breaks happen on the boundary of the observed time period. It also implies that there are sufficient number of observations between breaks for large sample approximation. However, Bai (2010) pointed out that the common breaks close to the boundary are allowed in a panel mean shift model when $T/N \to 0$. To simplify our proofs, we adopt this convenient assumption. We explore the performance of our break estimator in the case of boundary breaks in Monte Carlo experiments.

Assumption 2 on the rank condition guarantees that equation (7) is valid, see Pesaran (2006) and KPY who discuss the situation of rank deficiency. This assumption can be relaxed to accommodate more empirical situations. For example, Karabiyik, Urbain and Westerlund (2019) consider the case of $p < q$. When $p < q$, additional combinations of regressors (Karabiyik, Urbain and Westerlund, 2019) or additional exogenous variables (Bai and Ng, 2010) should be included to proxy the unobserved error factors. **Karabiyik, Reese and Westerlund (2017) provide a new analytical framework to address the problem that too many observables cause the second moment matrix of the estimated factors to become asymptotically singular.** Juodis, Karabiyik and Westerlund (2021) establish the theory of pooled CCE, while the true number of common factors can be larger than the number of estimated factors. Our theoretical results can be extended to the case of rank deficiency by following the papers mentioned above. We will explore the performance of the estimators in case that Assumption 2 is not satisfied in the Monte Carlo simulations.

In Assumption 3, we assume that Γ_i and γ_i are independent, so the regressors and the error factor loadings are uncorrelated. Different from Pesaran (2006), we use the cross-section averages of regressors only to proxy the unobserved factors f_t in this paper, thus γ_i does not appear in equation (7) above, implying that whether γ_i is correlated with Γ_i or not does not affect the rank condition. In addition, breaks $\mathcal{K}_1$ in γ_i in Assumption 3 will not affect the rank condition as well.

Assumptions 4, 5 **on the identification condition for the individual slopes are borrowed from KPY.** Under Assumptions 7 and 8, the idiosyncratic errors ε_{it} and v_{it} follow a general linear stationary process with heteroscedasticity and autocorrelation for each i. Assumption 9 specifies the relationship between T/N and the magnitude of breaks $\phi_{N,j}, j = 1, ..., m$. $\phi_{N,j}$ can grow slower or faster than N, depending on the relative rate of T/N. The condition on the magnitude of breaks in Assumption 9 generalizes Assumption 2 in stationary panels considered by BFK (2019) to the multiple breaks case.

Under these assumptions, we can show that the multiple breaks are estimated consistently, as summarized in the following theorem:

Theorem 1 *Under Assumptions 1-9,* $\lim_{(N,T)\to\infty} P\left(\hat{k}_j = k_j^0\right) = 1, j = \{1, \cdots, m\}$.

The rate of convergence and the distribution of the estimated structural breaks in stationary or nonstationary homogeneous panels have been discussed by Bai (2010), Baltagi, Kao and Liu (2017) and others. As pointed out by Bai (2010), Theorem 1 implies a degenerate limiting distribution for $\hat{k}_j$. To obtain a non-degenerate distribution, a different framework of shrinking magnitude of breaks is usually assumed. Baltagi, Kao and Liu (2017) show the convergence rates of break estimators in homogeneous cointegrated panels and stationary panel regression are $O_p(1/NT)$ and $O_p(1/N)$, respectively, suggesting the benefit of using observations in the cross-sectional dimension under the common break assumption in panels. In our model, similar insights can be carried over. However, when the slopes are heterogeneous, the derivation of convergence rate and limiting distribution of the break point estimators is technically nontrivial. In addition, as shown in the following proposition, the convergence rate of $\hat{k}_j$ is not required for the asymptotic distribution of the slope estimators, so we leave it for future research.

Denote $\underline{V}_i(\mathcal{K}_0) = diag\left(V_{i1}, V_{i2}, \cdots, V_{i,m_0+1}\right)$ **with** $V_{i1} = (v_{i1}, ..., v_{i,K_{0,1}})'$, $V_{i2} = (v_{i,K_{0,1}+1}, ..., v_{i,K_{0,2}})', \cdots, V_{i,m_0+1} = (v_{i,K_{0,m_0}+1}, ..., v_{iT})'$. Given the consistency of estimated structural breaks $\hat{\mathcal{K}}$ above, we can obtain consistent estimators of the slope parameters.

Proposition 1 ***Under Assumptions 1-9, as*** *$(N,T) \to \infty$, and $\frac{\sqrt{T}}{N} \to 0$, for each $i = \{1, \cdots, N\}$,*

$$\sqrt{T}\left(\hat{b}_i - b_i\right) \xrightarrow{d} N\left(0, \Sigma_{X,i}^{-1}\Sigma_{X\varepsilon,i}\Sigma_{X,i}^{-1}\right),$$

where *$\Sigma_{X,i} = plim_{T\to\infty}\frac{1}{T}\underline{V}_i(\mathcal{K}_0)'\underline{V}_i(\mathcal{K}_0)$ and $\Sigma_{X\varepsilon,i} = plim_{T\to\infty}\frac{1}{T}\underline{V}_i(\mathcal{K}_0)'\Sigma_{\varepsilon,i}\underline{V}_i(\mathcal{K}_0)'$.*

According to Lemma 6 in the Appendix, $\Sigma_{X,i}$ can be estimated by $\frac{1}{T}\underline{X}_i(\mathcal{K}_0)'M_{\underline{\overline{X}}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0)$, which can be easily computed when $\mathcal{K}_0$ and $\mathcal{K}_1$ are replaced with their least squares estimates. It has a well-behaved probability limit when $T \to \infty$. **Similarly,** as in Pesaran (2006), KPY and BFK, a consistent Newey-West type estimator of $\Sigma_{X\varepsilon,i}$ can be obtained as

$$\widehat{\Sigma}_{X\varepsilon,i} = \widehat{\Lambda}_{i0} + \sum_{j=1}^{\omega}\left(1 - \frac{j}{\omega+1}\right)\left(\widehat{\Lambda}_{ij} + \widehat{\Lambda}_{ij}'\right), \; \widehat{\Lambda}_{ij} = \frac{1}{T}\sum_{t=j+1}^{\omega} e_{it}e_{i,t-j}\underline{X}_{it}(\hat{\mathcal{K}}_0, \hat{\mathcal{K}}_1)\underline{X}_{it}(\hat{\mathcal{K}}_0, \hat{\mathcal{K}}_1)', \tag{17}$$

where ω is the window size,[12] e_{it} is the t^{th} element of $e_i = M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}Y_i - M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)\hat{b}_i$ and $\underline{X}_{it}(\hat{\mathcal{K}}_0, \hat{\mathcal{K}}_1)$ is the t^{th} row of $M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{K}_0)$. Thus, a consistent Newey-West type estimator of $\Sigma_{X,i}^{-1}\Sigma_{X\varepsilon,i}\Sigma_{X,i}^{-1}$ is given by

$$\left[\frac{1}{T}\underline{X}_i(\hat{\mathcal{K}}_0)'M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)\right]^{-1}\widehat{\Sigma}_{X\varepsilon,i}\left[\frac{1}{T}\underline{X}_i(\hat{\mathcal{K}}_0)'M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)\right]^{-1}. \tag{18}$$

Proposition 2 *Under Assumptions 1-9, and $(N,T) \to \infty$,*

$$\sqrt{N}\left(\hat{b}_{MG} - b\right) \xrightarrow{d} N\left(0, \Sigma_b\right),$$

where Σ_b can be consistently estimated by

$$\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{b}_i - \hat{b}_{MG}\right)\left(\hat{b}_i - \hat{b}_{MG}\right)'.$$

[12] In practice, the selection of the window size ω is important. Pesaran and Timmermann (2007) propose the cross-validation methods for selection of a single estimation window in the presence of breaks.

This result suggests that Theorem 1 in KPY (2011) holds as if K_0 and K_1 were treated as known. Similarly, a pooled estimator

$$\hat{b}_P = \left[\sum_{i=1}^{N} \underline{X}_i(\hat{\mathcal{K}}_0)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} \underline{X}_i(\hat{\mathcal{K}}_0) \right]^{-1} \sum_{i=1}^{N} \underline{X}_i \left(\hat{\mathcal{K}}_0 \right)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} Y_i \tag{19}$$

can be defined as in equation (20) in KPY (2011).[13]

5 Additional Nonstationary Components in the Regressors

In this section, our analysis of nonstationary panels is extended to the case of both nonstationary f_t and v_{it}. Idiosyncratic errors ε_{it} remain $I(0)$. Compared with Section 5 of Phillips and Moon (1999), our model accommodates additional features of an error factor structure and multiple breaks in slopes. In equation (3) $x_{it} = \Gamma_i' f_t + v_{it}$, errors v_{it} follow $I(1)$ processes:

$$v_{it} = v_{i,t-1} + \varsigma_{it},\ i = 1, .., N, \tag{20}$$

where ς_{it} follows the assumption below:

Assumption 10 *ς_{it}, $i = 1, ..., N$, are cross-sectionally independent. For each i, (i) $\varsigma_{it} = \Psi_i(L)\epsilon_{it}$ with ϵ_{it} is IID random variables with zero mean and has a finite fourth-order moments; (ii) $Var(\epsilon_{it}) = \Sigma_{\epsilon,i} = P_i P_i'$, and $\Psi_i(L) = \sum_{j=0}^{\infty} \Psi_{ij} L^j$ with $\sum_{j=0}^{\infty} j \left\| \Psi_{ij} \right\| < \infty$, and $\Psi_i(1) = \sum_{j=0}^{\infty} \Psi_{ij}$.*

Different from Case 1 of stationary v_{it} considered in Section 4, in Case 2 of $I(1)$ v_{it}, the CCE transformed regressors in the partitioned regression (15) remain nonstationary. We will show that $\hat{\mathcal{K}}$ defined in equation (13) above are still consistent and $\hat{b}_i$ is T-consistent. In addition, different from Case 1, the restriction on the relative diverging rate between T and N in Assumption 9 is not required here.

[13] Its limiting distribution can be proved in line with Theorem 2 of KPY:

$$\sqrt{N} \left(\hat{b}_P - b \right) \overset{d}{\to} N\left(0, \Sigma_P\right).$$

Σ_P can be estimated consistently by $\hat{\Sigma}_P = \hat{\Psi}^{*-1} \hat{R} \hat{\Psi}^{*-1}$, where $\hat{R} = \frac{1}{N-1} \sum_{i=1}^{N} \left[\frac{1}{T} \underline{X}_i(\hat{\mathcal{K}}_0)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} \underline{X}_i(\hat{\mathcal{K}}_0) \right] \left(\hat{b}_i - \hat{b}_{MG} \right) \left(\hat{b}_i - \hat{b}_{MG} \right)' \left[\frac{1}{T} \underline{X}_i(\hat{\mathcal{K}}_0)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} \underline{X}_i(\hat{\mathcal{K}}_0) \right]$ and $\hat{\Psi}^* = \frac{1}{NT} \sum_{i=1}^{N} \underline{X}_i(\hat{K}_0)' M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)} \underline{X}_i(\hat{K}_0)$.

Theorem 2 *Under Assumptions 1-**8**, 10, as* $(N,T)\to\infty$, $\lim_{(N,T)\to\infty} P\left(\hat{k}_j = k_j^0\right) = 1, j=\{1,\cdots,m\}$.

With an additional Assumptions 11 and 12 on φ_t and identifying b_i below, respectively, we obtain the following Proposition 3. In line with equation (5.8) in Phillips and Moon (1999) in nonstationary heterogeneous panels without structural breaks and error factors, for each $i=1,...,N$, $\hat{b}_i$ is also super consistent in our model.

Assumption 11 φ_t *is linear stationary process, (i)* $\varphi_t = \Pi(L)u_t$ *with* μ_t, $t = 1,...,T$, *have a finite fourth-order moments; (ii)* $Var(u_t) = \Sigma_u = QQ'$, *and* $\Pi(L) = \sum_{j=0}^{\infty}\Pi_j L^j$ *with* $\sum_{j=0}^{\infty} j\left\|\Pi_j\right\| < \infty$, *and* $\Pi(1) = \sum_{j=0}^{\infty}\Pi_j$.

Assumption 12 $\frac{1}{T^2}\underline{X}_i(\mathcal{K}_0)'M_{\underline{\overline{X}}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0)$ *is non-singular, and its inverse has a finite second-order moment.*

Proposition 3 *Under Assumptions 1-7, 9-12, for each* i, $T(\hat{b}_i - b_i)$ *converges weakly to a non-degenerate distribution, as* $(N,T)\to\infty$.

Intercept estimator is not included in $\hat{b}_i$ above, and its convergence rate is $\sqrt{T}$ as in a cointegration model (Hamilton, 1994, p.588). The intercept can be wiped out by adding a vector of ones to $\underline{\overline{X}}(\hat{\mathcal{K}}_1)$ in the $M_{\underline{\overline{X}}(\hat{\mathcal{K}}_1)}$. The limiting distribution of $T(\hat{b}_i - b_i)$ is complicated and inconvenient in practice. It is of a similar form to Theorem 8 of Phillips and Moon (1999) and is reported in Appendix A.

In empirical applications of heterogeneous panels, the cross-section means of b_i are usually of interest, thus a popular estimator is either the mean-group estimator or the pooled estimator. For the mean group estimator of b,

$$\sqrt{N}\left(\hat{b}_{MG} - b\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} v_{b,i} + \frac{1}{\sqrt{N}T}\sum_{i=1}^{N}\left[\left(\frac{1}{T^2}\underline{X}_i(\mathcal{K}_0)'M_{\underline{\overline{X}}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0)\right)^{-1}\frac{1}{T}\underline{X}_i(\mathcal{K}_0)'M_{\underline{\overline{X}}(\mathcal{K}_1)}\varepsilon_i\right] + o_p(1) \tag{21}$$

The second term is $O_p(1/T)$, dominated by the first term in the equation above. Thus, we can obtain a similar result to Proposition 2 in Case 1: $\sqrt{N}\left(\hat{b}_{MG} - b\right) \xrightarrow{d} N(0,\Sigma_b)$ as $(N,T)\to\infty$.

In a special case of homogeneous slopes $b_i = b$ with $v_{b,i} = 0$, the first term in equation (21) disappears. Thus, equation (21) reduces to

$$\sqrt{N}T\left(\hat{b}_{MG} - b\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[\left(\frac{1}{T^2}\underline{X}_i(\mathcal{K}_0)'M_{\underline{\overline{X}}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0)\right)^{-1}\frac{1}{T}\underline{X}_i(\mathcal{K}_0)'M_{\underline{\overline{X}}(\mathcal{K}_1)}\varepsilon_i\right] + o_p(1)\,. \tag{22}$$

The convergence rate of $\hat{b}_{MG}$ in a homogeneous panel becomes $\sqrt{N}T$, same as in Bai, Kao and Ng (2009) and Huang et al. (2020).

We obtain the following Proposition 4.

Proposition 4 *Under Assumptions 1-7, 9 and 10-12, in a homogeneous panel with $b_i = b$, as $(N,T) \to \infty$,*

$$\sqrt{N}T\left(\hat{b}_{MG} - b\right) \xrightarrow{d} N\left(0, \Sigma_{MG}\right),$$

where

$$\Sigma_{MG} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} E[(T^{-2}\underline{X}_i(\mathcal{K}_0)' M_{\overline{\underline{X}}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0))^{-1}(T^{-1}\underline{X}_i(\mathcal{K}_0)' M_{\overline{\underline{X}}(\mathcal{K}_1)}\varepsilon_i)$$
$$\times (T^{-1}\varepsilon_i' M_{\overline{\underline{X}}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0)')(T^{-2}\underline{X}_i(\mathcal{K}_0)' M_{\overline{\underline{X}}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0))^{-1}].$$

More details can be found in Appendix A.4. For simplicity, the asymptotic bias mentioned in Theorem 8 of Phillips and Moon (1999) and Proposition 1 of Bai, Ng and Kao (2009) disappears here under the assumptions of no serial/ cross-sectional correlation and heteroskedasticity. Σ_{MG} can be estimated consistently by

$$\hat{\Sigma}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{1}{T^2}\underline{X}_i(\hat{\mathcal{K}}_0)' M_{\overline{\underline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)\right)^{-1}\left(\frac{1}{T}\underline{X}_i(\hat{\mathcal{K}}_0)' M_{\overline{\underline{X}}(\hat{\mathcal{K}}_1)}\hat{\varepsilon}_i\right)\right.$$
$$\left.\times\left(\frac{1}{T}\hat{\varepsilon}_i' M_{\overline{\underline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)\right)\left(\frac{1}{T^2}\underline{X}_i(\hat{\mathcal{K}}_0)' M_{\overline{\underline{X}}(\hat{\mathcal{K}}_1)}\underline{X}_i(\hat{\mathcal{K}}_0)\right)^{-1}\right],$$

where $\hat{\varepsilon}_i = Y_i - \underline{X}_i(\hat{K}_0)\hat{b}_{MG}$.

6 Monte Carlo Simulations

In this section, Monte Carlo experiments are conducted to examine the finite sample properties of the break estimators. We consider the case of three breaks, i.e., $m = 3$, including two common breaks in slopes (k_1^0, k_2^0) and a third one in error factor loadings k_3^0 in various scenarios. We find supporting results to the main findings in Theorems 1 and 2. This is done by looking at the frequency of choosing true breaks using the proposed break estimators. For nonstationary panels, nonstationarity could come from either f_t or v_{it} or both under the common factor assumption (3). Thus, we consider six different scenarios: i) Case 1 with $I(1)$ factors f_t and $I(0)$ v_{it}; ii) Case 1 under rank deficiency; iii) Case 2 with $I(1)$ f_t and $I(1)$ v_{it}; iv) Case 2 with $I(0)$ f_t and $I(1)$ v_{it}; v) Case 2 with $I(1)$ errors ε_{it}; vi) Case 1 with mixed stationary and nonstationary regressors and factors.

6.1 Data Generating Process

Our basic design is similar to the one used in KPY but now with multiple breaks:

$$y_{it} = \alpha_i + \beta_i \left(k_1^0, k_2^0\right) x_{it} + \gamma_{1,i} \left(k_3^0\right) f_t + \varepsilon_{it}, \; i = 1, ..., N; \; t = 1, ..., T, \qquad (23)$$

where $\alpha_i \sim iidN(1, 1)$. The scalar regressor x_{it} is affected by the common correlated effect f_t:

$$x_{it} = a_i + \gamma_{2,i} f_t + v_{it}, \qquad (24)$$

with $a_i \sim iidN(0.5, 0.5)$ and $\gamma_{2,i} \sim iidN(0.5, 0.5)$. The scalar factor f_t follows an $I(1)$ process:

$$f_t = f_{t-1} + v_{ft}, t = -49, ..., 0, 1, ..., T;$$

where $f_{-50} = 0$, $v_{ft} \sim iidN(0, 1)$.

Two common breaks k_1^0, k_2^0 in slopes are assumed at $[0.3T]$ and $[0.5T]$ of the time span:

$$\beta_i(k_1^0, k_2^0) = \begin{cases} \beta_i, & t = 1, ..., k_1^0, \\ \beta_i + \Delta\beta_i, & t = k_1^0 + 1, ..., T, \\ \beta_i + 2\Delta\beta_i, & t = k_2^0 + 1, ..., T \end{cases}$$

where $\beta_i \sim iidN(1, 0.04)$ and $\Delta\beta_i \sim iidN(0, 0.5)$. A third break $k_3^0 = [0.7T]$ occurs in the error factor loadings:

$$\gamma_{1,i}(k_3^0) = \begin{cases} \gamma_{1,i}, & t = 1, ..., k_3^0, \\ \gamma_{1,i} + \Delta\gamma_i, & t = k_3^0 + 1, ..., T, \end{cases} \qquad (25)$$

where $\gamma_{1,i} \sim iidN(1, 0.2)$ and $\Delta\gamma_i \sim iidN(0.5, 0.5)$.

In scenario (i) of Case 1, as in KPY, both ε_{it} and v_{it} are stationary. $\varepsilon_{it} = \rho_{i\varepsilon}\varepsilon_{i,t-1} + \sigma_i \left(1 - \rho_{i\varepsilon}^2\right)^{0.5} \omega_{it}$, for $i = 1, 2, ..., [N/2]$ and $\varepsilon_{it} = \sigma_i \left(1 + \theta_{i\varepsilon}^2\right)^{-0.5} \left(\omega_{it} + \theta_{i\varepsilon}\omega_{i,t-1}\right)$, for $i = [N/2] + 1, ..., N$, with $\omega_{it} \sim iidN(0, 1), \sigma_i^2 \sim iidU[0.5, 1.5], \rho_{i\varepsilon} = iidU[0.05, 0.95]$ and $\theta_{i\varepsilon} \sim iidU[0, 1]$. Similarly, $v_{it} = \rho_{vi}v_{i,t-1} + \psi_{it}, \psi_{it} \sim iidN(0, 1 - \rho_{vi}^2)$, with $v_{i,-49} = 0$, and $\rho_{vi} \sim iidU[0.05, 0.95]$.[14]

In scenario (ii), we consider the importance of rank deficiency in finite samples. The DGP here is the same as above, except that the means of a_i and $\gamma_{2,i}$ change to zero, i.e., $a_i \sim iidN(0, 0.5)$ and $\gamma_{2,i} \sim iidN(0, 0.5)$ in equation (24). In the current design, the rank condition is not satisfied asymptotically.

In scenario (iii) of Case 2, both v_{it} and ε_{it} follow $I(1)$ processes,

$$v_{it} = v_{i,t-1} + \psi_{it}, \; \psi_{it} \sim iidN(0, 1), \; t = -49, ..., 0, 1, ..., T.$$

[14] In this design, the signal-to-noise ratio is about 1.5.

We also allow for $I(0)$ f_t in the design above in scenario (iv). In addition, in scenario (v), we examine the impact of nonstationary errors on break point estimators, we also consider Case 2 with nonstationary errors, i.e., $I(1)$ ε_{it}, $\varepsilon_{it} = \varepsilon_{i,t-1} + \vartheta_{it}$, $\vartheta_{it} \sim iidN(0,1)$, $t = -49, ..., 0, 1, ..., T$.

Finally, in scenario (vi), we also consider the case of mixed stationary and nonstationary regressors and factors. To allow for a stationary regression, we add an additional regressor and factor in the regression (23) above. More specifically,

$$y_{it} = \alpha_i + \beta_{1,i}\left(k_1^0\right) x_{1,it} + \beta_{2,i}\left(k_2^0\right) x_{2,it} + \gamma_{11,i}\left(k_3^0\right) f_{1,t} + \gamma_{12,i}\left(k_3^0\right) f_{2,t} + \varepsilon_{it},$$

where both regressors are generated by

$$\begin{aligned} x_{1,it} &= a_i + \gamma_{21,i} f_{1,t} + \gamma_{22,i} f_{2,t} + v_{1,it}, \\ x_{2,it} &= a_i + 0 \cdot f_{1,t} + \gamma_{23,i} f_{2,t} + v_{2,it}. \end{aligned}$$

We assume that both $v_{1,it}$ and $v_{2,it}$ are $I(0)$ as v_{it} in Case 1 above. Two factors $f_{1,t}$ and $f_{2,t}$ are generated as $I(1)$ and $I(0)$ processes, respectively, as follows:

$$f_{1,t} = f_{1,t-1} + v_{1,ft}, \text{ and } f_{2,t} = 0.5 f_{2,t-1} + v_{2,ft}.$$

Thus, $x_{1,it}$ is $I(1)$ and $x_{2,it}$ is $I(0)$. Same as γ_{2i}, loadings $\gamma_{21,i}, \gamma_{22,i}, \gamma_{23,i} \sim iidN(0.5, 0.5)$. The break points $k_1^0 = [0.3T], k_2^0 = [0.5T]$ appear in the slopes:

$$\begin{aligned} \beta_{1,i}\left(k_1^0\right) &= \begin{cases} \beta_{11,i}, & t = 1, ..., k_1^0, \\ \beta_{11,i} + \Delta\beta_{1,i}, & t = k_1^0 + 1, ..., T, \end{cases} \\ \beta_{2,i}\left(k_2^0\right) &= \begin{cases} \beta_{21,i}, & t = 1, ..., k_2^0, \\ \beta_{21,i} + \Delta\beta_{2,i}, & t = k_2^0 + 1, ..., T, \end{cases} \end{aligned}$$

where $\Delta\beta_{1,i}, \Delta\beta_{2,i} \sim iidN(0, 0.16)$. Here $\gamma_{11,i}\left(k_3^0\right)$ and $\gamma_{12,i}\left(k_3^0\right)$ have the same design as $\gamma_{1,i}\left(k_3^0\right)$ in (25) but the variance of $\Delta\gamma_i$ changes from 0.5 to 0.16.

Different combinations of $T = 20, 50, 100$ and $N = 10, 50, 200$ are considered in the Monte Carlo experiments with 1,000 replications. Due to limited space, only the results with $T = 50$ are reported in the paper.

6.2 Results

Figure 1 presents the histograms of estimators $(\hat{k}_1, \hat{k}_2, \hat{k}_3)$ in Case 1 with nonstationary factors for $T = 50$. The true values of the break points are $k_1^0 = 15$, $k_2^0 = 25$, $k_3^0 = 35$. In each replication, a dynamic programming algorithm proposed by Bai and Perron (2003) is applied to obtain $\hat{k}_1$, $\hat{k}_2, \hat{k}_3$ simultaneously. The upper, middle and lower panels represent the empirical distributions of $\hat{k}_1$, $\hat{k}_2$ and $\hat{k}_3$, respectively. Figure 1 shows that the frequencies of choosing $(k_1^0,\ k_2^0, k_3^0)$ increase substantially as N increases from 10 to 200. For example, the probability of choosing k_1^0 increases from 36% for $N = 10$ to 69% for $N = 200$. This finding supports the results in Theorem 1.

Figure 2 reports the histograms of $(\hat{k}_1,\ \hat{k}_2, \hat{k}_3)$ in Case 1 for $T = 50$ under rank deficiency. The rank condition is required for the validity of the CCE approach to deal with unobserved common factors. We examine the finite sample properties of these break estimators when the rank condition is not satisfied asymptotically. Although the probabilities of choosing the true break points are smaller than those in Figure 1, they still increase substantially with N, showing that under rank deficiency, the estimators $(\hat{k}_1,\ \hat{k}_2, \hat{k}_3)$ are still very informative about choosing $(k_1^0,\ k_2^0, k_3^0)$ when N is large.

In Figure 3, we consider Case 2 with nonstationary regressors and both f_t and v_{it} are nonstationary in x_{it}. Similar patterns as in Figure 1 are observed. The probabilities of choosing true break dates increase with N, e.g., nearly 100% for choosing k_1^0 by $\hat{k}_1$ for $N = 200$ and $T = 50$. This finding supports the consistency of the break estimators in Theorem 2. In Figure 4, we also consider a scenario of an $I(0)$ f_t and $I(1)$ v_{it} in Case 2, where $f_t = 0.5 f_{t-1} + v_{ft}$ and $v_{ft} \sim iidN(0, 0.75)$. As expected, as long as x_{it} is still $I(1)$, $\hat{k}_1$, $\hat{k}_2, \hat{k}_3$ are consistent. Little impact is spotted from changing f_t from $I(1)$ to $I(0)$ in Figure 4.

In Figure 5, we consider the scenario of nonstationary errors ε_{it} in the design of Case 2 above. Under the current design, f_t, v_{it} and ε_{it} follow $I(1)$ processes. Different from Case 2, $I(1)$ ε_{it} could lead to a spurious regression and thus, the least squares estimators of slopes could be inconsistent. In addition, nonstationary ε_{it} could lead to a smaller signal-to-noise ratio in the DGP of Figure 5 than that of Figure 3 with $I(0)$ ε_{it}. Thus, we observe smaller probabilities of choosing (k_1^0, k_2^0, k_3^0) here, even though the same pattern remains. That is, big N helps to date the break points.

Lastly, we examine the scenario of mixed stationary and nonstationary regressors in Figure 6, as in Bai, Kao and Ng (2009), Huang, Jin, Phillips, Su (2021). Slightly different from the designs used in Figures 1-4, an additional regressor and factor are added to the design (23). In our modified design, given an $I(1)$ $f_{1,t}$ and an $I(0)$ $f_{2,t}$, $x_{1,it}$, $x_{2,it}$ are $I(1)$ and $I(0)$, respectively. We consider $I(0)$ v_{it} in this scenario to avoid potential spurious regression after $f_{1,t}$ and $f_{2,t}$ are partialled out from the regressors and y_{it}. As expected, the frequency of choosing k_2^0, the break point in the stationary regressors, is smaller than that of choosing k_1^0 under the same design parameters for a same N. After scaling up the magnitude of the break in $\beta_{2,i}(k_2^0)$, we find a similar pattern as in Figure 1, still observing increasing probabilities of dating the true break points with N in the histograms of $\hat{k}_1$, $\hat{k}_2, \hat{k}_3$.

Moreover, we also conduct additional robustness checks, including using $(\bar{y}_{\cdot t}, \bar{x}_{\cdot t})$ instead of $\bar{x}_{i\cdot}$, to proxy f_t, boundary breaks, fixed effects model, different magnitude of breaks in slopes and factor loadings, adding a time trend etc. These results can be found in Figures A1-A6 in the supplementary Appendix B. The results with $T = 20$ and 100 are in line with those with $T = 50$ reported above, and are available upon request from the authors.

[Insert Figures 1-6 Here]

Finally, we examine the finite sample properties of the slope estimators in Case 1. Table 1 reports the root mean squared error (RMSE) and bias of $\hat{b}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{b}_i(\hat{k}_1, \hat{k}_2, \hat{k}_3)$ defined in (16) under the design described in Figure 1. The size of the t test is also included. The results show that the RMSE as well as bias decrease notably with (N, T), in line with the simulation results in KPY.

7 Application: International R&D Spillovers

In this section, we apply our approach to an empirical example of international R&D spillovers, which was studied by Coe and Helpman (1995) and Coe, Helpman, Hoffmaister (2009, CHH hereafter). Huang et al. (2021) find a latent group structure in the long-run relationship between technological change, domestic R&D stock, foreign R&D stock for 24 OECD countries during 1971-2004. **Different from CHH (2009) and Huang et al. (2021) who emphasize heterogeneous international R&D spillovers in different countries, we focus on the heterogeneous effects in different time periods along with the changing global economic conditions.**

As pointed out by Coe et al. (2009), the total factor productivity (TFP) and domestic R&D stock accelerated after 1990 for some countries. To accommodate this pattern, we allow common breaks in their long-run relationships. We follow the specification considered by Huang et al. (2021, model (5.1)),

$$\log(y_{it}) = \beta_i^d\left(k_1^0\right)\log(s_{it}^d) + \beta_i^f\left(k_1^0\right)\log(s_{it}^f) + \gamma_i' f_t + \varepsilon_{it}, \tag{26}$$

Table 1: $RMSE$, $Bias$ and inference of the Mean group estimators $\hat{b}_{MG}$

		$RMSE \times 100$				$Bias \times 100$				$Size$ (5%)			
Method	T\N	30	50	100	200	30	50	100	200	30	50	100	200
$\hat{\beta}_{1,MG}$	30	8.78	6.32	5.65	5.12	0.25	-0.20	0.19	0.17	7.34	6.23	5.43	5.03
	30	6.98	5.32	4.99	4.55	-0.18	0.15	-0.13	-0.11	6.85	6.65	5.21	4.87
	100	4.65	3.32	3.12	2.98	-0.09	0.08	0.09	-0.03	5.31	6.78	5.43	5.05
	200	2.52	2.43	1.97	1.75	0.08	0.06	0.02	0.01	5.11	5.32	5.16	4.89
$\hat{\beta}_{2,MG}$	20	7.54	6.21	5.23	5.25	-0.28	-0.19	0.20	-0.18	7.24	6.11	5.54	5.31
	50	6.64	5.46	4.83	4.64	0.22	-0.15	-0.15	0.12	7.01	6.01	5.15	5.21
	100	3.46	3.65	3.60	2.72	0.17	0.09	0.07	-0.02	6.32	6.03	5.40	5.05
	200	2.74	2.49	1.94	1.83	-0.11	0.03	0.01	0.00	5.89	5.43	5.53	5.12
$\hat{\beta}_{3,MG}$	20	7.84	6.43	5.65	5.35	0.26	0.22	-0.18	0.15	6.99	6.14	5.43	4.86
	50	5.75	5.64	4.75	4.13	-0.17	0.16	0.08	-0.04	6.56	6.57	5.51	4.59
	100	3.73	3.39	3.48	3.54	0.10	-0.08	0.03	-0.03	5.46	6.64	5.43	5.23
	200	2.86	2.65	2.32	1.93	0.07	0.09	0.01	0.01	4.87	5.43	5.05	5.05

Note: (1) the DGP is same as that of Figure 1. (2) the estimated structural breaks $(\hat{k}_1, \hat{k}_2)$ split the whole time period into three regimes, such that $\hat{b}_i = (\hat{\beta}_{i1}, \hat{\beta}_{i2}, \hat{\beta}_{i3})$. The mean group estimators are defined as $\hat{\beta}_{1,MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_{i1}$, $\hat{\beta}_{2,MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_{i2}$, $\hat{\beta}_{3,MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_{i3}$ respectively. (3) For $s = \{1,2,3\}$, the standard error of $\hat{\beta}_{s,MG}$ is be consistently estimated by $\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\beta}_{i1} - \hat{\beta}_{s,MG}\right)^2$ as in Proposition 2.

where y_{it} is the TFP in country i in year t. s_{it}^d and s_{it}^f are real domestic and foreign R&D capital stocks, respectively. $\beta_i^d(k_1^0)$ and $\beta_i^f(k_1^0)$ represent heterogeneous effects of domestic and foreign R&D stocks on the TFP. We allow a common break k_1^0 in the slopes. Detailed data information is provided by Coe et al. (2009) and Huang et al. (2021), who found a single nonstationary common factor in the data. Here, we also assume an I(1) factor f_t.

Table 2 columns (1) and (2) include the dynamic OLS estimates of CHH (2009) and pooled FM-OLS by Huang et al. (2021, Table 7) without considering a latent group structure in the slopes for comparison. Using the cross-sectional averages of log(s_{it}^d) and log(s_{it}^f) to proxy the unobserved common factor, we estimate the common break and slopes in (26) with the least squares estimation method proposed in Section 3. There are two key findings in our estimation results. First, we find that there is a common break in the slopes in 1992.[15] **It splits the sample period into two regimes, 1971-1992 and 1993-2004, and the estimation results in these two sample periods are reported in columns (3) and (4), respectively. Second, the coefficients of** $\log(s_{it}^d)$ **and** $\log(s_{it}^f)$ **are significantly different in these two periods, with a doubling effect of foreign R&D spillovers during 1993-2004.** [16]

The doubling effect of **foreign R&D spillovers suggests that international technology diffusion via importing foreign R&D plays a more important role in boosting domestic productivity growth than domestic R&D in the OECD countries, and this effect is more pronounced starting from 1993. Following the German reunification in October 1990, the collapse of the former Soviet Union in Decomember 1991, and more importantly, the formal establishment of the European Union in 1993, globalization accelerated in the early 1990s. According to the well cited KOF Globalisation Index, the world overall index sped up starting from 1991.** [17]

In a globalized economy, R&D activities concentrate in a few rich countries. For example, Keller (2004) documented that 84 percent of

[15]**The CUSUM test of common breaks proposed by Jiang and Kurozumi (2023) suggests that there is one comon break in the data.**

[16]When there are two common breaks, our second estimated break date occurs in 1976. The second break of 1976 splits the period 1971-1992 into two sub-regimes.

[17]The KOF Globalisation Index is provided by KOF Swiss Economic Institute at ETH Zurich. The link: https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html

Table 2: Structural Change In International R&D Spillover

Dependent Variable:	Total Factor Productivity			
Periods	1971-2004	1971-2004	1971-1992	1993-2004
Columns	(1)	(2)	(3)	(4)
Methods	CHH2009	FM-OLS	CCEMG	CCEMG
$\log(s_{it}^d)$	0.095***	0.099***	0.084***	0.098***
	(0.005)	(0.027)	(0.005)	(0.005)
$\log(s_{it}^f)$	0.213***	0.121***	0.123***	0.251***
	(0.014)	(0.044)	(0.035)	(0.054)

Note: (1) Standard errors are reported in parentheses. (2) The stars, *, ** and *** indicate the significance level at 10%, 5% and 1%, respectively.

the world's R&D spending was contributed by the G-7 countries in 1995. With more free trade and foreign direct investment, small and developing economies depend more on foreign technologies than domestic R&D in their productivity growth. According to Keller (2004), "for most countries, foreign sources of technology account for 90 percent or more of domestic productivity growth. " Our estimates in columns (3) and (4) indicate that this is also the case for OECD countries.

8 Conclusion

This paper proposes the estimation of unknown multiple structural breaks both in slopes and factor loadings in nonstationary panels with common factors. Based on KPY's approach for dealing with nonstationary factors in panels, we extend Bai and Perron's least squares estimator for multiple breaks in time series regression to nonstationary heterogeneous panels with unobserved factors in errors. We show that the proposed estimators, including the estimated structural breaks and slopes, are consistent in both cases of nonstationary factors and nonstationary regressors. These main findings are supported by the Monte Carlo simulations.

There are potentially two important issues to explore in the current framework. One is testing for multiple structural changes in nonstationary panels. In this paper, we only assume multiple breaks in slopes and factor loadings and estimate these break points. It would be meaningful to test the existence of the breaks in many empirical studies before applying our estimation methods. A candidate is to extend Bai and Perron's (1998) sup$F$ or double maximum tests into nonstationary panels. Another important issue is related to sequential estimation of the break points. In this paper, we estimate multiple breaks simultaneously. In the case of mixed stationary and nonstationary factors and regressors as considered in Figures 4 and 5, it would matter a lot whether breaks are estimated simultaneously or sequentially. It would be interesting to explore the asymptotic properties of sequential estimation of multiple breaks as in Bai and Perron (1998) and Pang, Du and Chong (2021). We leave these research questions for future research.

References

Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, 79(4), 551-563.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4), 1229-1279.

Bai, J. (2010). Common breaks in means and variances for panel data. *Journal of Econometrics*, 157(1): 78-92.

Bai, J., & Li, K. (2014). Theory and methods of panel data models with interactive effects. *Annals of Statistics*, 42(1), 142-170.

Bai, J., & Kao, C. (2006). On the estimation and inference of a panel cointegration model with cross-sectional dependence. *Contributions to Economic Analysis*, 274, 3-30.

Bai, J., & Ng, S. (2010). Instrumental variable estimation in a data rich environment. *Econometric Theory*, 26(6), 1577-1606.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 47-78.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1-22.

Bai, J., Han, X., & Shi, Y. (2020). Estimation and inference of change points in high-dimensional factor models. *Journal of Econometrics*, 219(1), 66-100.

Bai, J., Kao, C, and Ng, S. (2009). Panel cointegration with global stochastic trends. *Journal of Econometrics*, 149(1): 82-99.

Bai, J., and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1), 47-78.

Baltagi, B. H., Feng, Q., & Kao, C. (2016). Estimation of heterogeneous panels with structural breaks. *Journal of Econometrics*, 191(1), 176-195.

Baltagi, B. H., Feng, Q., & Kao, C. (2019). Structural changes in heterogeneous panels with endogenous regressors. *Journal of Applied Econometrics*, 34(6), 883-892.

Baltagi, B. H., Kao, C., & Liu, L. (2017). Estimation and identification of change points in panel models with nonstationary or stationary regressors and error term. *Econometric Reviews*, 36(1-3), 85-102.

Baltagi, B. H., Kao, C., & Wang, F. (2015). Change point estimation in large heterogeneous panels. Manuscript, Syracuse University.

Baltagi, B. H., Kao, C., & Wang, F. (2017). Identification and estimation of a large factor model with structural instability. *Journal of Econometrics*, 197(1), 87-100.

Baltagi, B. H., Kao, C., & Wang, F. (2021). Estimating and testing high dimensional factor models with multiple structural changes. *Journal of Econometrics*, 220(2), 349-365.

Bergamelli, M., Bianchi, A., Khalaf, L., & Urga, G. (2019). Combining p-values to test for multiple structural breaks in cointegrated regressions. *Journal of Econometrics*, 211(2), 461-482.

Breitung, J., & Eickmeier, S. (2011). Testing for structural breaks in dynamic factor models. *Journal of Econometrics*, 163(1), 71-84.

Boldea, O., Drepper, B., & Gan, Z. (2020). Change point estimation in panel data with time-varying individual effects. *Journal of Applied Econometrics*, 35(6), 712-727.

Chen, L., Dolado, J. J., & Gonzalo, J. (2014). Detecting big structural breaks in large factor models. *Journal of Econometrics*, 180(1), 30-48.

Cheng, X., Liao, Z., & Schorfheide, F. (2016). Shrinkage estimation of high-dimensional factor models with structural instabilities. *Review of Economic Studies*, 83(4), 1511-1543.

Coe, D. T., & Helpman, E. (1995). International R&D spillovers. *European Economic Review*, 39(5), 859-887.

Coe, D. T., Helpman, E., & Hoffmaister, A. W. (2009). International R&D spillovers and institutions. *European Economic Review*, 53(7), 723-741.

Dong, C., Gao, J., & Peng, B. (2021). Varying-coefficient panel data models with nonstationarity and partially observed factor structure. *Journal of Business & Economic Statistics*, 39(3), 700-711.

Ditzen, J., Karavias, Y., & Westerlund, J. (2023). Multiple structural breaks in interactive effects panel data and the impact of quantitative easing on bank lending. *arXiv:2211.06707*.

Duan, J., Bai, J., & Han, X. (2023). Quasi-maximum likelihood estimation of break point in high-dimensional factor models. *Journal of Econometrics*, 233(1), 209-236.

Feng, Q., and C. Kao (2020), Large-Dimensional Panel Data Econometrics: Testing, Estimation and Structural Changes, *World Scientific* ***Publishing Co. Pte.***

***Ltd.*.**

Hamilton, J. (1994), Time Series Analysis, *Princeton University Press.*

Holly, S., Pesaran, M. H., & Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics*, 158(1), 160-173.

Hsiao, C. (2018). Panel models with interactive effects. *Journal of Econometrics*, 206(2), 645-673.

Hsiao, C., Shi, Z., & Zhou, Q. (2022). Transformed estimation for panel interactive effects models. *Journal of Business & Economic Statistics*, 40(4), 1831-1848.

Huang, W., Jin, S., Phillips, P. C. B., & Su, L. (2021). Nonstationary panel models with latent group structures and cross-section dependence. *Journal of Econometrics*, 221(1), 198-222.

Huang, W., Jin, S., & Su, L. (2020). Identifying latent grouped patterns in cointegrated panels. *Econometric Theory*, 36(3), 410.

Jiang, P., & Kurozumi, E. (2023). A new test for common breaks in heterogeneous panel data models. *Econometrics and Statistics*, forthcoming.

Juodis, A., Karabiyik, H., & Westerlund, J. (2021). On the robustness of the pooled CCE estimator. *Journal of Econometrics*, 220(2), 325-348.

Juodis, A., Reese S., (2022). The Incidental Parameters Problem in Testing for Remaining Cross-Section Correlation, *Journal of Business & Economic Statistics*, 40(3), 1191-1203.

Kaddoura, Y., & Westerlund, J. (2023). Estimation of panel data models with random interactive effects and multiple structural breaks when T is fixed. *Journal of Business & Economic Statistics*, 41(3), 778-790.

Kapetanios, G., Pesaran, M. H., & Yamagata, T. (2011). Panels with non-stationary multifactor error structures. *Journal of Econometrics*, 160(2), 326-348.

Karavias, Y., Narayan, P. K., & Westerlund, J. (2023). Structural breaks in interactive effects panels and the stock market reaction to COVID-19. *Journal of Business & Economic Statistics*, 41(3), 653-666.

Karabiyik, H., Reese, S., & Westerlund, J. (2017). On the role of the rank condition in CCE estimation of factor-augmented panel regressions. *Journal of Econometrics*, 197(1), 60-64.

Karabiyik, H., Urbain, J. P., & Westerlund, J. (2019). CCE estimation of factor-augmented regression models with more factors than observables. *Journal of Applied Econometrics*, 34(2), 268-284.

Kejriwal, M., & Perron, P. (2008). The limit distribution of the estimates in cointegrated regression models with multiple structural changes. *Journal of Econometrics*, 146(1), 59-73.

Keller, W., 2004. International technology diffusion. *Journal of Economic Literature*, 42, 752-782.

Kim, D. (2011). Estimating a common deterministic time trend break in large panels with cross sectional dependence. *Journal of Econometrics*, 164(2), 310-330.

Kim, D. (2014). Common breaks in time trends for large panel data with a factor structure. *Econometrics Journal*, 17(3), 301-337.

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

Lumsdaine, R. L., Okui, R., & Wang, W. (2023). Estimation of panel group structure models with structural breaks in group memberships and coefficients. *Journal of Econometrics*, 233(1), 45-65.

Ma, C., & Tu, Y. (2023). Group fused Lasso for large factor models with multiple structural breaks. *Journal of Econometrics*, 233(1), 132-154.

Maheu, J. M., & Song, Y. (2018). An efficient Bayesian approach to multiple structural change in multivariate time series. *Journal of Applied Econometrics*, 33(2), 251-270.

Oka, T., & Perron, P. (2018). Testing for common breaks in a multiple equations system. *Journal of Econometrics*, 204(1), 66-85.

Pang, T., Du, L and Chong, T. (2021), Estimating multiple breaks in nonstationary autoregressive models, *Journal of Econometrics*, 221 (1), 2021, 277-311.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967-1012.

Pesaran, M. H., & Timmermann, A. (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics*, 137(1), 134-161.

Phillips, P. C. B., and Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, 67(5), 1057-1111.

Qian, J., & Su, L. (2016). Shrinkage estimation of common breaks in panel data models via adaptive group fused lasso. *Journal of Econometrics*, 191(1), 86-109.

Qu, Z., & Perron, P. (2007). Estimating and testing structural changes in multivariate regressions. *Econometrica*, 75(2), 459-502.

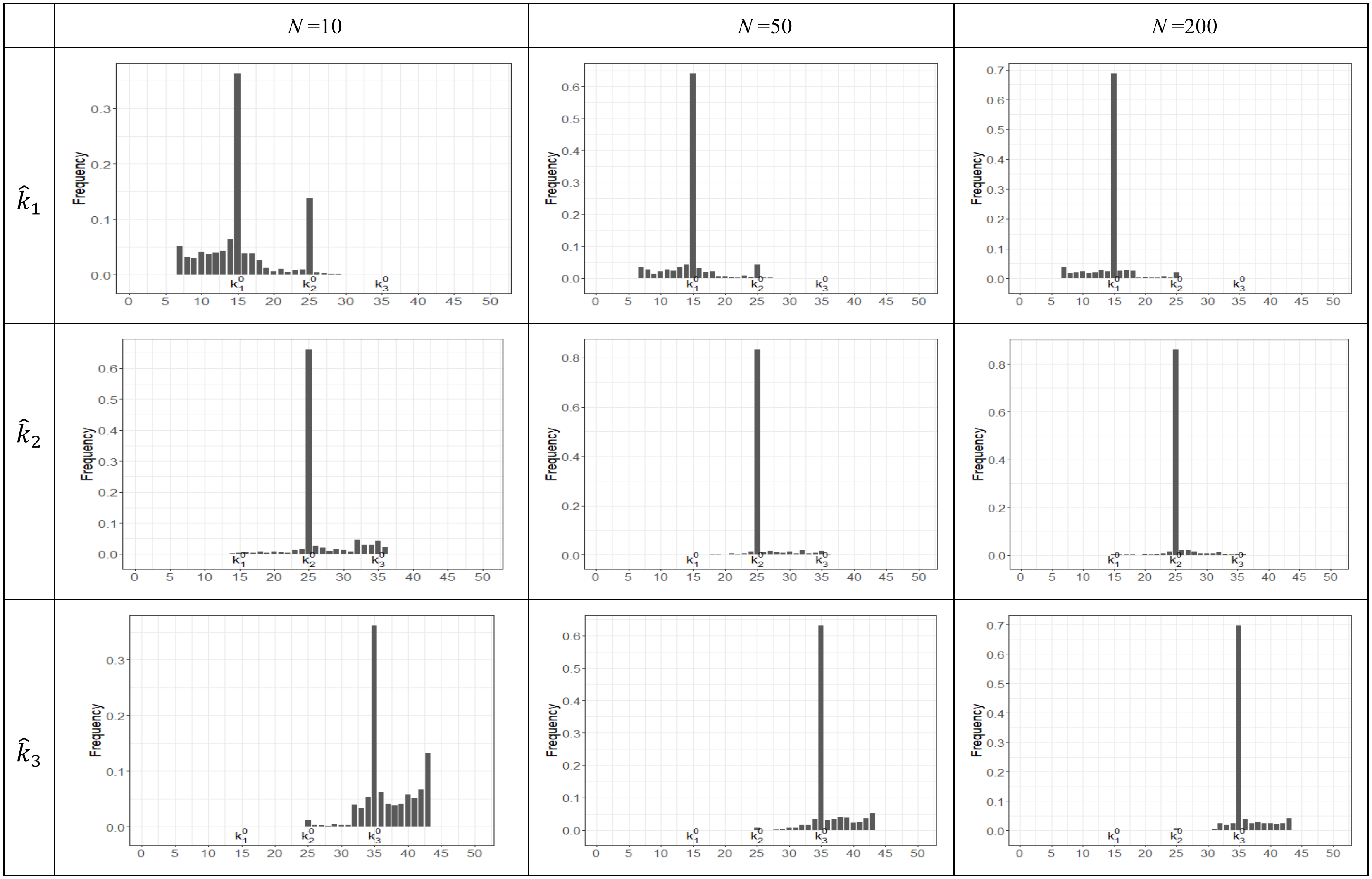
Stock, J. H., & Watson, M. (2009). Forecasting in dynamic factor models subject to structural instability. *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, 173, 205.

Stock, J. H., & Watson, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. *In Handbook of macroeconomics* (Vol. 2, pp. 415-525). Elsevier.

Smith, S. C. (2024). Noncommon breaks. *Journal of Business & Economic Statistics*, 1-24.

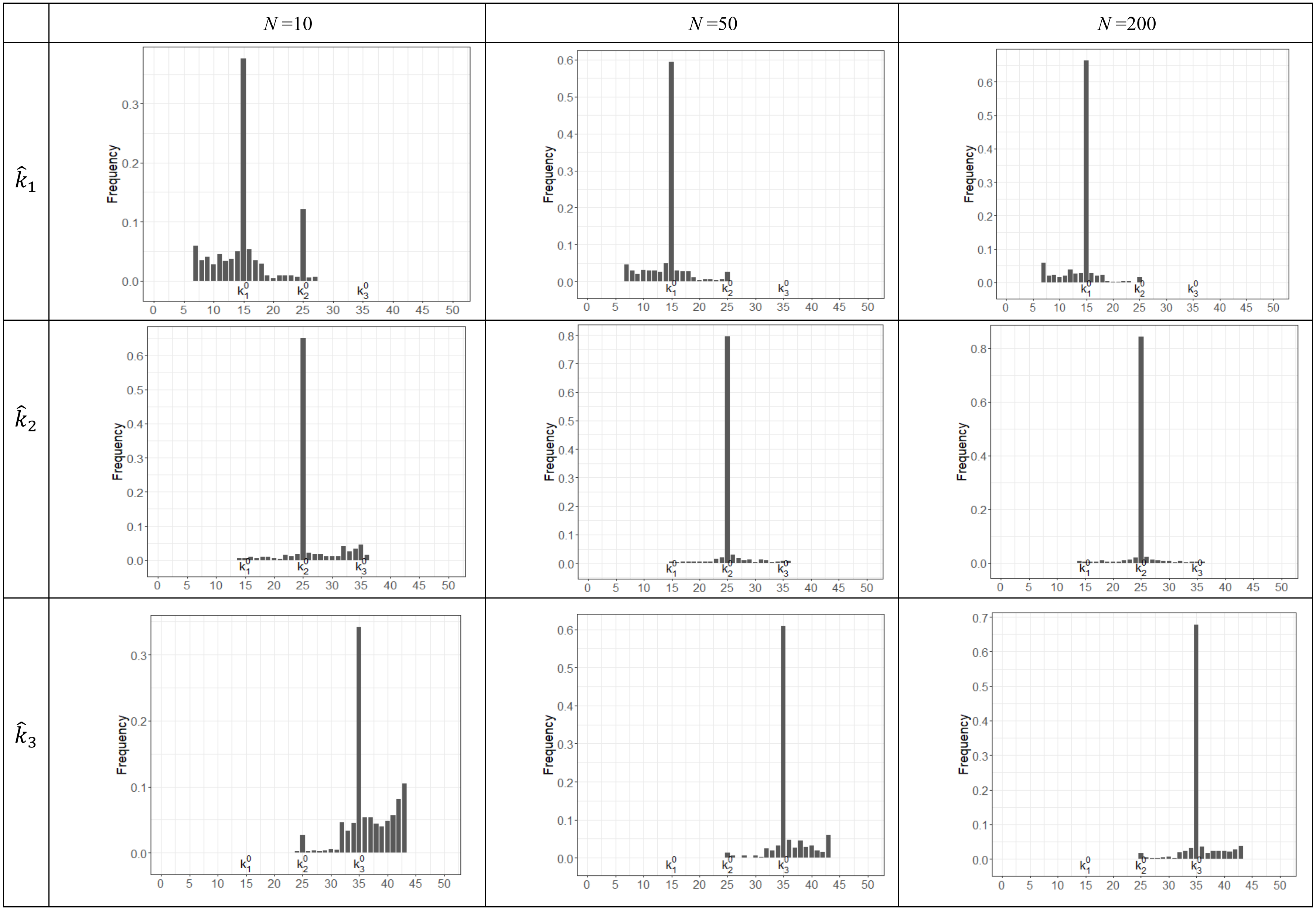
Yamamoto, Y., & Tanaka, S. (2015). Testing for factor loading structural change under common breaks. *Journal of Econometrics*, 189(1), 187-206.

Figure 1: Histograms of Break Point Estimators in Case 1 with Nonstationary Factors ($T = 50$)



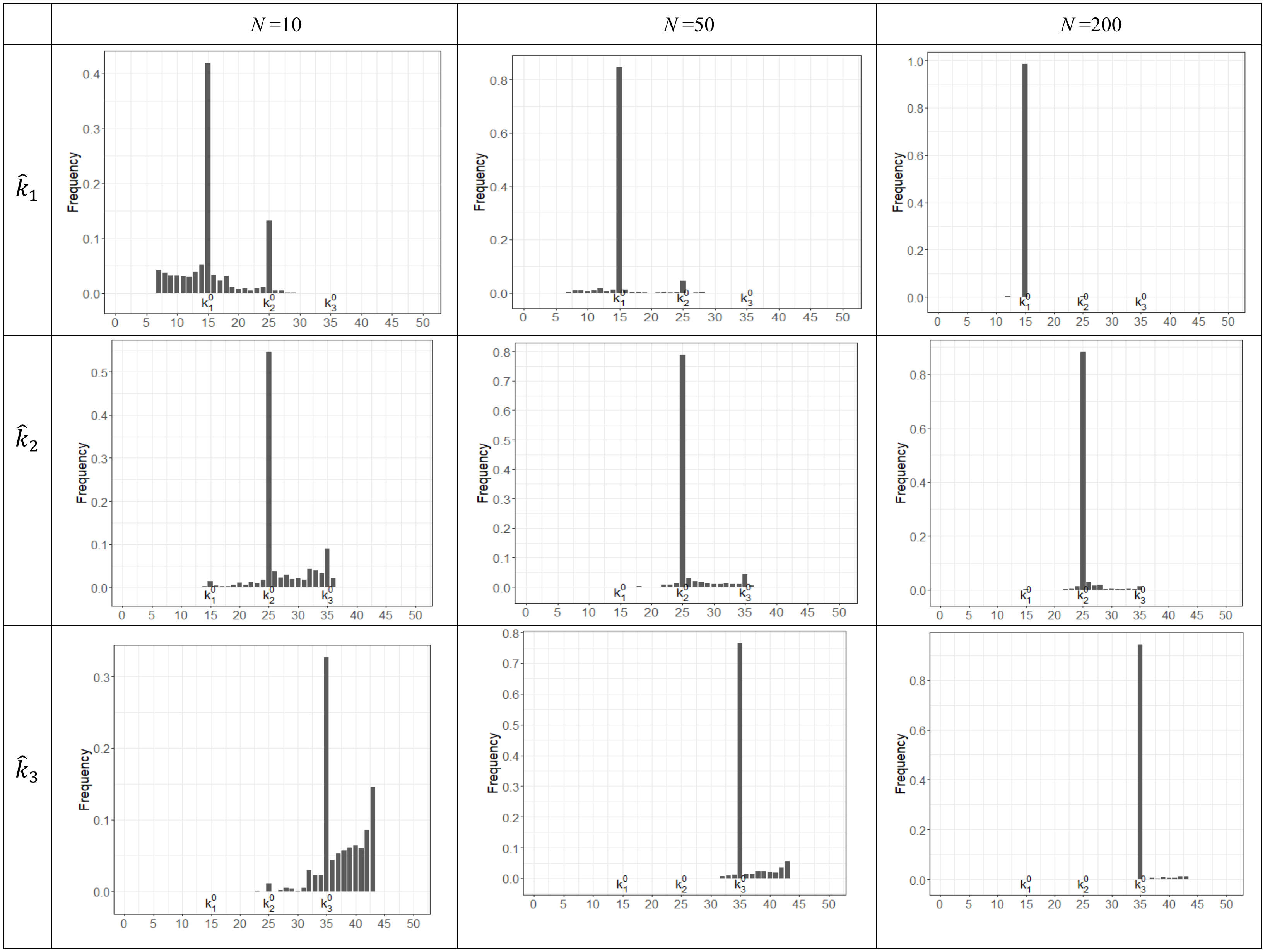
Note: $y_{it} = \alpha_i + \beta_{it}(k_1^0, k_2^0)x_{i,t} + \gamma_{1,it}(k_3^0)f_t + \varepsilon_{it}$, with $\beta_{it}(k_1^0, k_2^0) = \begin{cases} \beta_i, & t = 1, \dots, k_1^0 \\ \beta_i + \Delta\beta_i, & t = k_1^0 + 1, \dots, k_2^0, \\ \beta_i + 2\Delta\beta_i, & t = k_2^0 + 1, \dots, T. \end{cases}$ $\alpha_i \sim iidN(1,1)$, $\beta_i \sim iidN(1, 0.04)$, $\Delta\beta_i \sim iidN(0, 0.5)$, $x_{it} = a_i + \gamma_{2,i}f_t + v_{it}$, where $a_i \sim iidN(0.5, 0.5)$, $\gamma_{2,i} \sim iidN(0.5,0.5)$ and $v_{it} = \rho_{vi}v_{i,t-1} + \psi_{it}, \psi_{it} \sim iidN(0, 1 - \rho_{vi}^2), v_{i,-50} = 0$, $\rho_{vi} \sim iidU[0.05,0.95]$. $f_t = f_{t-1} + v_{ft}, t = -50, \dots, T$, $v_{ft} \sim iidN(0, 1 - \rho_f^2)$, $f_{-50} = 0$. $\gamma_{1,it}(k_3^0) = \begin{cases} \gamma_{1,i}, & t = 1, \dots, k_3^0, \\ \gamma_{1,i} + \Delta\gamma_i, & t = k_3^0 + 1, \dots, T, \end{cases}$ $\gamma_{1,i} \sim iidN(1,0.2)$, $\Delta\gamma_i \sim iidN(0.5, 0.5)$. $\varepsilon_{it} = \begin{cases} \rho_{i\varepsilon}\varepsilon_{i,t-1} + \sigma_i(1 - \rho_{i\varepsilon}^2)^{0.5}\omega_{it}, & i = 1, \dots, [N/2], \\ \sigma_i(1 + \theta_{i\varepsilon}^2)^{-0.5}(\omega_{it} + \theta_{i\varepsilon}\omega_{i,t-1}), & i = [N/2] + 1, \dots, T, \end{cases}$ $\sigma_i^2 \sim iidU(0.5,1.5), \omega_{1i} \sim iidN(0,1)$, $\rho_{i\varepsilon} \sim iidU[0.05, 0.95]$, $\theta_{i\varepsilon} \sim iidU[0,1]$. These variables are mutually independent. The replication number is 1,000. $T = 50, k_1^0 = [0.3T] = 15, k_2^0 = [0.5T] = 25, k_3^0 = [0.7T] = 35$. $\hat{k}_j$: The OLS estimator of the change point $k_j^0, j = 1,2,3$.

Figure 2: Histograms of Break Point Estimators in Case 1 Under Rank Deficiency ($T = 50$)



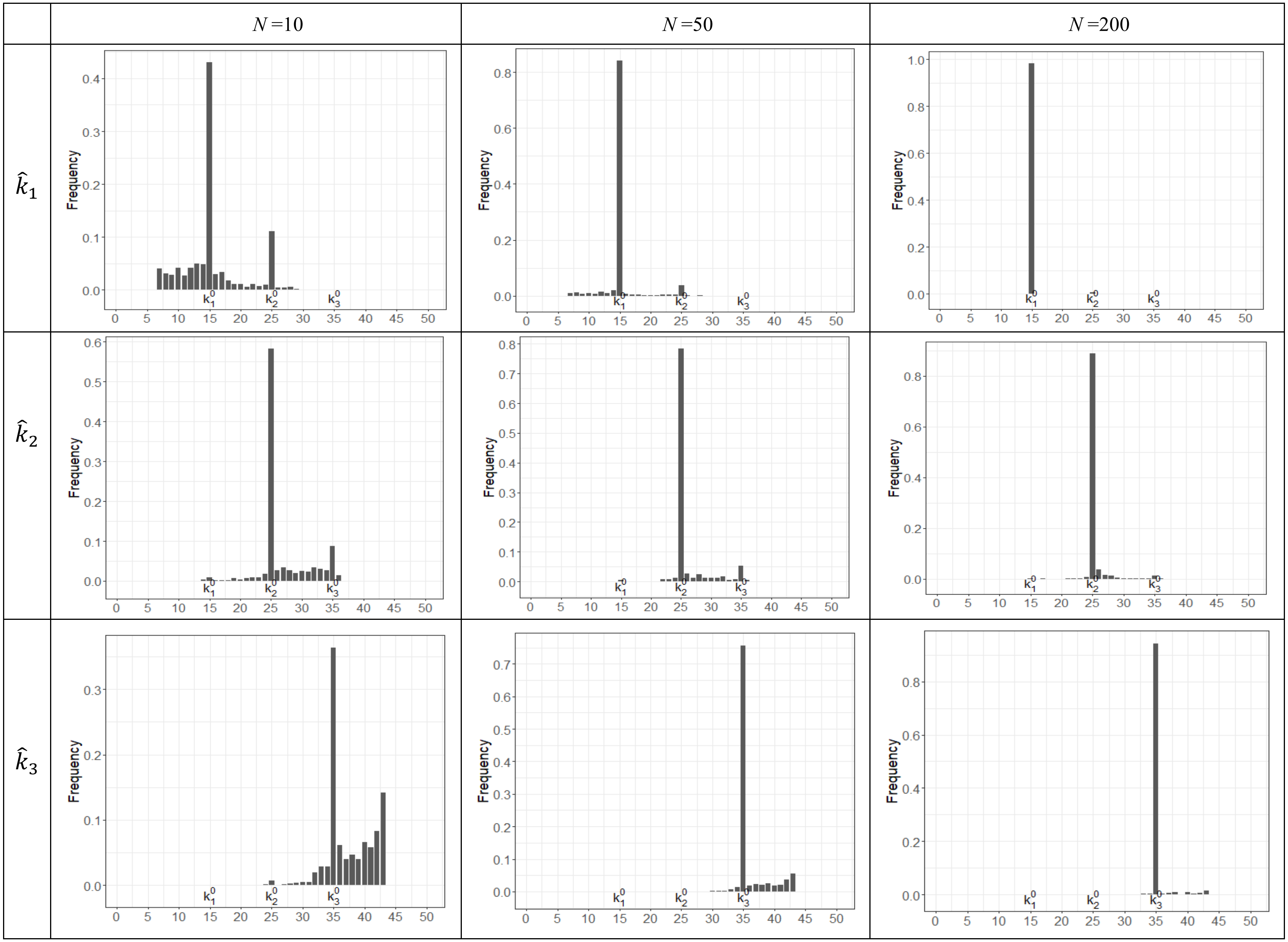
Note: The DGP is the same as that in Figure 1, except that the means of a_i and $\gamma_{2,i}$ change to zero, i.e., $a_i \sim iidN(0, 0.5)$, $\gamma_{2i} \sim iidN(0, 0.5)$. In the current design, the rank condition is not satisfied asymptotically.

Figure 3: Histograms of Break Point Estimators in Case 2 ($T = 50$)



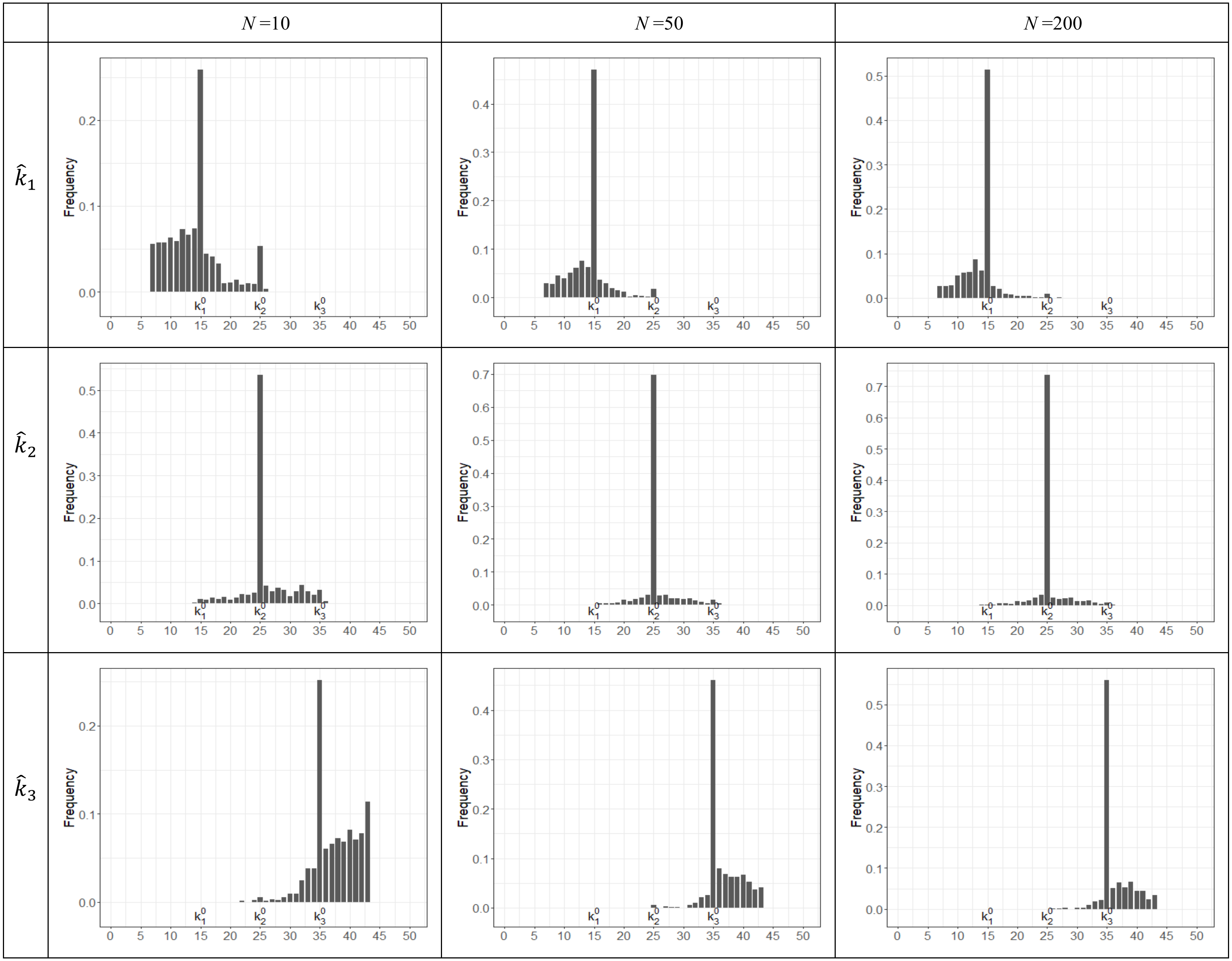
Note: The DGP is **the** same as that of Figure 1, except nonstationary $v_{it} = v_{i,t-1} + \psi_{it}, \psi_{it} \sim iidN(0, 1 - \rho_{vi}^2), v_{i,-50} = 0$,.

Figure 4: Histograms of Break Point Estimators in Case 2 with Stationary Factors ($T = 50$)



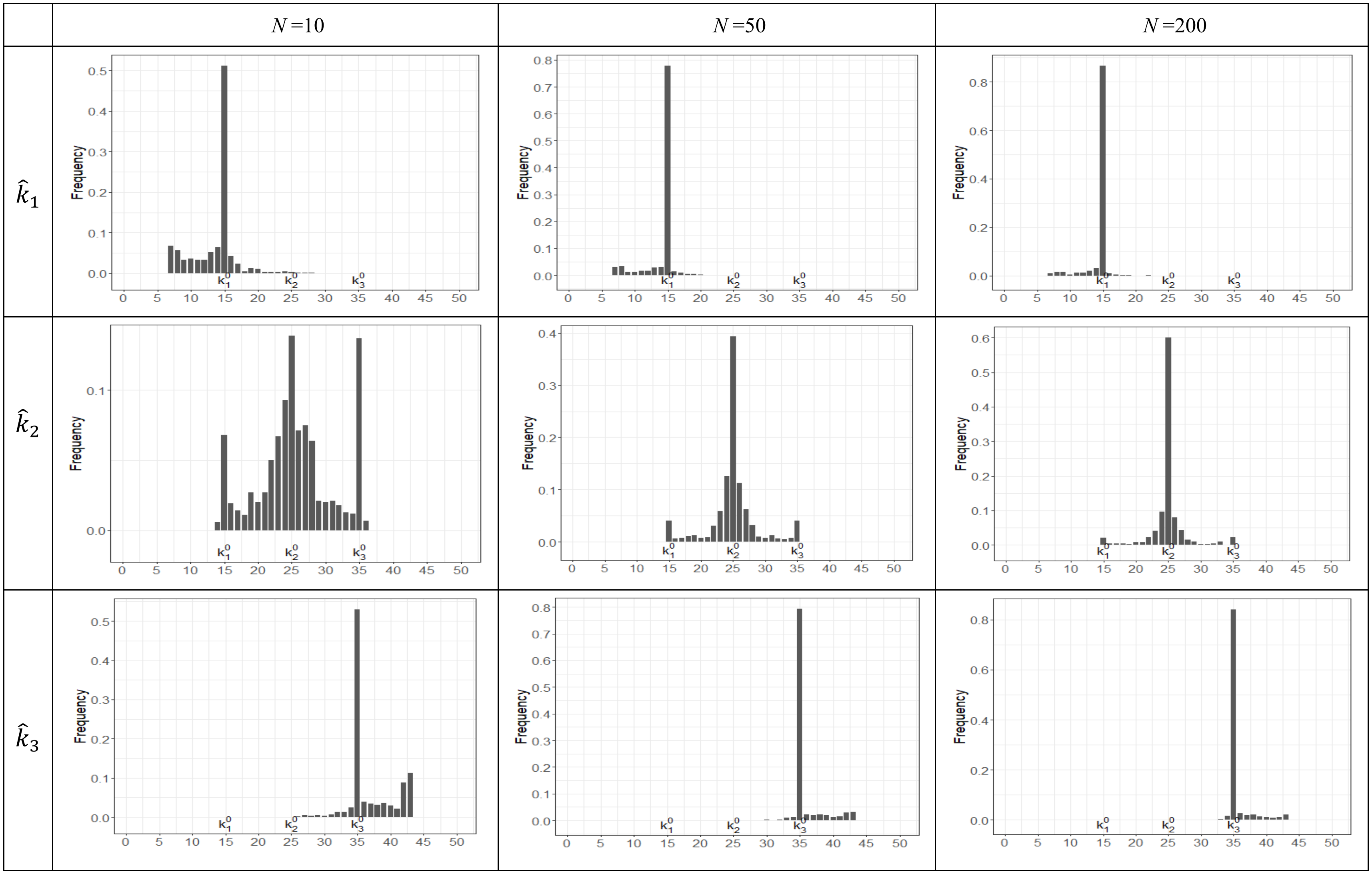
Note: The DGP **is the** same as that of Figure 3, except stationary factors, $f_t = 0.5f_{t-1} + v_{ft}, t = -49, \dots, 0, 1, \dots T$, $v_{ft} \sim iidN(0, 1 - \rho_f^2)$ with $\rho_f = 0.5$, $f_{-50} = 0$.

Figure 5: Histograms of Break Point Estimators in Case 2 with Nonstationary Error ($T = 50$)



Note: The DGP is **the** same as that of Figure 3, except nonstationary errors, $\varepsilon_{it} = \varepsilon_{i,t-1} + \vartheta_{it}, t = -49, \dots, 0, 1, \dots T$, $\vartheta_{it} \sim iidN(0, 1 - \rho_{\varepsilon i}^2), \varepsilon_{i,-50} = 0$.

Figure 6: Histograms of Break Point Estimators in Case 1 with Mixed Stationary and Nonstationary Regressors ($T = 50$)



Note: An additional regressor and factor are added in the DGP used in Figure 1 to allow for mixed stationary and nonstationary regressors. $y_{it} = \alpha_i + \beta_{1,it}(k_1^0)x_{1,it} + \beta_{2,it}(k_2^0)x_{2,it} + \gamma_{11,it}(k_3^0)f_{1,t} + \gamma_{12,it}(k_3^0)f_{2,t} + \varepsilon_{it}$, where $x_{1,it} = a_i + \gamma_{21,i}f_{1,t} + \gamma_{22,i}f_{2,t} + v_{1,it}$, $x_{2,it} = a_i + \gamma_{23,i}f_{2,t} + v_{2,it}$. $\gamma_{21,i}, \gamma_{22,i}, \gamma_{23,i} \sim iidN(0.5, 0.5)$. Two factors $f_{1,t} = f_{1,t-1} + v_{1,ft}$, $f_{2,t} = 0.5f_{2,t-1} + v_{2,ft}$, $f_{1,-50} = f_{2,-50} = 0$. $k_1^0 = 0.3T, k_2^0 = 0.5T, k_3^0 = 0.7T$. $\gamma_{11,it}(k_3^0)$, $\gamma_{12,it}(k_3^0)$ have the same design as $\gamma_{1,it}(k_3^0)$ in Figure 1 except the variance of the shift term changes from 0.5 to 0.16. $\beta_{1,it}(k_1^0) = \begin{cases} \beta_{11,i}, & t = 1, \ldots, k_1^0 \\ \beta_{11,i} + \Delta\beta_{1,i}, & t = k_1^0 + 1, \ldots, T, \end{cases}$ and $\beta_{2,it}(k_2^0) = \begin{cases} \beta_{21,i}, & t = 1, \ldots, k_2^0 \\ \beta_{21,i} + \Delta\beta_{2,i}, & t = k_2^0 + 1, \ldots, T, \end{cases}$ with $\Delta\beta_{1,i} \sim iidN(0, 0.16), \Delta\beta_{2,i} \sim iidN(0, 0.16)$.

Supplementary Appendices: Proofs of Theorems and Lemmas (not for publication)

Nonstationary Heterogeneous Panels with Multiple Structural Changes

by Badi H. Baltagi, Qu Feng, Wei Wang

The supplementary appendices include detailed proofs of the main results in the text. To simplify notation, in this section we consider the case of three breaks, $m = 3$, including two in the slopes, (k_1^0, k_2^0), and one in the error factor loadings, k_3^0. The proofs of the general case in model (10) can be presented at the cost of additional notation.

Specifically, Appendix A includes detailed proofs of Theorems 1 and 2, Propositions sub-regimes1-4. Subsection A.1 provides necessary Lemmas and detailed proof of Theorem 1. Similarly, subsection A.2 provides necessary lemmas and proof of Theorem 2. Lastly, Subsection A.3 provides proofs of Propositions 1- 2, and A.4 provides proofs of Propositions 3-4 respectively. Detailed proofs of lemmas are collected in the supplementary Appendix B. Additional figures of simulations are also attached in the last.

Appendix A: Proofs of Theorems and Propositions

A.1 Proof of Theorem 1

Proof of Theorem 1.

Following Bai and Perron (1998), we decompose the analysis of multiple breaks into several problems involving a single structural change in each. Without loss of generality, we only provide the proof of $\lim_{(N,T)\to\infty} P(\hat{k}_1 = k_1^0) = 1$. The proof of $\lim_{(N,T)\to\infty} P(\hat{k}_j = k_j^0) = 1$, $j = 2, 3$, can be shown similarly and is omitted.

To show $\hat{k}_1 - k_1^0 \xrightarrow{p} 0$, it is equivalent to show that for any given $\epsilon > 0$, for both large T and N, $P(|\hat{k}_1 - k_1^0| \geq 1) < \epsilon$. As in BFK (2016), we assume that $\hat{k}_1 - k_1^0$, $\hat{k}_2 - k_2^0$ and $\hat{k}_3 - k_3^0$ are bounded here for simplicity.[18]

Under Assumption 1 and that the estimators of break fractions are consistent, we consider the set $K(C_k) = \{(k_1, k_2, k_3) : 1 \leq |k_1 - k_1^0|, |k_j - k_j^0| \leq C_k, aT \leq k_j \leq$

[18] With an observed proxy for f_t, each series is considered a time series model with multiple breaks. Based on Bai and Perron's (1998) finding, here we assume that estimated breaks are bounded.

$(1-a)T, j=1,2,3\}$ for a finite constant C_k and $a>0$. By definition, $S(k_1,k_2,k_3) = \sum_{i=1}^N SSR_i(k_1,k_2,k_3)$ is minimized globally at $(\hat{k}_1,\hat{k}_2,\hat{k}_3)$, i.e., $S(\hat{k}_1,\hat{k}_2,\hat{k}_3) \leq S(k_1^0,\hat{k}_2,\hat{k}_3)$ with probability 1.

Therefore, we examine the behavior of $S(k_1,k_2,k_3)$ on the set $K(C_k)$. It is sufficient to show that for each $\epsilon>0$, for both large T and N, $P(\min_{K(C_k)}[S(k_1,k_2,k_3) - S(k_1^0,k_2,k_3)] \leq 0) < \epsilon$. Without loss of generality, assume $k_1 < k_1^0 < k_2$,

$$\begin{aligned}
&S(k_1,k_2,k_3) - S(k_1^0,k_2,k_3)\\
=&[S(k_1,k_2,k_3) - S(k_1,k_0^0,k_2,k_3)] - [S(k_1^0,k_2,k_3) - S(k_1,k_1^0,k_2,k_3)]. \qquad (27)\\
=&\textstyle\sum_{i=1}^N \left[SSR_i(k_1,k_2,k_3) - SSR_i(k_1,k_1^0,k_2,k_3)\right]\\
&-\textstyle\sum_{i=1}^N \left[SSR_i(k_1^0,k_2,k_3) - SSR_i(k_1,k_1^0,k_2,k_3)\right],
\end{aligned}$$

where, $SSR_i(k_1,k_1^0,k_2,k_3)$ is the sum of squared residuals in the regression with four breaks at (k_1,k_1^0,k_2,k_3) for series i and $S(k_1,k_1^0,k_2,k_3) = \sum_{i=1}^N SSR_i(k_1,k_1^0,k_2,k_3)$. Thus, the analysis of a three-break or multiple break problem can be decomposed into two problems involving a single break. The first term $SSR_i(k_1,k_2,k_3) - SSR_i(k_1,k_1^0,k_2,k_3)$ allows an additional fourth break k_1^0 between k_1 and k_2, and the second term $SSR_i(k_1^0,k_2,k_3) - SSR_i(k_1,k_1^0,k_2,k_3)$ adds an additional fourth break at k_1 between 1 and k_1^0. Thus, it is convenient to derive each part above as a single common break issue in panel data as in BFK (2016).

Following Bai and Perron (1998), we denote $\hat{\delta}_i(\hat{k}_1,\hat{k}_2,\hat{k}_3) = \left(\hat{\delta}_{i1}',\hat{\delta}_{i2}',\hat{\delta}_{i3}',\hat{\delta}_{i4}'\right)'$ the estimator of $(\delta_{i1},\delta_{i2},\delta_{i3},\delta_{i4})$ in the regression with three breaks k_1, k_2 and k_3, and $(\hat{\delta}_{i1}^*,\hat{\delta}_{i\Delta},\hat{\delta}_{i2}^*,\hat{\delta}_{i3}^*,\hat{\delta}_{i4}^*)$ the estimator of $(\delta_{i1},\delta_{i1},\delta_{i2},\delta_{i3},\delta_{i4})$ based on the partition $(k_1,\ k_1^0,\ k_2,\ k_3)$. In particular, $\hat{\delta}_{i1}^*$ is an estimate of δ_{i1} associated with regressor $(z_{i1},...,z_{i,k_1},0,...,0)'$, $\hat{\delta}_{i\Delta}$ is the estimate of δ_{i1} associated with regressor $Z_{i\Delta} = (0,...,0,z_{i,k_1+1},...,z_{i,k_1^0},0,...,0)'$, and $\hat{\delta}_{i2}^*$ is the estimate of δ_{i2} associated with regressor $(0,...,0,z_{i,k_1^0+1},...,z_{i,k_2},0,...,0)'$. $\hat{\delta}_{i3}^*$, $\hat{\delta}_{i4}^*$ can be defined similarly.

By definition,

$$\begin{aligned}
SSR_i(k_1,k_2,k_3) \;=\;& \sum_{t=1}^{k_1}\left(y_{it} - z_{it}'\hat{\delta}_{i1}\right)^2 + \sum_{t=k_1+1}^{k_2}\left(y_{it} - z_{it}'\hat{\delta}_{i2}\right)^2\\
&+\sum_{t=k_2+1}^{k_3}\left(y_{it} - z_{it}'\hat{\delta}_{i3}\right)^2 + \sum_{t=k_3+1}^{T}\left(y_{it} - z_{it}'\hat{\delta}_{i4}\right)^2,
\end{aligned}$$

and

$$
\begin{aligned}
SSR_i(k_1, k_1^0, k_2, k_3) &= \sum_{t=1}^{k_1} \left(y_{it} - z_{it}'\hat{\delta}_{i1}^*\right)^2 + \sum_{t=k_1+1}^{k_1^0} \left(y_{it} - z_{it}'\hat{\delta}_{i\Delta}\right)^2 \\
&+ \sum_{t=k_1^0+1}^{k_2} \left(y_{it} - z_{it}'\hat{\delta}_{i2}^*\right)^2 + \sum_{t=k_2+1}^{k_3} \left(y_{it} - z_{it}'\hat{\delta}_{i3}^*\right)^2 + \sum_{t=k_3+1}^{T} \left(y_{it} - z_{it}'\hat{\delta}_{i4}^*\right)^2.
\end{aligned}
$$

It's worth noting that $\hat{\delta}_{i1}$ and $\hat{\delta}_{i1}^*$ are the estimators associated with same regressor $(z_{i1}, ..., z_{i,k_1}, 0, ..., 0)'$, thus, $\hat{\delta}_{i1} = \hat{\delta}_{i1}^*$. Similarly, $\hat{\delta}_{i3} = \hat{\delta}_{i3}^*$, $\hat{\delta}_{i4} = \hat{\delta}_{i4}^*$. Thus,

$$
\begin{aligned}
& SSR_i(k_1, k_2, k_3) - SSR_i(k_1, k_1^0, k_2, k_3) \qquad (28) \\
= & \sum_{t=k_1+1}^{k_2} \left(y_{it} - z_{it}'\hat{\delta}_{i2}\right)^2 - \left[\sum_{t=k_1+1}^{k_1^0} \left(y_{it} - z_{it}'\hat{\delta}_{i\Delta}\right)^2 + \sum_{t=k_1^0+1}^{k_2} \left(y_{it} - z_{it}'\hat{\delta}_{i2}^*\right)^2\right].
\end{aligned}
$$

Since the term $SSR_i(k_1, k_2, k_3) - SSR_i(k_1, k_1^0, k_2, k_3)$ involves a regression with a break k_1^0 between k_1 and k_2, we focus on the interval $[k_1+1, k_2]$. k_1^0 splits $[k_1+1, k_2]$ into two parts $[k_1+1, k_1^0]$ and $[k_1^0+1, k_2]$. These three intervals are referred to as $\bigstar$, Δ and $\bigstar - \Delta$, respectively, i.e., $\bigstar = [\Delta, \bigstar - \Delta]$. Under the current assumptions, the number of observations on interval Δ is finite, different from that on $\bigstar$ or $\bigstar - \Delta$. Define $Y_{i\bigstar} = (y_{i,k_1+1}, ..., y_{i,k_2})'$, $Y_{i\Delta} = (y_{i,k_1+1}, ..., y_{i,k_1^0}, 0, ..., 0)'$ and $Y_{i(\bigstar-\Delta)} = Y_{i\bigstar} - Y_{i\Delta} = (0, ..., 0, y_{i,k_1^0+1}, ..., y_{i,k_2})'$. $Z_{i\bigstar}$, $\varepsilon_{i\bigstar}^*$, $Z_{i\Delta}$, $Z_{i(\bigstar-\Delta)}$ can be defined in the same fashion. By construction, $Y_{i\Delta}'Y_{i(\bigstar-\Delta)} = 0$ and $Z_{i\Delta}'Z_{i(\bigstar-\Delta)} = 0$.

Recall that the OLS estimators of $(\delta_{i1}, \delta_{i2})$ on intervals of $[k_1+1, k_1^0]$ and $[k_1^0+1, k_2]$ are $\hat{\delta}_{i\Delta}, \hat{\delta}_{i2}^*$, respectively. Without considering a break in slopes on the interval $[k_1 + 1, k_2]$, the OLS estimator for δ_{i2} is $\hat{\delta}_{i2}$. The first term in (28), $\sum_{t=k_1+1}^{k_2}(y_{it} - z_{it}'\hat{\delta}_{i2})^2 = [Y_{i\bigstar} - Z_{i\bigstar}\hat{\delta}_{i2}]'[Y_{i\bigstar} - Z_{i\bigstar}\hat{\delta}_{i2}]$ is the sum of squared residuals in the regression of y on z for series i using time series sample on the interval $[k_1 + 1, k_2]$. The second term in equation (28)

$$
\begin{aligned}
& \textstyle\sum_{t=k_1+1}^{k_1^0}(y_{it} - z_{it}'\hat{\delta}_{i\Delta})^2 + \sum_{t=k_1^0+1}^{k_2}(y_{it} - z_{it}'\hat{\delta}_{i2}^*)^2 \\
& = \textstyle\sum_{t=k_1+1}^{k_1^0}(y_{it} - z_{it}'\hat{\delta}_{i\Delta})^2 + \sum_{t=k_1^0+1}^{k_2}(y_{it} - z_{it}'\hat{\delta}_{i\Delta} + z_{it}'(\hat{\delta}_{i\Delta} - \hat{\delta}_{i2}^*))^2 \\
& = [Y_{i\bigstar} - Z_{i\bigstar}\hat{\delta}_{i\Delta} - Z_{i(\bigstar-\Delta)}(\hat{\delta}_{i2}^* - \hat{\delta}_{i\Delta})]'[Y_{i\bigstar} - Z_{i\bigstar}\hat{\delta}_{i\Delta} - Z_{i(\bigstar-\Delta)}(\hat{\delta}_{i2}^* - \hat{\delta}_{i\Delta})]
\end{aligned}
$$

is the sum of squared residuals in the regression of y on z for series i with a break k_1^0 on the interval $[k_1 + 1, k_2]$. Thus, according to Amemiya (1985, p. 31),

$$
\begin{aligned}
SSR_i(k_1, k_2, k_3) - SSR_i(k_1, k_1^0, k_2, k_3) &= (\hat{\delta}_{i2}^* - \hat{\delta}_{i\Delta})'Z_{i(\bigstar-\Delta)}'M_{Z_{i\bigstar}}Z_{i(\bigstar-\Delta)}(\hat{\delta}_{i2}^* - \hat{\delta}_{i\Delta}) \\
&= (\hat{\delta}_{i2}^* - \hat{\delta}_{i\Delta})'Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}(\hat{\delta}_{i2}^* - \hat{\delta}_{i\Delta}).
\end{aligned}
$$

The second equality above is due to the facts of $Z_{i(\bigstar-\Delta)} = Z_{i\bigstar} - Z_{i\Delta}$ and

$$Z'_{i(\bigstar-\Delta)} M_{Z_{i\bigstar}} Z_{i(\bigstar-\Delta)} = Z'_{i\Delta} M_{Z_{i\bigstar}} Z_{i\Delta},$$

where $M_{Z_{i\bigstar}} = I_{k_2-k_1+1} - Z_{i\bigstar}\left(Z'_{i\bigstar} Z_{i\bigstar}\right) Z'_{i\bigstar}$ and $I_{(k_2-k_1+1)}$ is the $(k_2-k_1+1)\times(k_2-k_1+1)$ identity matrix. Next, following BFK (2016) we derive the expression of $SSR_i(k_1,k_2,k_3) - SSR_i(k_1,k_1^0,k_2,k_3)$.

For $t \in [k_1+1, k_1^0]$, $\hat{\delta}_{i\Delta} = (Z'_{i\Delta} Z_{i\Delta})^{-1} Z'_{i\Delta} Y_{i\Delta}$ and $\hat{\delta}^*_{i2} = (Z'_{i(\bigstar-\Delta)} Z_{i(\bigstar-\Delta)})^{-1} Z'_{i(\bigstar-\Delta)} Y_{i(\bigstar-\Delta)}$ for $t \in [k_1^0+1, k_2]$. Partitioned regression gives

$$\begin{aligned}\hat{\delta}^*_{i2} - \hat{\delta}_{i\Delta} &= (Z'_{i(\bigstar-\Delta)} M_{Z_{i\bigstar}} Z_{i(\bigstar-\Delta)})^{-1} Z'_{i(\bigstar-\Delta)} M_{Z_{i\bigstar}} Y_{i\bigstar} \\ &= -(Z'_{i\Delta} M_{Z_{i\bigstar}} Z_{i\Delta})^{-1} Z'_{i\Delta} M_{Z_{i\bigstar}} Y_{i\bigstar}.\end{aligned}$$

Plugging $Y_{i\bigstar} = Z_{i\bigstar}\delta_{i1} + Z_{i(\bigstar-\Delta)}(\delta_{i2}-\delta_{i1}) + \varepsilon^*_{i\bigstar}$ into the equation above gives,

$$\begin{aligned}\hat{\delta}^*_{i2} - \hat{\delta}_{i\Delta} &= (\delta_{i2}-\delta_{i1}) + \left(Z'_{i(\bigstar-\Delta)} M_{Z_{i\bigstar}} Z_{i(\bigstar-\Delta)}\right)^{-1} Z'_{i(\bigstar-\Delta)} M_{Z_{i\bigstar}} \varepsilon^*_{i\bigstar} \\ &= (\delta_{i2}-\delta_{i1}) - \left(Z'_{i\Delta} M_{Z_{i\bigstar}} Z_{i\Delta}\right)^{-1} Z'_{i\Delta} M_{Z_{i\bigstar}} \varepsilon^*_{i\bigstar}.\end{aligned} \quad (29)$$

Thus, we can get

$$\begin{aligned}SSR_i(k_0,k_1) - SSR_i(k_0,k_0^0,k_1) &= (\delta_{i2}-\delta_{i1})' Z'_{i\Delta} M_{Z_{i\bigstar}} Z_{i\Delta}(\delta_{i2}-\delta_{i1}) \\ &\quad -2(\delta_{i2}-\delta_{i1})' Z'_{i\Delta} M_{Z_{i\bigstar}} \varepsilon^*_{i\bigstar} \\ &\quad +\varepsilon^{*\prime}_{i\bigstar} M_{Z_{i\bigstar}} Z_{i\Delta}\left(Z'_{i\Delta} M_{Z_{i\bigstar}} Z_{i\Delta}\right)^{-1} Z'_{i\Delta} M_{Z_{i\bigstar}} \varepsilon^*_{i\bigstar}.\end{aligned} \quad (30)$$

Similarly, the second term $SSR_i(k_1^0,k_2,k_3) - SSR_i(k_1,k_1^0,k_2,k_3)$ in (27) involves a regression with a break at k_1 between 1 and k_1^0. Denote the interval $[1,k_1^0]$ by $\Diamond$. k_1 splits $[1,k_1^0]$ into two parts $[1,k_1]$ and $[k_1+1,k_1^0]$. Note that the latter interval has been denoted as Δ above. Similarly, define $Y_{i\Diamond} = (y_{i,1},...,y_{i,k_1^0})'$, $Z_{i\Diamond}$ and $\varepsilon^*_{i\Diamond}$ on the interval $\Diamond$. The number of observations on the interval $\Diamond$ is unbounded under Assumption 1 as $T\to\infty$. Note that there is no true break in slopes on the interval $[1,k_1^0]$ and the corresponding true slope parameter is δ_{i1}. The OLS estimators of $(\delta_{i1},\delta_{i1})$ on intervals of $[1,k_1]$ and $[k_1+1,k_1^0]$ are $\hat{\delta}^*_{i1}$, $\hat{\delta}_{i\Delta}$, respectively. As in equation (27), we can obtain

$$SSR_i(k_1^0,k_2,k_3) - SSR_i(k_1,k_1^0,k_2,k_3) = (\hat{\delta}_{i\Delta} - \hat{\delta}^*_{i1})' Z'_{i\Delta} M_{Z_{i\Diamond}} Z_{i\Delta}(\hat{\delta}_{i\Delta} - \hat{\delta}^*_{i1}).$$

Partitioned regression gives $\hat{\delta}_{i\Delta} - \hat{\delta}^*_{i1} = (Z'_{i\Delta} M_{Z_{i\Diamond}} Z_{i\Delta})^{-1} Z'_{i\Delta} M_{Z_{i\Diamond}} Y_{i\Diamond}$, where $M_{Z_{i\Diamond}} = I_{k_1^0} - Z_{i\Diamond}(Z'_{i\Diamond} Z_{i\Diamond})^{-1} Z'_{i\Diamond}$. Plugging $Y_{i\Diamond} = Z_{i\Diamond}\delta_{i1} + \varepsilon^*_{i\Diamond}$ into the equation above gives

$$\hat{\delta}_{i\Delta} - \hat{\delta}^*_{1i} = (Z'_{i\Delta} M_{Z_{i\Diamond}} Z_{i\Delta})^{-1} Z'_{i\Delta} M_{Z_{i\Diamond}} \varepsilon^*_{i\Diamond}. \quad (31)$$

Since there is no break in slopes on the interval $[1, k_1^0]$, no slope shift term appears in (31), which is different from (29). Thus, we can get

$$SSR_i(k_1^0, k_2, k_3) - SSR_i(k_1, k_1^0, k_2, k_3) = \varepsilon_{i\diamondsuit}^{*\prime} M_{Z_{i\diamondsuit}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\diamondsuit}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\diamondsuit}} \varepsilon_{i\diamondsuit}^{*}. \tag{32}$$

Combining equations (30) and (32), we obtain,

$$\begin{aligned}
& S(k_1, k_2, k_3) - S(k_1^0, k_2, k_3) \\
= \; & \textstyle\sum_{i=1}^N \left[S_i(k_1, k_2, k_3) - S_i(k_1, k_1^0, k_2, k_3)\right] - \sum_{i=1}^N \left[S_i(k_1^0, k_2, k_3) - S_i(k_1, k_1^0, k_2, k_3)\right] \\
= \; & \textstyle\sum_{i=1}^N \left(\delta_{i2} - \delta_{i1}\right)' Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta} \left(\delta_{i2} - \delta_{i1}\right) - 2\sum_{i=1}^N \left(\delta_{i2} - \delta_{i1}\right)' Z_{i\Delta}' M_{Z_{i\bigstar}} \varepsilon_{i\bigstar}^{*} \\
& \textstyle + \sum_{i=1}^N \varepsilon_{i\bigstar}^{*\prime} M_{Z_{i\bigstar}} Z_{i\Delta} \left(Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta}\right)^{-1} Z_{i\Delta}' M_{Z_{i\bigstar}} \varepsilon_{i\bigstar}^{*} \\
& \textstyle - \sum_{i=1}^N \varepsilon_{i\diamondsuit}^{*\prime} M_{Z_{i\diamondsuit}} Z_{i\Delta} \left(Z_{i\Delta}' M_{Z_{i\diamondsuit}} Z_{i\Delta}\right)^{-1} Z_{i\Delta}' M_{Z_{i\diamondsuit}} \varepsilon_{i\diamondsuit}^{*}.
\end{aligned}$$

Like in Bai (1997) and BFK (2016), here $S(k_1, k_2, k_3) - S(k_1^0, k_2, k_3)$ can be expressed as the sum of a deterministic part $\sum_{i=1}^N J_{1i}(k_1, k_2, k_3)$ and a stochastic term $-\sum_{i=1}^N J_{2i}(k_1, k_2, k_3)$, where $J_{1i}(k_1, k_2, k_3) = (\delta_{i2} - \delta_{i1})' Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta}(\delta_{i2} - \delta_{i1})$,

$$\begin{aligned}
J_{2i}(k_1, k_2, k_3) = & [2(\delta_{i2} - \delta_{i1})' Z_{i\Delta}' M_{Z_{i\bigstar}} \varepsilon_{i\bigstar}^{*}] - [\varepsilon_{i\bigstar}^{*\prime} M_{Z_{i\bigstar}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\bigstar}} \varepsilon_{i\bigstar}^{*}] \\
& + [\varepsilon_{i\diamondsuit}^{*\prime} M_{Z_{i\diamondsuit}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\diamondsuit}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\diamondsuit}} \varepsilon_{i\diamondsuit}^{*}].
\end{aligned}$$

Thus, $S(k_1, k_2, k_3) - S(k_1^0, k_2, k_3) = \sum_{i=1}^N J_{1i}(k_1, k_2, k_3) - \sum_{i=1}^N J_{2i}(k_1, k_2, k_3)$.

To prove Theorem 1 and the statement $P(\min_{K(C_k)}[S(k_1, k_2, k_3) - S(k_1^0, k_2, k_3)] \leq 0) < \epsilon$ for both large T and N, it suffices to show

$$P(\sup_{K(C_k)} |\frac{1}{T} \textstyle\sum_{i=1}^N J_{2i}(k_1, k_2, k_3)| \geq \inf_{K(C_k)} \frac{1}{T} \sum_{i=1}^N J_{1i}(k_1, k_2, k_3)) < \epsilon. \tag{33}$$

Consider the term $\frac{1}{T} Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta}$ in $J_{1i}(k_1, k_2, k_3)$. Since $Z_{i\bigstar} = Z_{i\Delta} + Z_{i(\bigstar-\Delta)}$ and $Z_{i\Delta}' Z_{i(\bigstar-\Delta)} = 0$,

$$\begin{aligned}
T^{-1} Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta} &= T^{-1} Z_{i\Delta}' Z_{i\Delta} - T^{-1} Z_{i\Delta}' Z_{i\bigstar}(Z_{i\bigstar}' Z_{i\bigstar})^{-1} Z_{i\bigstar}' Z_{i\Delta} \\
&= T^{-1} Z_{i\Delta}' Z_{i\Delta} - T^{-2} Z_{i\Delta}' Z_{i\Delta}(T^{-2} Z_{i\bigstar}' Z_{i\bigstar})^{-1} T^{-1} Z_{i\Delta}' Z_{i\Delta}.
\end{aligned}$$

Note that the numbers of observations on the intervals of $\bigstar$ and Δ are $k_2 - k_1$ and $k_1^0 - k_1$. On the set $K(C_k)$, $k_1^0 - k_1$ is finite, while $k_2 - k_1$ is unbounded as $T \to \infty$. By Lemma 1(i), $\frac{1}{T} Z_{i\Delta}' Z_{i\Delta} = O_p(1)$ and $\frac{1}{T^2} Z_{i\Delta}' Z_{i\Delta}(\frac{1}{T^2} Z_{i\bigstar}' Z_{i\bigstar})^{-1} \frac{1}{T} Z_{i\Delta}' Z_{i\Delta} = o_p(1)$ on $K(C_k)$, thus, $T^{-1} Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta} = T^{-1} Z_{i\Delta}' Z_{i\Delta} + o_p(1)$. Last,

$$\inf_{K(C_k)} \frac{1}{T} \textstyle\sum_{i=1}^N J_{1i}(k_1, k_2, k_3) = \inf_{K(C_k)} \sum_{i=1}^N \left(\delta_{i2} - \delta_{i1}\right)' \left(\frac{1}{T} Z_{i\Delta}' Z_{i\Delta}\right) \left(\delta_{i2} - \delta_{i1}\right) + o_p(1).$$

Under Assumption 6, let a finite $\varrho_{min} > 0$ be the minimum eigenvalue of $\frac{1}{N}\sum_{i=1}^{N}(\frac{1}{T}Z'_{i\Delta}Z_{i\Delta})$ uniformly on $K(C_k)$. Following the proof of Lemma 1 in BFK's (2016) appendix, we obtain

$$\inf_{K(C_k)}\frac{1}{T}\sum_{i=1}^{N}J_{1i}(k_1,k_2,k_3) \geq \varrho_{min}\phi_{N,1},$$

with probability tending to 1 and $\phi_{N,1} = \sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'(\delta_{i2}-\delta_{i1})$. Thus, from equation (33), to prove Theorem 1, it is sufficient to show

$$P(\sup_{K(C_k)}\frac{1}{T\phi_{N,1}}|\sum_{i=1}^{N}J_{2i}(k_1,k_2,k_3)| \geq \varrho_{min}) < \epsilon. \tag{34}$$

By Lemma 2,

$$\begin{aligned} |\sum_{i=1}^{N}J_{2i}(k_1,k_2,k_3)| &\leq |\sum_{i=1}^{N}[2(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar}]| \\ &\quad +|\sum_{i=1}^{N}[\varepsilon^{*\prime}_{i\bigstar}M_{Z_{i\bigstar}}Z_{i\Delta}(Z'_{i\Delta}M_{Z_{i\bigstar}}Z_{i\Delta})^{-1}Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar}]| \\ &\quad +|\sum_{i=1}^{N}[\varepsilon^{*\prime}_{i\diamondsuit}M_{Z_{i\diamondsuit}}Z_{i\Delta}(Z'_{i\Delta}M_{Z_{i\diamondsuit}}Z_{i\Delta})^{-1}Z'_{i\Delta}M_{Z_{i\diamondsuit}}\varepsilon^*_{i\diamondsuit}]| \\ &= O_p(T^{1/2}\phi_{N,1}^{1/2}) + O_p(N). \end{aligned}$$

Thus, $\frac{1}{T\phi_{N,1}}\left|\sum_{i=1}^{N}J_{2i}(k_1,k_2,k_3)\right| = O_p(\frac{1}{\sqrt{T\phi_{N,1}}}) + O_p(\frac{N}{T\phi_{N,1}})$. Under Assumption 9 that $\frac{N}{T\phi_{N,1}} \to 0$, as $(N,T)\to\infty$, the term $\frac{1}{T\phi_{N,1}}|J_2(k_1,k_2,k_3)|$ vanishes for any $(k_1,k_2,k_3)\in K(C_k)$. Therefore, (34) and then Theorem 1 are established.

The following Lemmas 1 and 2 are needed to prove Theorem 1.

Lemma 1 *Under Assumptions 1-5, 7,8, and uniformly over $K(C_k)$, as $(N,T)\to\infty$, for $i=1,...,N$,*

(i) $\frac{1}{T}Z'_{i\Delta}Z_{i\Delta} = O_p(1)$, $\frac{1}{T^2}Z'_{i\bigstar}Z_{i\bigstar} = O_p(1)$;

(ii) $\frac{1}{\sqrt{T}}Z'_{i\Delta}\varepsilon_{i\bigstar} = \frac{1}{\sqrt{T}}Z'_{i\Delta}\varepsilon_{i\Delta} = O_p(1)$, $\frac{1}{T}Z'_{i\bigstar}\varepsilon_{i\bigstar} = O_p(1)$;

(iii) $\frac{1}{\sqrt{T}}Z'_{i\Delta}\varepsilon_{i\diamondsuit} = \frac{1}{\sqrt{T}}Z'_{i\Delta}\varepsilon_{i\Delta} = O_p(1)$, $\frac{1}{T}Z'_{i\diamondsuit}\varepsilon_{i\diamondsuit} = O_p(1)$;

(iv) $\frac{1}{T}\bar{V}'_{\bigstar}\bar{V}_{\bigstar} = O_p\left(\frac{1}{N}\right)$, $\frac{1}{\sqrt{T}}Z'_{i\Delta}\bar{V}_{\bigstar} = O_p\left(\frac{1}{\sqrt{N}}\right)$, $\frac{1}{T}Z'_{i\bigstar}\bar{V}_{\bigstar} = O_p\left(\frac{1}{\sqrt{N}}\right)$.

Lemma 2 *Under Assumptions 1-8, uniformly on $K(C_k)$,*

(i) $\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar} = O_p\left(\sqrt{T\phi_{N,1}}\right)$;

(ii) $\sum_{i=1}^{N}\varepsilon^{*\prime}_{i\bigstar}M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z'_{i\Delta}M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar} = O_p(N)$;

(iii) $\sum_{i=1}^{N}\varepsilon^{*\prime}_{i\diamondsuit}M_{Z_{i\diamondsuit}}Z_{i\Delta}\left(Z'_{i\Delta}M_{Z_{i\diamondsuit}}Z_{i\Delta}\right)^{-1}Z'_{i\Delta}M_{Z_{i\diamondsuit}}\varepsilon^*_{i\diamondsuit} = O_p(N)$.

The proofs of Lemmas 1 and 2 can be found in the supplementary Appendix B.

A.2 Proof of Theorem 2

The proof of Theorem 2 is similar to that of Theorem 1. To obtain the inequality (33) in Case 2 of $I(1)$ v_{it}, Lemmas 3 and 4 are needed.

Lemma 3 *Under Assumptions 1-9 and 10, uniformly on $K(C_k)$ and for each $i = 1, ..., N$, as $(N,T) \to \infty$,*

(i) $\frac{1}{T} Z'_{i\Delta} Z_{i\Delta} = O_p(1)$, $\frac{1}{T^2} Z'_{i\bigstar} Z_{i\bigstar} = O_p(1)$;

(ii) $\frac{1}{\sqrt{T}} Z'_{i\Delta} \varepsilon_{i\bigstar} = \frac{1}{\sqrt{T}} Z'_{i\Delta} \varepsilon_{i\Delta} = O_p(1)$, $\frac{1}{T} Z'_{i\bigstar} \varepsilon_{i\bigstar} = O_p(1)$;

(iii) $\frac{1}{\sqrt{T}} Z'_{i\Delta} \varepsilon_{i\diamondsuit} = \frac{1}{\sqrt{T}} Z'_{i\Delta} \varepsilon_{i\Delta} = O_p(1)$, $\frac{1}{T} Z'_{i\diamondsuit} \varepsilon_{i\diamondsuit} = O_p(1)$;

(iv) $\frac{1}{T^2} \bar{V}'_{\bigstar} \bar{V}_{\bigstar} = O_p\left(\frac{1}{N}\right)$, $\frac{1}{T} Z'_{i\Delta} \bar{V}_{\bigstar} = O_p\left(\frac{1}{\sqrt{N}}\right)$, $\frac{1}{T\sqrt{T}} Z'_{i\bigstar} \bar{V}_{\bigstar} = O_p\left(\frac{1}{\sqrt{N}}\right)$.

Lemma 4 *Under Assumptions 1-9 and 10, uniformly on $K(C_k)$,*

(i) $\sum_{i=1}^N (\delta_{i2} - \delta_{i1})' Z'_{i\Delta} M_{Z_{i\star}} \varepsilon^*_{i\bigstar} = O_p\left(\sqrt{T\phi_{N,1}}\right) + O_p\left(T\sqrt{\frac{\phi_{N,1}}{N}}\right)$;

(ii) $\sum_{i=1}^N \varepsilon^{*\prime}_{i\bigstar} M_{Z_{i\star}} Z_{i\Delta} \left(Z'_{i\Delta} M_{Z_{i\star}} Z_{i\Delta}\right)^{-1} Z'_{i\Delta} M_{Z_{i\star}} \varepsilon^*_{i\bigstar} = O_p(N) + O_p(T)$;

(iii) $\sum_{i=1}^N \varepsilon^{*\prime}_{i\diamondsuit} M_{Z_{i\diamondsuit}} Z_{i\Delta} \left(Z'_{i\Delta} M_{Z_{i\diamondsuit}} Z_{i\Delta}\right)^{-1} Z'_{i\Delta} M_{Z_{i\diamondsuit}} \varepsilon^*_{i\diamondsuit} = O_p(N) + O_p(T)$.

Proof of Theorem 2. As in the proof of Theorem 1, it is suffices to show for any $\epsilon > 0$, for large N and T,

$$P(\sup_{K(C_k)} |\frac{1}{T} \sum_{i=1}^N J_{2i}(k_1, k_2, k_3)| \geq \inf_{K(C_k)} \frac{1}{T} \sum_{i=1}^N J_{1i}(k_1, k_2, k_3)) < \epsilon.$$

In Case 2, the only difference lies in that v_{it} changes from $I(0)$ to $I(1)$. Since $x_{it} = \Gamma'_i f_t + v_{it}$ and $\bar{x}_t = \bar{\Gamma}' f_t + \bar{v}_t$, $z_{it} = (x'_{it}, \bar{x}'_{\cdot t})'$ remains $I(1)$ for $I(1)$ f_t. Thus, with Lemma 3, the following result remains unchanged, $\inf_{K(C_k)} \frac{1}{T} \sum_{i=1}^N J_{1i}(k_1, k_2, k_3) \geq \varrho_{min} \phi_{N,1}$ with probability tending to 1. As in the proof of Theorem 1, we need to show

$$P(\sup_{K(C_k)} \frac{1}{T\phi_{N,1}} |\sum_{i=1}^N J_{2i}(k_1, k_2, k_3)| \geq \varrho_{min}) < \epsilon. \quad (35)$$

By Lemma 4,

$$\begin{aligned} |\sum_{i=1}^N J_{2i}(k_1, k_2, k_3)| &\leq |\sum_{i=1}^N [(\delta_{i2} - \delta_{i1})' Z'_{i\Delta} M_{Z_{i\star}} \varepsilon^*_{i\bigstar}]| \\ &\quad + |\sum_{i=1}^N [\varepsilon^{*\prime}_{i\bigstar} M_{Z_{i\star}} Z_{i\Delta} (Z'_{i\Delta} M_{Z_{i\star}} Z_{i\Delta})^{-1} Z'_{i\Delta} M_{Z_{i\star}} \varepsilon^*_{i\bigstar}]| \\ &\quad + |\sum_{i=1}^N [\varepsilon^{*\prime}_{i\diamondsuit} M_{Z_{i\diamondsuit}} Z_{i\Delta} (Z'_{i\Delta} M_{Z_{i\diamondsuit}} Z_{i\Delta})^{-1} Z'_{i\Delta} M_{Z_{i\diamondsuit}} \varepsilon^*_{i\diamondsuit}]| \\ &= O_p(\sqrt{T\phi_{N,1}}) + O_p(T\phi_{N,1}^{1/2} N^{-1}) + O_p(N) + O_p(T). \end{aligned}$$

Thus,

$$\frac{1}{T\phi_{N,1}} |J_2(k_1, k_2, k_3)| = O_p(T^{-1/2}\phi_{N,1}^{-1/2}) + O_p(N^{-1/2}\phi_{N,1}^{-1/2}) + O_p(NT^{-1}\phi_{N,1}^{-1}) + O_p(\frac{1}{\phi_{N,1}}).$$

Under Assumption 9, $\phi_{N,1} \to \infty$ and $\frac{N}{T\phi_{N,1}} \to 0$, as as $(N,T) \to \infty$, $\frac{1}{T\phi_{N,1}} |J_2(k_1, k_2, k_3)|$ vanishes for any $(k_1, k_2, k_3) \in K(C_k)$. Therefore, (35) is established, and Theorem 2 is proved.

A.3 Proofs of Propositions 1 and 2

In this subsection, we also assume $m=3$, including two breaks k_1^0, k_2^0 in slopes and a third one k_3^0 in error factor loadings. Let

$$\underline{V}_i(k_1^0,k_2^0) = diag\left((v'_{i1},...,v'_{i,k_1^0})',(v'_{i,(k_1^0+1)},...,v'_{i,k_2^0})',(v'_{i,k_2^0+1},...,v'_{iT})'\right).$$

in order to prove Propositions 1 and 2, we first give the following Lemma.

Lemma 5 ***Under Assumptions 1-5, 7, 8, and uniformly over*** $K(C_k)$ ***and for each*** $i=1,...,N$***, as*** $(N,T)\to\infty$***,***

(i) $||\frac{1}{T}\bar{V}'(k_1^0,k_2^0)M_{\underline{\overline{X}}(k_1^0,k_2^0)}\bar{V}(k_1^0,k_2^0)|| = O_p(N^{-1})$, $||\frac{1}{T}\underline{V}'_i(k_1^0,k_2^0)M_{\underline{\overline{X}}(k_1^0,k_2^0)}\underline{V}_i(k_1^0,k_2^0)|| = O_p(1)$

(ii)$||\frac{1}{T}\mathbb{F}(k_1^0,k_2^0)'M_{\underline{\overline{X}}(k_1^0,k_2^0)}\mathbb{F}(k_1^0,k_2^0)|| = O_p\left(N^{-1}\right)$, $||\frac{1}{T}\underline{V}'_i(k_1^0,k_2^0)M_{\underline{\overline{X}}(k_1^0,k_2^0)}\mathbb{F}(k_1^0,k_2^0)|| = O_p(N^{-1/2})$;

(iii) $||\frac{1}{T}\bar{V}(k_1^0,k_2^0)'\varepsilon_i|| = O_p(\frac{1}{N}) + O_p(\frac{1}{\sqrt{NT}})$, $||\frac{1}{T}\bar{V}(k_1^0,k_2^0)'\mathbb{F}^0(k_1^0,k_2^0)|| = \frac{1}{\sqrt{N}}$;

(iv) $||\frac{1}{T}\mathbb{F}^{0\prime}(k_1^0,k_2^0)\varepsilon_i|| = O_p(1)$.

Lemma 6 ***Under the Assumptions 1-5, 7, 8 and*** $q\le p$,***as*** $(N,T)\to\infty$,

$$\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\underline{\overline{X}}(k_3^0)}\underline{X}_i(k_1^0,k_2^0) = \frac{1}{T}\underline{V}'_i(k_1^0,k_2^0)\underline{V}_i(k_1^0,k_2^0) + O_p(\frac{1}{N}) + O_p(\frac{1}{\sqrt{T}}).$$

Proof of Lemma 6. We consider the case of two common breaks in the slopes and one in the error factor loadings, i.e., $m_0=2$, $m_1=1$. **In matrix form,**

$$\begin{aligned} Y_i &= \begin{pmatrix} x'_{i1}\beta_{i1} \\ \vdots \\ x'_{iT}\beta_{i,m_0+1}\end{pmatrix} + \begin{pmatrix} f'_1\gamma_{i1} \\ \vdots \\ f'_T\gamma_{i,m_1+1}\end{pmatrix} + \varepsilon_i \\ &= \underline{X}_i(k_1^0,k_2^0)b_i + \underset{T\times[(m_1+1)p]}{\mathbb{F}(k_3^0)}\ g_i + \varepsilon_i, \end{aligned}$$

where $\underline{X}_i(k_1^0,k_2^0) = diag\left(X_{i1},X_{i2},X_{i,3}\right)$, $\underset{T\times[(m_1+1)p]}{\mathbb{F}(k_3^0)} = diag((f_1,\cdots,f_{k_3^0})',(f_{k_3^0+1},\cdots,f_T)')$ **and** $g_i = (\gamma'_{i1},\gamma'_{i2})'$.

We use $\underset{T\times[(m_1+1)p]}{\underline{\overline{X}}(k_3^0)} = diag\left((\bar{x}'_1,...,\bar{x}'_{k_3^0})',(\bar{x}'_{k_3^0+1},...,\bar{x}'_T)'\right)$ **to proxy** $\mathbb{F}(k_3^0)$,

$$\begin{aligned}\hat{\mathbb{F}}(k_3^0) = \underset{T\times[(m_1+1)p]}{\underline{\overline{X}}(k_3^0)} &= \underset{T\times[(m_1+1)q]}{\mathbb{F}(k_3^0)}\ \underset{[(m_1+1)q]\times[(m_1+1)p]}{\bar{\mathbf{\Gamma}}} + \underset{T\times[(m_1+1)p]}{\bar{V}(k_3^0)} \\ &= \mathbb{F}(k_3^0)(I_2\otimes\bar{\Gamma}) + \bar{V}(k_3^0),\end{aligned}$$

where $\underset{T\times[(m_1+1)q]}{\bar{V}(k_3^0)} = diag((\bar{v}_1', ..., \bar{v}_{k_3^0}')', (\bar{v}_{k_3^0+1}', ..., \bar{x}_T')')$, **and** $\underset{[(m_1+1)q]\times[(m_1+1)q]}{\bar{\mathbf{\Gamma}}} = diag\left(\bar{\Gamma}, \cdots, \bar{\Gamma}\right) = I_{m_1+1} \otimes \bar{\Gamma}$.

Denote $\mathbb{F}^0(k_3^0) = [\mathbb{F}(k_3^0), 0_{T\times[(m_1+1)(p-q)]}]$ **and the full rank matrix**

$$\begin{aligned}\underset{[(m_1+1)p]\times[(m_1+1)p]}{B} &= [B_{[(m_1+1)q]}, B_{-(m_1+1)q}] \\ &= \begin{bmatrix} \bar{\mathbf{\Gamma}}_{(m_1+1)q}^{-1} & -\bar{\mathbf{\Gamma}}_{(m_1+1)q}^{-1}\bar{\mathbf{\Gamma}}_{-(m_1+1)q} \\ 0_{(m_1+1)(p-q)\times q} & I_{(m_1+1)p-(m_1+1)q} \end{bmatrix}.\end{aligned}$$

Define$\bar{\mathbf{\Gamma}} = [\bar{\mathbf{\Gamma}}_{(m_1+1)q}, \bar{\mathbf{\Gamma}}_{-(m_1+1)q}]$ **and** $\bar{V}(\mathcal{K}_1) = [\bar{V}_{(m_1+1)q}(\mathcal{K}_1), \bar{V}_{-(m_1+1)q}(\mathcal{K}_1)]$, **similar to the definitions of** $\bar{C} = [\bar{C}_m, \bar{C}_{-m}]$ **and** $\bar{U} = [\bar{U}_m, \bar{U}_{-m}]$ **in P.62 of Karabiyik et al. (2017). Thus,**

$$\begin{aligned}\hat{\mathbb{F}}(\mathcal{K}_1)B = \overline{X}(\mathcal{K}_1)B &= \mathbb{F}(\mathcal{K}_1)\bar{\mathbf{\Gamma}} + \bar{V}(\mathcal{K}_1)B \\ &= \mathbb{F}^0(\mathcal{K}_1) + [\bar{V}_{(m_1+1)q}(\mathcal{K}_1)\bar{\mathbf{\Gamma}}_{(m_1+1)q}^{-1}, \bar{V}_{-(m_1+1)q}(\mathcal{K}_1) \\ &\quad - \bar{V}_{(m_1+1)q}(\mathcal{K}_1)\bar{\mathbf{\Gamma}}_{(m_1+1)q}^{-1}\bar{\mathbf{\Gamma}}_{-(m_1+1)q}]\end{aligned}$$

and$\hat{\mathbb{F}}^0(\mathcal{K}_1) = \mathbb{F}^0(\mathcal{K}_1) + \bar{V}^0(\mathcal{K}_1)$ **with** $\bar{V}^0(\mathcal{K}_1) = \bar{V}(\mathcal{K}_1)BD_N = [\bar{V}_{(m_1+1)q}^0(\mathcal{K}_1), \bar{V}_{-(m_1+1)q}^0(\mathcal{K}_1)]$. **Since** BD_N **is positive definite,** $M_{\hat{\mathbb{F}}(\mathcal{K}_1)} = M_{\hat{\mathbb{F}}^0(\mathcal{K}_1)}$.

Define the pseudo-inverse $\bar{\mathbf{\Gamma}}^+ = \bar{\mathbf{\Gamma}}'(\bar{\mathbf{\Gamma}}\bar{\mathbf{\Gamma}}')^{-1}$ **such that** $\bar{\mathbf{\Gamma}}\bar{\mathbf{\Gamma}}^+ = I_{(m_1+1)q}$. **Following equation (S20) of Karabiyik et al. (2017), we obtain**

$$\begin{aligned}\underline{X}_i(\mathcal{K}_0) &= \mathbb{F}(\mathcal{K}_0)\mathbf{\Gamma}_i + V_i = \mathbb{F}(\mathcal{K}_0)\bar{\mathbf{\Gamma}}\bar{\mathbf{\Gamma}}^+\mathbf{\Gamma}_i + \underline{V}_i(\mathcal{K}_0) \\ &= \hat{\mathbb{F}}(\mathcal{K}_0)\bar{\mathbf{\Gamma}}^+\mathbf{\Gamma}_i - (\hat{\mathbb{F}}(\mathcal{K}_0) - \mathbb{F}(\mathcal{K}_0)\bar{\mathbf{\Gamma}})\bar{\mathbf{\Gamma}}^+\mathbf{\Gamma}_i + \underline{V}_i(\mathcal{K}_0) \\ &= \hat{F}^0(\mathcal{K}_0)D_N^{-1}B^{-1}\bar{\mathbf{\Gamma}}^+\mathbf{\Gamma}_i - \bar{V}(\mathcal{K}_0)\bar{\mathbf{\Gamma}}^+\mathbf{\Gamma}_i + \underline{V}_i(\mathcal{K}_0), \end{aligned} \tag{36}$$

and

$$\begin{aligned}\frac{1}{T}\underline{X}_i'(\mathcal{K}_0)M_{\overline{X}(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0) &= \frac{1}{T}\underline{X}_i'(\mathcal{K}_0)M_{\hat{\mathbb{F}}^0(\mathcal{K}_1)}\underline{X}_i(\mathcal{K}_0) \\ &= \frac{1}{T}\underline{X}_i'(\mathcal{K}_0)M_{\hat{\mathbb{F}}^0(\mathcal{K}_0)}\underline{X}_i(\mathcal{K}_0) \\ &\quad + \frac{1}{T}\underline{X}_i'(\mathcal{K}_0)\left[M_{\hat{\mathbb{F}}^0(\mathcal{K}_1)} - M_{\hat{\mathbb{F}}^0(\mathcal{K}_0)}\right]\underline{X}_i(\mathcal{K}_0)\end{aligned} \tag{37}$$

By following the proof of Lemma S.2 in Karabiyik et al. (2017), we show that the first term above is as follows:

$$\frac{1}{T}\underline{X}_i'(k_1^0, k_2^0)M_{\hat{\mathbb{F}}^0(k_3^0)}\underline{X}_i(k_1^0, k_2^0) = \frac{1}{T}\underline{V}_i'(k_1^0, k_2^0)\underline{V}_i(k_1^0, k_2^0) + O_p(\frac{1}{N}) + O_p(\frac{1}{\sqrt{T}}).$$

Similarly, since

$$(M_{\hat{\mathbb{F}}^0(k_3^0)} - M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)}) = \frac{1}{T^2}\mathbb{F}^0(k_1^0,k_2^0)\left(\frac{1}{T^2}\mathbb{F}^{0\prime}(k_1^0,k_2^0)\mathbb{F}^0(k_1^0,k_2^0)\right)^+\mathbb{F}^{0\prime}(k_1^0,k_2^0)$$
$$-\frac{1}{T^2}\mathbb{F}^0(k_3^0)\left(\frac{1}{T^2}\mathbb{F}^{0\prime}(k_3^0)\mathbb{F}^0(k_3^0)\right)^+\mathbb{F}^{0\prime}(k_3^0) + o_p(1),$$

there exists at least $[T\times\min\{\lambda_1^0,\lambda_2^0,...,\lambda_m^0\}]^2$ **elements equal to 0. Thus, given equation (S62) of Karabiyik et al. (2017),** $||M_{\hat{\mathbb{F}}^0(k_3^0)} - M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)}|| = O_p(\frac{1}{T})$,**and the second term in equation (37) above shrinks to 0 as** $T\to\infty$. **Combining these terms together, we show that**

$$\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\underline{X}_i(k_1^0,k_2^0) = \frac{1}{T}\underline{V}_i'(k_1^0,k_2^0)\underline{V}_i(k_1^0,k_2^0) + O_p(\frac{1}{N}) + O_p(\frac{1}{\sqrt{T}}).$$

Proof of Proposition 1. Following **equation (S17) of Karabiyik et al. (2017), we obtain** $\hat{\mathbb{F}}(k_3^0) = \underline{X}(k_3^0) = \mathbb{F}(k_3^0)\bar{\mathbf{\Gamma}} + \bar{V}(k_3^0)$. **Thus,**

$$\mathbb{F}(k_3^0) = \underline{X}(k_3^0)\bar{\mathbf{\Gamma}}^+ - \bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^+. \tag{38}$$

For the individual series $i = 1,...,N$, **plugging equation (38) into (14) gives,**

$$\begin{aligned} Y_i &= \underline{X}_i(k_1^0,k_2^0)b_i + \mathbb{F}(k_3^0)g_i + \varepsilon_i \\ &= \underline{X}_i(k_1^0,k_2^0)b_i + \overline{X}(k_3^0)\bar{\mathbf{\Gamma}}^+g_i - \bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^+g_i + \varepsilon_i \\ &= \underline{X}_i(\hat{k}_1,\hat{k}_2)b_i + [\underline{X}_i(k_1^0,k_2^0) - \underline{X}_i(\hat{k}_1,\hat{k}_2)]b_i \\ &\quad + \overline{X}(\hat{k}_3)\bar{\mathbf{\Gamma}}^+g_i + [\overline{X}(k_3^0) - \overline{X}(\hat{k}_3)]\bar{\mathbf{\Gamma}}^+g_i \\ &\quad + \varepsilon_i - \bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^+g_i. \end{aligned} \tag{39}$$

Plugging equation (39) above into the expression of $\hat{b}_i$ gives,

$$\begin{aligned} \hat{b}_i &= \hat{b}_i(\hat{k}_1,\hat{k}_2) = [\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}Y_i \qquad (40) \\ &= b_i + [\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1} \\ &\quad \times\Big\{\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}[\underline{X}_i(k_1^0,k_2^0) - \underline{X}_i(\hat{k}_1,\hat{k}_2)]b_i \\ &\quad + \underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}[\overline{X}(k_3^0) - \overline{X}(\hat{k}_3)]g_i + \underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\varepsilon_i \\ &\quad - \underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^+g_i\Big\}. \end{aligned}$$

Thus, we decompose $\sqrt{T}(\hat{b}_i - b_i)$ into six terms,

$$\begin{aligned}\sqrt{T}(\hat{b}_i - b_i) = [\frac{1}{T}\underline{X}_i(\hat{k}_1, \hat{k}_2)' M_{\overline{X}(\hat{k}_3)} \underline{X}_i(\hat{k}_1, \hat{k}_2)]^{-1} \Bigg\{ & \frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)' M_{\overline{X}(\hat{k}_3)} \varepsilon_i \\ & - \frac{1}{\sqrt{T}}[\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] M_{\overline{X}(\hat{k}_3)} \varepsilon_i \\ & \frac{1}{\sqrt{T}}\underline{X}_i\left(k_1^0, k_2^0\right)' M_{\overline{X}(\hat{k}_3)} \bar{V}(k_3^0) \bar{\mathbf{\Gamma}}^+ g_i \\ & + \frac{1}{\sqrt{T}}[\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] M_{\overline{X}(\hat{k}_3)} \bar{V}(k_3^0) \bar{\mathbf{\Gamma}}^+ g_i \\ & + \frac{1}{\sqrt{T}}\underline{X}_i(\hat{k}_1, \hat{k}_2)' M_{\overline{X}(\hat{k}_3)} [\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] b_i \\ & + \frac{1}{\sqrt{T}}\underline{X}_i(\hat{k}_1, \hat{k}_2)' M_{\overline{X}(\hat{k}_3)} [\overline{X}(k_3^0) - \overline{X}(\hat{k}_3)] g_i \Bigg\}. \end{aligned} \quad (41)$$

Under Theorem 1, $\hat{k}_1 - k_1^0 = o_p(1)$, $\hat{k}_2 - k_2^0 = o_p(1)$, **and** $\hat{k}_3 - k_3^0 = o_p(1)$, **for each** i, $\underline{X}_i(\hat{k}_1, \hat{k}_2) \xrightarrow{p} \underline{X}_i(k_1^0, k_2^0)$ **and** $M_{\overline{X}(\hat{k}_3)} \xrightarrow{p} M_{\overline{X}(k_3^0)}$. **Thus, under Assumption 6 (ii),**

$$\frac{1}{T}\underline{X}_i(\hat{k}_1, \hat{k}_2)' M_{\overline{X}(\hat{k}_3)} \underline{X}_i(\hat{k}_1, \hat{k}_2) - \frac{1}{T}\underline{X}_i(k_1^0, k_2^0)' M_{\overline{X}(k_3^0)} \underline{X}_i(k_1^0, k_2^0) \xrightarrow{p} 0.$$

As in KPY, in the model considered in Case 1, after the transformation using $M_{\overline{X}(\hat{k}_3)}$, $M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1, \hat{k}_2)$ **becomes stationary since I(1)** f_t **is removed asymptotically in regressors** x_{it}**, as shown in Lemma 6, that is,**

$$\begin{aligned}\Sigma_{X,i} &= plim_{T\to\infty}\frac{1}{T}\underline{X}_i(k_1^0, k_2^0)' M_{\overline{X}(k_3^0)} \underline{X}_i(k_1^0, k_2^0) \\ &= plim_{T\to\infty}\frac{1}{T}\underline{V}_i'(k_1^0, k_2^0)\underline{V}_i(k_1^0, k_2^0).\end{aligned}$$

For the second term of equation (41) inside the curly braces, by Theorem 1, $P\left(\hat{k}_1 \neq k_1^0, \hat{k}_2 \neq k_2^0\right) = P\left(|\hat{k}_1 \neq k_1^0| \geq 1, |\hat{k}_2 \neq k_2^0| \geq 1\right) \to 0$. For any $\eta > 0$**,**

$$\begin{aligned} & P(||T^{-1/2}[\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] M_{\overline{X}(\hat{k}_3)} \varepsilon_i|| > \eta) \\ = \;& P(||T^{-1/2}[\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] M_{\overline{X}(\hat{k}_3)} \varepsilon_i|| > \eta, \hat{k}_1 = k_1^0, \hat{k}_2 = k_2^0) \\ & + P(||T^{-1/2}[\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] M_{\overline{X}(\hat{k}_3)} \varepsilon_i|| > \eta, \hat{k}_1 \neq k_1^0, \hat{k}_2 \neq k_2^0) \\ = \;& P(0 > \eta) P(\hat{k}_1 = k_1^0, \hat{k}_2 = k_2^0) \\ & + P(||T^{-1/2}[\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] M_{\overline{X}(\hat{k}_3)} \varepsilon_i|| > \eta \Big| \hat{k}_1 \neq k_1^0, \hat{k}_2 \neq k_2^0) P(\hat{k}_1 \neq k_1^0, \hat{k}_2 \neq k_2^0) \\ \leq \;& P(0 > \eta)(\hat{k}_1 = k_1^0, \hat{k}_2 = k_2^0) + P(\hat{k}_1 \neq k_1^0, \hat{k}_2 \neq k_2^0) \to 0. \end{aligned}$$

Thus, $T^{-1/2}[\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] M_{\overline{X}(\hat{k}_3)} \varepsilon_i = o_p(1)$. **Similar arguments show that** , $T^{-1/2}\underline{X}_i(\hat{k}_1, \hat{k}_2)' M_{\overline{X}(\hat{k}_3)} [\underline{X}_i(k_1^0, k_2^0) - \underline{X}_i(\hat{k}_1, \hat{k}_2)] b_i = o_p(1)$ and $T^{-1/2}[\underline{X}_i(k_1^0, k_2^0) -$

$\underline{X}_i(\hat{k}_1, \hat{k}_2)]M_{\underline{X}(\hat{k}_3)}\bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^+ g_i = o_p(1)$, **and**

$$\frac{1}{\sqrt{T}}\underline{X}_i(\hat{k}_1, \hat{k}_2)' M_{\underline{X}(\hat{k}_3)}[\underline{X}(k_3^0) - \underline{X}(\hat{k}_3)]g_i = o_p(1).$$

Thus, the expression $\sqrt{T}(\hat{b}_i - b_i)$ **of equation (41) reduces to**

$$\sqrt{T}(\hat{b}_i - b_i) = [\frac{1}{T}\underline{X}_i(k_1^0, k_2^0)' M_{\underline{X}(k_3^0)}\underline{X}_i(k_1^0, k_2^0)]^{-1}[\frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)' M_{\underline{X}(k_3^0)}\varepsilon_i \\ - \frac{1}{\sqrt{T}}\underline{X}_i\left(k_1^0, k_2^0\right)' M_{\underline{X}(k_3^0)}\bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^+ g_i] + o_p(1).$$

Next, we need to consider the asymptotic distribution of

$$\frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)' M_{\underline{X}(k_3^0)}\varepsilon_i - \frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)' M_{\underline{X}(k_3^0)}\bar{V}(\hat{k}_3)\bar{\mathbf{\Gamma}}^+ g_i.$$

Our proof proceeds by following that of Theorem 3 of Karabiyik et al. (2017). First, consider

$$\frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)' M_{\underline{X}(k_3^0)}\varepsilon_i = Q_{NT} + \frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)'[M_{\underline{X}(k_3^0)} - M_{\underline{X}(k_1^0, k_2^0)}]\varepsilon_i \tag{42}$$

where

$$\begin{aligned} Q_T &= \frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)' M_{\mathbb{F}^0(k_1^0, k_2^0)}\varepsilon_i \\ &= \frac{1}{\sqrt{T}}[\underline{V}_i(k_1^0, k_2^0)' - \Gamma_i'(\bar{\mathbf{\Gamma}}')^+\bar{V}'(k_1^0, k_2^0)]\varepsilon_i \\ &- \frac{1}{\sqrt{T}}[\underline{V}_i(k_1^0, k_2^0)' - \Gamma_i'(\bar{\mathbf{\Gamma}}')^+\bar{V}'(k_1^0, k_2^0)]P_{\mathbb{F}^0(k_1^0, k_2^0)}\varepsilon_i \\ &+ \frac{1}{\sqrt{T}}[\underline{V}_i(k_1^0, k_2^0)' - \mathbf{\Gamma}_i'(\bar{\mathbf{\Gamma}}')^+\bar{V}'(k_1^0, k_2^0)][M_{\hat{\mathbb{F}}^0(k_1^0, k_2^0)} - M_{\mathbb{F}^0(k_1^0, k_2^0)}]\varepsilon_i \\ &= Q_{0T} - Q_{1T} + Q_{2T}. \end{aligned}$$

Since

$$\begin{aligned} ||\frac{1}{\sqrt{T}}\underline{X}_i(k_1^0, k_2^0)'[M_{\underline{X}(k_3^0)} - M_{\underline{X}(k_1^0, k_2^0)}]\varepsilon_i|| &\leq \sqrt{T}||\frac{1}{T}\underline{X}_i(k_1^0, k_2^0)'\varepsilon_i|| \times ||M_{\underline{X}(k_3^0)} - M_{\underline{X}(k_1^0, k_2^0)}|| \\ &= \sqrt{T}O_P(1)O_p(\frac{1}{T}) = O_p(\frac{1}{\sqrt{T}}), \end{aligned}$$

we focus on $Q_T = Q_{0T} - Q_{1T} + Q_{2T}$. **By Lemma 5(iii), we have**

$$\begin{aligned} Q_{0T} &= \frac{1}{\sqrt{T}}\underline{V}_i(k_1^0, k_2^0)'\varepsilon_i + \frac{1}{\sqrt{T}}\mathbf{\Gamma}_i(\bar{\mathbf{\Gamma}}')^+\bar{V}(k_1^0, k_2^0)'\varepsilon_i \\ &= \frac{1}{\sqrt{T}}\underline{V}_i(k_1^0, k_2^0)'\varepsilon_i + O_p(\frac{\sqrt{T}}{N}) + O_p(\frac{1}{\sqrt{N}}). \end{aligned}$$

Similarly, by Lemma5(iii) and (iv),

$$\begin{aligned}||Q_{1T}|| &= \frac{1}{\sqrt{T}}||[\underline{V}_i(k_1^0,k_2^0)' - \mathbf{\Gamma}_i(\bar{\mathbf{\Gamma}}^+)'\bar{V}(k_1^0,k_2^0)']P_{\mathbb{F}^0(k_1^0,k_2^0)}\varepsilon_i|| \\ &\leq \frac{1}{\sqrt{T}}||\frac{1}{T}\underline{V}_i(k_1^0,k_2^0)'\mathbb{F}^0(k_1^0,k_2^0)][\frac{1}{T^2}\mathbb{F}^{0\prime}(k_1^0,k_2^0)\mathbb{F}^0(k_1^0,k_2^0)]^{-1}[\frac{1}{T}\mathbb{F}^{0\prime}(k_1^0,k_2^0)\varepsilon_i]|| \\ &+ \frac{1}{\sqrt{T}}||\mathbf{\Gamma}_i(\bar{\mathbf{\Gamma}}^+)'\frac{1}{T}\bar{V}(k_1^0,k_2^0)'\mathbb{F}^0(k_1^0,k_2^0)][\frac{1}{T^2}\mathbb{F}^{0\prime}(k_1^0,k_2^0)\mathbb{F}^0(k_1^0,k_2^0)]^{-1}[\frac{1}{T}\mathbb{F}^{0\prime}(k_1^0,k_2^0)\varepsilon_i]|| \\ &= O_p(\frac{1}{\sqrt{T}}) + O_p(\frac{1}{\sqrt{TN}}) = O_p(\frac{1}{\sqrt{T}}).\end{aligned}$$

For the term Q_{2T}, according to equation (S29) of Karabiyik et al. (2017), we first obtain that

$$\begin{aligned}&T^2(M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)} - M_{\mathbb{F}^0(k_1^0,k_2^0)}) \\ =&T\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)[\frac{1}{T}\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)'\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)]^+\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)' \\ &+\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0)(\frac{1}{T^2}\mathbb{F}(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0))^+\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0)' \\ &+\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0)(\frac{1}{T^2}\mathbb{F}(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0))^+\mathbb{F}(k_1^0,k_2^0)' \\ &+\mathbb{F}(k_1^0,k_2^0)'(\frac{1}{T^2}\mathbb{F}(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0))^+\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0) \\ &+\hat{\mathbb{F}}^0(k_1^0,k_2^0)[(\frac{1}{T^2}\hat{\mathbb{F}}^0(k_1^0,k_2^0)'\hat{\mathbb{F}}^0(k_1^0,k_2^0))^+ - \Sigma^+_{\mathbb{F}^0}]\hat{\mathbb{F}}^0(k_1^0,k_2^0)',\end{aligned} \tag{43}$$

where $\Sigma_{\mathbb{F}^0} = \begin{bmatrix} \frac{1}{T^2}\mathbb{F}(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0) & \mathbf{0}_{[(m_1+1)q]\times[(m_1+1)(q-p)]} \\ \mathbf{0}_{[(m_1+1)(q-p)]\times[(m_1+1)q]} & \frac{1}{T}\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)'\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0) \end{bmatrix}$.

We plug the above expression (43) into Q_{2T},

$$\begin{aligned}Q_{2T} =& \frac{1}{\sqrt{T}}V_i(k_1^0,k_2^0)'(M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)} - M_{\mathbb{F}^0(k_1^0,k_2^0)})\varepsilon_i \\ &- \frac{1}{\sqrt{T}}\mathbf{\Gamma}_i(\bar{\mathbf{\Gamma}}^+)'\bar{V}(k_1^0,k_2^0)'(M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)} - M_{\mathbb{F}^0(k_1^0,k_2^0)})\varepsilon_i \\ =&Q_{2T,1} - Q_{2T,2}.\end{aligned}$$

First,

$$
\begin{aligned}
Q_{2T,1} =& \frac{1}{T^2\sqrt{T}} V_i(k_1^0,k_2^0)' T^2 (M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)} - M_{\mathbb{F}^0(k_1^0,k_2^0)})\varepsilon_i \\
=& [\frac{1}{T} V_i(k_1^0,k_2^0)' \bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)][\frac{1}{T}\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)'\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)]^+ \\
& \times [\frac{1}{\sqrt{T}}\bar{V}^0_{-(m_1+1)q}(k_1^0,k_2^0)'\varepsilon_i] \\
+& \frac{1}{\sqrt{T}}[\frac{1}{T}V_i(k_1^0,k_2^0)'\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0)](\frac{1}{T^2}\mathbb{F}(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0))^+[\frac{1}{T}\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0)'\varepsilon_i] \\
+& \frac{1}{\sqrt{T}}[\frac{1}{T}V_i(k_1^0,k_2^0)'\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0)](\frac{1}{T^2}\mathbb{F}(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0))^+[\frac{1}{T}\mathbb{F}(k_1^0,k_2^0)'\varepsilon_i] \\
+& \frac{1}{\sqrt{T}}[\frac{1}{T}V_i(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0)'](\frac{1}{T^2}\mathbb{F}(k_1^0,k_2^0)'\mathbb{F}(k_1^0,k_2^0))^+[\frac{1}{T}\bar{V}^0_{(m_1+1)q}(k_1^0,k_2^0)\varepsilon_i] \\
+& \frac{1}{\sqrt{T}}[\frac{1}{T}V_i(k_1^0,k_2^0)'\hat{\mathbb{F}}^0(k_1^0,k_2^0)][(\frac{1}{T^2}\hat{\mathbb{F}}^0(k_1^0,k_2^0)'\hat{\mathbb{F}}^0(k_1^0,k_2^0))^+ - \Sigma^+_{\mathbb{F}^0}][\frac{1}{T}\hat{\mathbb{F}}^0(k_1^0,k_2^0)'\varepsilon_i].
\end{aligned}
$$

According to Lemma 5 (iii) and (iv), $||Q_{2T,1}|| = O_p(\frac{1}{\sqrt{T}})$. **Similar argument show that** $||Q_{2T,2}|| = O_p(\frac{\sqrt{T}}{N}) + O_p(\frac{1}{\sqrt{N}})$. **Thus,** $||Q_{2T}|| = O_p(\frac{\sqrt{T}}{N}) + O_p(\frac{1}{\sqrt{N}}) + O_p(\frac{1}{\sqrt{T}})$.

According to Lemma 5(ii),$\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\bar{V}(k_3^0) = O_p(N^{-1})$, $T^{-1/2}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\bar{V}(k_3^0) = O_p\left(T^{1/2}N^{-1}\right)$, thus, the third term

$$||T^{-1/2}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^+ g_i|| = O_p(T^{1/2}N^{-1}).$$

Combining all these terms together, we obtain

$$\sqrt{T}(\hat{b}_i - b_i) = [\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\underline{X}_i(k_1^0,k_2^0)]^{-1}\frac{1}{\sqrt{T}}\underline{V}_i(k_1^0,k_2^0)'\varepsilon_i + O_p\left(T^{1/2}N^{-1}\right) + o_p(1).$$

According to Lemma 6,

$$
\begin{aligned}
\Sigma_{X,i} &= plim_{T\to\infty}\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\underline{X}_i(k_1^0,k_2^0) \\
&= plim_{T\to\infty}\frac{1}{T}\underline{V}_i(k_1^0,k_2^0)'\underline{V}_i(k_1^0,k_2^0)
\end{aligned}
$$

and then

$$\Sigma_{X\varepsilon,i} = plim_{T\to\infty}\frac{1}{T}\underline{V}_i(k_1^0,k_2^0)'\Sigma_{\varepsilon,i}\underline{V}_i(k_1^0,k_2^0)',$$

as $T^{1/2}N^{-1} \to 0$, we obtain $\sqrt{T}(\hat{b}_i - b_i) \xrightarrow{d} N(0, \Sigma^{-1}_{X,i}\Sigma_{X\varepsilon,i}\Sigma^{-1}_{X,i})$.

Proof of Proposition 2. Under Assumption 4, the asymptotic distribution of mean-group estimator can be derived similarly. Thus, we obtain

$$\sqrt{N}\left(\hat{b}_{MG}-b\right) = N^{-1/2}\textstyle\sum_{i=1}^{N} v_{b,i}$$
$$+\frac{1}{\sqrt{N}}\textstyle\sum_{i=1}^{N}[\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(\hat{k}_3)}\varepsilon_i$$
$$+\frac{1}{\sqrt{N}}\textstyle\sum_{i=1}^{N}[\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}[\underline{X}_i(\hat{k}_1,\hat{k}_2)-\underline{X}_i(k_1^0,k_2^0)]'M_{\overline{X}(\hat{k}_3)}\varepsilon_i$$
$$+\frac{1}{\sqrt{N}}\textstyle\sum_{i=1}^{N}[\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}[\underline{X}_i(\hat{k}_1,\hat{k}_2)-\underline{X}_i(k_1^0,k_2^0)]b_i$$
$$+\frac{1}{\sqrt{N}}\textstyle\sum_{i=1}^{N}[\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}[\overline{X}(k_3^0)-\overline{X}(\hat{k}_3)]g_i$$
$$+\frac{1}{\sqrt{N}}\textstyle\sum_{i=1}^{N}[\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^{+}g_i.$$

By Assumption 4, the limiting distribution of the first term is $N(0,\Sigma_b)$. **For the second term, equation (42) in the Proof of Proposition 1 implies that**

$$\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\varepsilon_i = \frac{1}{\sqrt{T}}(\frac{1}{\sqrt{T}}V_i(k_1^0,k_2^0)'\varepsilon_i + O_p(\frac{1}{\sqrt{N}})+O_p(\frac{1}{\sqrt{T}})) = O_p(\frac{1}{\sqrt{T}}),$$

as $(N,T)\to\infty$. **Thus,** $E[\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\varepsilon_i\varepsilon_i'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(k_1^0,k_2^0)] = O(\frac{1}{T})$, **and**

$$Var(N^{-1/2}\textstyle\sum_{i=1}^{N}[T^{-1}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}T^{-1}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(\hat{k}_3)}\varepsilon_i)$$
$$=\frac{1}{NT}\textstyle\sum_{i=1}^{N}(T^{-1}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2))^{-1}(T^{-1}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(\hat{k}_3)}Var(\varepsilon_i)M_{\overline{X}(\hat{k}_3)}\underline{X}_i(k_1^0,k_2^0))$$
$$\times(T^{-1}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2))^{-1} = O_p(T^{-1}).$$

Thus, $N^{-1/2}\sum_{i=1}^{N}[T^{-1}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\overline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}T^{-1}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(\hat{k}_3)}\varepsilon_i = O_p(T^{-1/2})$. Similarly, the last term is $O_p(N^{-1/2}T^{-1})$.

As in the proof of Proposition 1, the second, third and fourth terms are also $o_p(1)$. According to Lemma 5(ii) **and o the definition of** R_{NT} **in the equation (S21) of Karabiyik et al. (2017), we can follow the proof of Lemma S.1 of Karabiyik et al. (2017) to show that**

$$\frac{1}{T\sqrt{N}}\textstyle\sum_{i=1}^{N}\underline{X}_i(k_1^0,k_2)'M_{\overline{X}(\hat{k}_3)}\bar{V}(k_3^0) = O_p(\frac{1}{N})$$

without any restriction on the rate at which N and T tend to infinity.

Therefore, as $(N,T)\to\infty$**,**

$$\sqrt{N}(\hat{b}_{MG}-b) = N^{-1/2}\textstyle\sum_{i=1}^{N}v_{b,i} + o_p(1) \xrightarrow{d} N\left(0,\Sigma_b\right).$$

A.4 Proofs of Propositions 3, 4

Proof of Proposition 3. We will show that the convergence rate of $\hat{b}_i$ is T. From equation (34), $T(\hat{b}_i - b_i)$ can be decomposed into five terms,

$$\begin{aligned}T(\hat{b}_i - b_i) = & [T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\underline{\overline{X}}(\hat{k}_3)}\varepsilon_i \\ & -[T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}[\underline{X}_i(k_1^0,k_2^0) - \underline{X}_i(\hat{k}_1,\hat{k}_2)]M_{\underline{\overline{X}}(\hat{k}_3)}\varepsilon_i \\ & -[T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^{+}g_i \\ & +[T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}[\underline{X}_i(k_1^0,k_2^0) - \underline{X}_i(\hat{k}_1,\hat{k}_2)]'M_{\underline{\overline{X}}(\hat{k}_3)}\bar{V}(k_3^0)\bar{\mathbf{\Gamma}}^{+}g_i \\ & +[T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}[\underline{X}_i(k_1^0,k_2^0) - \underline{X}_i(\hat{k}_1,\hat{k}_2)]b_i.\end{aligned}$$

Under Theorem 2, $\hat{k}_1 - k_1^0 = o_p(1)$, $\hat{k}_2 - k_2^0 = o_p(1)$ and $\hat{k}_3 - k_3^0 = o_p(1)$, for each i, $\underline{X}_i\left(\hat{k}_1,\hat{k}_2\right) - \underline{X}_i\left(k_1^0,k_2^0\right) \xrightarrow{p} 0$ and $M_{\underline{\overline{X}}(\hat{k}_3)} \xrightarrow{p} M_{\underline{\overline{X}}(k_3^0)}$. Thus, similar to equation (37) in the proof of Proposition 1, except the first term, the other four terms above are $o_p(1)$, i.e.,

$$T(\hat{b}_i - b_i) = [T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)]^{-1}T^{-1}\underline{X}_i\left(\hat{k}_1,\hat{k}_2\right)'M_{\underline{\overline{X}}(\hat{k}_3)}\varepsilon_i + o_p(1).$$

Thus, to prove Proposition 3, we need to show that the first term above converges weakly to a non-degenerate distribution. Given that

$$T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{\overline{X}}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2) - T^{-2}\underline{X}_i(k_1^0,k_2^0)'M_{\underline{\overline{X}}(k_3^0)}\underline{X}_i(k_1^0,k_2^0) \xrightarrow{p} 0,$$

it is equivalent to show that $\left[\frac{1}{T^2}\underline{X}_i(k_1^0,k_2^0)'M_{\underline{\overline{X}}(k_3^0)}\underline{X}_i(k_1^0,k_2^0)\right]^{-1}\frac{1}{T}\underline{X}_i\left(k_1^0,k_2^0\right)'M_{\underline{\overline{X}}(k_3^0)}\varepsilon_i$ converges weakly to a non-degenerate distribution.

Following Phillips and Moon (1999), we will show that as $T \to \infty$,

$$\begin{aligned}\frac{1}{T^2}\underline{X}_i(k_1^0,k_2^0)'M_{\underline{\overline{X}}(k_3^0)}\underline{X}_i(k_1^0,k_2^0) & \Rightarrow G_i, \\ \frac{1}{T}\underline{X}_i\left(k_1^0,k_2^0\right)'M_{\underline{\overline{X}}(k_3^0)}\varepsilon_i & \Rightarrow H_i,\end{aligned}$$

where G_i and H_i are two non-degenerate distributions, respectively, which will be specified below. Therefore, as $T \to \infty$, $T(\hat{b}_i - b_i) \Rightarrow G_i^{-1}H_i$.

Consider the term $\frac{1}{T^2}\underline{X}_i(k_1^0,k_2^0)'M_{\underline{\overline{X}}(k_3^0)}\underline{X}_i(k_1^0,k_2^0)$ first. Denote $\underline{X}_i(k_1^0,k_2^0) = diag\left(X_{i1}, X_{i2}, X_{i3}\right)$ with $\underset{(k_1^0\times p)}{X_{i1}(k_1^0)} = (x_{i1},...,x_{i,k_1^0})'$, $\underset{(k_2^0-k_1^0)\times p}{X_{i2}(k_1^0,k_2^0)} = (x_{i,k_1^0+1},...,x_{ik_2^0})'$, $\underset{(T-k_2^0)\times p}{X_{i3}(k_2^0)} = (x_{i,k_2^0+1},...,x_{iT})'$. $F_1 = (f_1,...,f_{k_1^0})'$, $F_2 = (f_{k_1^0+1},...,f_{k_2^0})'$, and $F_3 = (f_{k_2^0+1},...,f_T)'$, and V_{i1}, V_{i2}, V_{i3}, ε_{1i}, ε_{2i}, ε_{3i} are similarly defined. Thus, $\underline{X}_i(k_1^0,k_2^0) = diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})$.

Let $F_4 = (f_1,...,f_{k_3^0})'$ and $F_5 = (f_{k_3^0+1},...,f_T)'$, we define $\mathbb{F}(k_3^0) = diag(F_4, F_5)$, and $\bar{\mathbb{V}}(k_3^0) = diag(\bar{V}_1, \bar{V}_2)$ with $\bar{V}_1 = (\bar{v}_1,...,\bar{v}_{k_3^0})'$ and $\bar{V}_2 = (\bar{v}_{k_3^0+1},...,\bar{v}_T)'$. When the rank condition is satisfied and $\underline{\overline{X}}(k_3^0) = \mathbb{F}(k_3^0)\bar{\Gamma} + \bar{\mathbb{V}}(k_3^0)$, $M_{\underline{\overline{X}}(k_3^0)}\underline{X}_i(k_1^0, k_2^0) = M_{\mathbb{F}(k_3^0)}\underline{X}_i(k_1^0, k_2^0) + o_p(1)$, as $(N, T) \to \infty$. Thus,

$$\begin{aligned}
&T^{-2}\underline{X}_i(k_1^0, k_2^0)' M_{\mathbb{F}(k_3^0)}\underline{X}_i(k_1^0, k_2^0) \\
&= T^{-2} diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})' \times diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3}) \\
&- [T^{-2} diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})'\mathbb{F}(k_3^0)](T^{-2}\mathbb{F}'(k_3^0)\mathbb{F}(k_3^0))^{-1} \\
&\times [T^{-2}\mathbb{F}'(k_3^0) diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})]. \qquad (44)
\end{aligned}$$

According to Phillips and Moon (1999, P.1062), under Assumption 12, for any $0 \leq \tau_1 \leq \tau_2 \leq 1$,

$$T^{-2}\sum_{[\tau_1 T]}^{[\tau_2 T]} v_{it}v_{it}' \Rightarrow \Psi_i(1)P_i(\int_{\tau_1}^{\tau_2} W_{\varsigma,i}W_{\varsigma,i}')P_i'\Psi_i(1)' = \int_{\tau_1}^{\tau_2} B_{\varsigma,i}B_{\varsigma,i}', \qquad (45)$$

where $B_{\varsigma,i}$ is a Brownian motion with covariance $\Psi_i(1)P_iP_i'\Psi_i(1)'$. Similarly, under Assumptions 5, 12 and 13,

$$T^{-2}\sum_{[\tau_1 T]}^{[\tau_2 T]} f_t f_t' \Rightarrow \Pi(1)Q(\int_{\tau_1}^{\tau_2} W_\varphi W_\varphi')Q'\Pi(1)' = \int_{\tau_1}^{\tau_2} B_\varphi B_\varphi', \qquad (46)$$

$$T^{-2}\sum_{[\tau_1 T]}^{[\tau_2 T]} f_t v_{it}' \Rightarrow \Pi(1)Q(\int_{\tau_1}^{\tau_2} W_\varphi W_{\varsigma,i}')P_i'\Psi_i(1)' = \int_{\tau_1}^{\tau_2} B_\varphi B_{\varsigma,i}'. \qquad (47)$$

In addition, under Assumptions 5, 8, and Lemma 8 of Phillips and Moon (1999),

$$\begin{aligned}
T^{-1}\sum_{[cT]}^{[dT]} f_t\varepsilon_{it} &\Rightarrow \Pi(1)Q(\int_c^d W_\varphi d(W_{\varepsilon.i}))\sigma_i + \textstyle\sum_{t=0}^{\infty}\sum_{s=0}^{\infty} E(\varphi_t\varepsilon_{i,t+s}) \\
&= \int_c^d B_\varphi d(B_{\varepsilon.i}) + \textstyle\sum_{t=0}^{\infty}\sum_{s=0}^{\infty} E(\varphi_t\varepsilon_{i,t+s}). \qquad (48)
\end{aligned}$$

Moreover, under Assumptions 5, 12 and 13,

$$\begin{aligned}
T^{-1}\sum_{[cT]}^{[dT]} v_{it}\varepsilon_{it} &\Rightarrow \Psi_i(1)P_i(\int_c^d W_{\varsigma,i} d(W_{\varepsilon.i}))\sigma_i + \textstyle\sum_{t=0}^{\infty}\sum_{s=0}^{\infty} E(\varsigma_{it}\varepsilon_{i,t+s}) \\
&= \int_c^d B_{\varsigma,i} d(B_{\varepsilon,i}) + \textstyle\sum_{t=0}^{\infty}\sum_{s=0}^{\infty} E(\varsigma_{it}\varepsilon_{i,t+s}). \qquad (49)
\end{aligned}$$

Consider the first term in equation (44) above,

$$\begin{aligned}
&T^{-2} diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})' \times diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3}) \\
=&T^{-2} diag\left((F_1\Gamma_i + V_{i1})'(F_1\Gamma_i + V_{i1}), (F_2\Gamma_i + V_{i2})'(F_2\Gamma_i + V_{i2}), (F_3\Gamma_i + V_{i3})'(F_3\Gamma_i + V_{i3})\right).
\end{aligned}$$

According to equations (45)-(47),

$$T^{-2}(F_j\Gamma_i + V_{j1})'(F_j\Gamma_i + V_{ij}) \Rightarrow \Gamma_i'\int_{\lambda_{j-1}^0}^{\lambda_j^0} B_\varphi B_\varphi'\Gamma_i + (\int_{\lambda_{j-1}^0}^{\lambda_j^0} B_{\varsigma,i}B_\varphi')\Gamma_i + \Gamma_i'(\int_{\lambda_{j-1}^0}^{\lambda_j^0} B_\varphi B_{\varsigma,i}') + \int_{\lambda_{j-1}^0}^{\lambda_j^0} B_{\varsigma,i}B_{\varsigma,i}',$$

for $j = \{1, 2, 3\}$ with $\lambda_0^0 = 0$ and $\lambda_3^0 = 1$. Thus,

$$
\begin{aligned}
&T^{-2}diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})' \cdot diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})\\
\Rightarrow & diag(\Gamma_i' \int_0^{\lambda_1^0} B_\varphi B_\varphi' \Gamma_i + (\int_0^{\lambda_1^0} B_{\varsigma,i} B_\varphi')\Gamma_i + \Gamma_i'(\int_0^{\lambda_1^0} B_\varphi B_{\varsigma,i}') + \int_0^{\lambda_1^0} B_{\varsigma,i} B_{\varsigma,i}',\\
&\Gamma_i' \int_{\lambda_1^0}^{\lambda_2^0} B_\varphi B_\varphi' \Gamma_i + (\int_{\lambda_1^0}^{\lambda_2^0} B_{\varsigma,i} B_\varphi')\Gamma_i + \Gamma_i'(\int_{\lambda_1^0}^{\lambda_2^0} B_\varphi B_{\varsigma,i}') + \int_{\lambda_1^0}^{\lambda_2^0} B_{\varsigma,i} B_{\varsigma,i}',\\
&\Gamma_i' \int_{\lambda_2^0}^{1} B_\varphi B_\varphi' \Gamma_i + (\int_{\lambda_2^0}^{1} B_{\varsigma,i} B_\varphi')\Gamma_i + \Gamma_i'(\int_{\lambda_2^0}^{1} B_\varphi B_{\varsigma,i}') + \int_{\lambda_2^0}^{1} B_{\varsigma,i} B_{\varsigma,i}') \equiv G_{i1}.
\end{aligned}
$$

Similarly, according to equations (46) and (47), the second term in equation (44)

$$
\begin{aligned}
&T^{-2}diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})'\mathbb{F}(k_3^0)\\
=&T^{-2}\left(F_1'\mathbb{F}(k_3^0)\Gamma_i + \mathbb{F}'(k_3^0)V_{i1}, F_2'\mathbb{F}(k_3^0)\Gamma_i + \mathbb{F}'(k_3^0)V_{i2}, F_3'\mathbb{F}(k_3^0)\Gamma_i + \mathbb{F}'(k_3^0)V_{i3}\right)'.
\end{aligned}
$$

Next, we derive the limiting distributions of the terms above. Without loss of generality, we assume that $k_1^0 < k_2^0 < k_3^0$ and define $F_\Delta = (f_{k_2^0+1}, \cdots, f_{k_3^0})'$. Since $\mathbb{F}(k_3^0) = diag(F_4, F_5)$,

$$
\begin{aligned}
&T^{-2}diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})'diag(F_4, F_5)\\
=&\frac{1}{T^2}\begin{pmatrix} \underset{p\times k_1^0}{\Gamma_i'F_1' + V_{i1}'} & & \\ & \underset{p\times(k_2^0-k_1^0)}{\Gamma_i'F_2' + V_{i2}'} & \\ & & \underset{p\times(T-k_2^0)}{\Gamma_i'F_3' + V_{i3}'} \end{pmatrix}\begin{pmatrix} \underset{k_3^0\times q}{F_4} & \\ & \underset{(T-k_3^0)\times q}{F_5} \end{pmatrix}\\
=&\begin{pmatrix} \frac{1}{T^2}\Gamma_i'F_1'F_1 + \frac{1}{T^2}V_{i1}'F_1 & 0_{p\times q}\\ \frac{1}{T^2}\Gamma_i'F_2'F_2 + \frac{1}{T^2}V_{i2}'F_2 & 0_{p\times q}\\ \frac{1}{T^2}\Gamma_i'F_\Delta'F_\Delta + \frac{1}{T^2}V_{i\Delta}'F_\Delta & \frac{1}{T^2}\Gamma_i'F_5'F_5 + \frac{1}{T^2}V_{i5}'F_5 \end{pmatrix},
\end{aligned}
$$

then

$$
\begin{aligned}
&T^{-2}diag(F_1\Gamma_i + V_{i1}, F_2\Gamma_i + V_{i2}, F_3\Gamma_i + V_{i3})'diag(F_4, F_5)\\
\Rightarrow&\begin{pmatrix} \Gamma_i'\int_0^{\lambda_1^0} B_\varphi B_\varphi' + \int_0^{\lambda_1^0} B_{\varsigma,i}B_\varphi' & 0_{p\times q}\\ \Gamma_i'\int_{\lambda_1^0}^{\lambda_2^0} B_\varphi B_\varphi' + \int_{\lambda_1^0}^{\lambda_2^0} B_{\varsigma,i}B_\varphi' & 0_{p\times q}\\ \Gamma_i'\int_{\lambda_2^0}^{\lambda_3^0} B_\varphi B_\varphi' + \int_{\lambda_2^0}^{\lambda_3^0} B_{\varsigma,i}B_\varphi' & \Gamma_i'\int_{\lambda_3^0}^{1} B_\varphi B_\varphi' + \int_{\lambda_3^0}^{1} B_{\varsigma,i}B_\varphi' \end{pmatrix} \equiv G_{i2}.
\end{aligned}
$$

In addition, $T^{-2}\mathbb{F}'(k_3^0)\mathbb{F}(k_3^0) = T^{-2}diag(F_4'F_4, F_5'F_5) \Rightarrow diag(\int_0^{\lambda_3^0} B_\varphi B_\varphi', \int_{\lambda_3^0}^1 B_\varphi B_\varphi')$. Thus, we obtain

$$
\begin{aligned}
&T^{-2}\underline{X}_i(k_1^0, k_2^0)'M_{\mathbb{F}(k_3^0)}\underline{X}_i(k_1^0, k_2^0) \qquad (50)\\
\Rightarrow&G_{i1} - G_{i2}diag\left((\int_0^{\lambda_3^0} B_\varphi B_\varphi')^{-1}, (\int_{\lambda_3^0}^1 B_\varphi B_\varphi')^{-1}\right)G_{i2}' \equiv G_i.
\end{aligned}
$$

Likewise, $\frac{1}{T}\mathbb{F}(k_3^0)'\varepsilon_i = \frac{1}{T}diag(F_4', F_5')\varepsilon_i \Rightarrow diag\left(\int_0^{k_3^0} B_\varphi d(B_{\varepsilon.i}), \int_{k_3^0}^1 B_\varphi d(B_{\varepsilon.i})\right)$ and then

$$\begin{aligned}
& T^{-1}\underline{X}_i\left(k_1^0,k_2^0\right)' M_{\underline{X}(k_3^0)}\varepsilon_i \\
\xrightarrow{p} \; & T^{-1}\underline{X}_i\left(k_1^0,k_2^0\right)' M_{\mathbb{F}(k_3^0)}\varepsilon_i \\
= \; & T^{-1}diag(F_1\Gamma_i+V_{i1}, F_2\Gamma_i+V_{i2}, F_3\Gamma_i+V_{i3})'\varepsilon_i \\
- \; & T^{-1}diag(F_1\Gamma_i+V_{i1}, F_2\Gamma_i+V_{i2}, F_3\Gamma_i+V_{i3})'\mathbb{F}(k_3^0)(\mathbb{F}(k_3^0)'\mathbb{F}(k_3^0))^{-1}\mathbb{F}(k_3^0)\varepsilon_i \\
= \; & \begin{bmatrix} T^{-1}\Gamma_i'F_1'\varepsilon_{1i}+T^{-1}V_{i1}'\varepsilon_{1i} \\ T^{-1}\Gamma_i'F_2'\varepsilon_{2i}+T^{-1}V_{i2}'\varepsilon_{2i} \\ T^{-1}\Gamma_i'F_3'\varepsilon_{3i}+T^{-1}V_{i3}'\varepsilon_{3i} \end{bmatrix} \\
& -T^{-2}diag(F_1\Gamma_i+V_{i1}, F_2\Gamma_i+V_{i2}, F_3\Gamma_i+V_{i3})'\mathbb{F}(k_3^0)\left(T^{-2}\mathbb{F}(k_3^0)'\mathbb{F}(k_3^0)\right)^{-1}\left(\frac{1}{T}\mathbb{F}(k_3^0)'\varepsilon_i\right).
\end{aligned}$$

According to equations (46), (48), and (49),

$$\begin{aligned}
& T^{-1}\underline{X}_i\left(k_1^0,k_2^0\right)' M_{\underline{X}(k_3^0)}\varepsilon_i \qquad (51)\\
\Rightarrow & \begin{bmatrix} \Gamma_i'\int_0^{\lambda_1^0} B_\varphi d(B_{\varepsilon.i}) + \Gamma_i'\sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varphi_t\varepsilon_{i,t+s}\right) + \int_0^{\lambda_1^0} B_{\varsigma,i}d(B_{\varepsilon.i}) + \sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varsigma_{it},\varepsilon_{i,t+s}\right) \\ \Gamma_i'\int_{\lambda_1^0}^{\lambda_2^0} B_\varphi d(B_{\varepsilon.i}) + \Gamma_i'\sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varphi_t\varepsilon_{i,t+s}\right) + \int_{\lambda_1^0}^{\lambda_2^0} B_{\varsigma,i}d(B_{\varepsilon.i}) + \sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varsigma_{it},\varepsilon_{i,t+s}\right) \\ \Gamma_i'\int_{\lambda_1^0}^{\lambda_2^0} B_\varphi d(B_{\varepsilon.i}) + \Gamma_i'\sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varphi_t\varepsilon_{i,t+s}\right) + \int_{\lambda_1^0}^{\lambda_2^0} B_{\varsigma,i}d(B_{\varepsilon.i}) + \sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varsigma_{it},\varepsilon_{i,t+s}\right) \end{bmatrix} \\
& - G_{i1}diag\left(\left(\int_0^{\lambda_3^0} B_\varphi B_\varphi'\right)^{-1}, \left(\int_{\lambda_3^0}^1 B_\varphi B_\varphi'\right)^{-1}\right) \\
& \times diag\left(\int_0^{k_3^0} B_\varphi d(B_{\varepsilon.i}) + \sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varphi_t\varepsilon_{i,t+s}\right), \int_{k_3^0}^1 B_\varphi d(B_{\varepsilon.i}) + \sum_{t=0}^\infty\sum_{s=0}^\infty E\left(\varphi_t\varepsilon_{i,t+s}\right)\right) \equiv H_i. \qquad (52)
\end{aligned}$$

Proof of Proposition 4. By the same argument in the proof of Proposition 2, we can obtain equation (20),

$$\sqrt{N}\left(\hat{b}_{MG}-b\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^N v_{b,i} + \frac{1}{\sqrt{N}T}\sum_{i=1}^N\left[\left(\frac{1}{T^2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)\right)^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{X}(\hat{k}_3)}\varepsilon_i\right]$$

In a special case of homogeneous slopes $b_i = b$ with $v_{b,i} = 0$, we have,

$$\sqrt{N}T\left(\hat{b}_{MG}-b\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^N\left[\left(\frac{1}{T^2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)\right)^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{X}(\hat{k}_3)}\varepsilon_i\right] + o_p(1).$$

As in the proof of Proposition 3 above, $\left(\frac{1}{T^2}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{X}(\hat{k}_3)}\underline{X}_i(\hat{k}_1,\hat{k}_2)\right)^{-1}\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{X}(\hat{k}_3)}\varepsilon_i$ weakly converges a non-degenerate distribution $G_i^{-1}H_i$.

Under the assumptions that $\varepsilon_{it}, \varphi_s, \varsigma_{jt'}$ are independent for all (i,j) and (t,s,t'), and $E(\varepsilon_{it}) = 0$, $E\left[\frac{1}{T}\underline{X}_i(\hat{k}_1,\hat{k}_2)'M_{\underline{X}(\hat{k}_3)}\varepsilon_i\right] = 0$. Thus, $\sqrt{N}T\left(\hat{b}_{MG}-b\right)$ is consistent, as $(N,T)\to\infty$. In addition, under the assumption cross-sectional independence of ε_{it}, $G_i^{-1}H_i$ are independent across i. Thus, by the Central Limit Theory, the limiting distribution of $\sqrt{N}T\left(\hat{b}_{MG}-b\right)$ is multivariate normal, i.e., as $(N,T)\to\infty$, $\sqrt{N}T\left(\hat{b}_{MG}-b\right)\xrightarrow{d}N(0,\Sigma_{MG})$. Next, we derive the expression of Σ_{MG}. For simplicity, asymptotic bias mentioned in Theorem 8 of Phillips and Moon (1999) and Proposition 1 of Bai, Ng and Kao (2009) disappears here under the assumptions of no serial/ cross-sectional correlation and heteroskedasticity.

Let $w_{it}=(\varepsilon_{it},\varphi_t',\varsigma_{it}')'$. Denote the long-run covariance matrix of w_{it}, partitioned comfortably for w_{it}, by

$$\Omega_i=\sum_{j=-\infty}^{\infty}E(w_{i0}w_{ij}')=\begin{bmatrix}\Omega_{\varepsilon.i} & \Omega_{\varepsilon\varphi i} & \Omega_{\varepsilon\varsigma i}\\ \Omega_{\varphi\varepsilon i} & \Omega_{\varphi} & \Omega_{\varphi\varsigma i}\\ \Omega_{\varsigma\varepsilon i} & \Omega_{\varsigma\varphi i} & \Omega_{\varsigma.i}\end{bmatrix}.$$

Denote $L_1\sim N(0,I_r)$ and $L_2\sim N(0,I_p)$. thus, as $T\to\infty$, $\frac{1}{T}F'\varepsilon_i\Rightarrow\int_0^1 B_\varphi d(B_{\varepsilon.i})\equiv\underset{r\times1}{\xi_{i1}}\sim\Omega_{\varepsilon.i}^{1/2}\Omega_{\varphi}^{1/2}\times L_1$, $\frac{1}{T}V_i'\varepsilon_i\Rightarrow\int_0^1 B_{\varsigma,i}d(B_{\varepsilon.i})\equiv\underset{p\times1}{\xi_{i2}}\sim\Omega_{\varepsilon.i}^{1/2}\Omega_{\varsigma.i}^{1/2}\times L_2$, where ξ_{i1} and ξ_{i2} are Gaussian processes, independent across i. Similarly, as $T\to\infty$,

$$\frac{1}{T^2}V_i'F_1\Rightarrow\int_0^1 B_\varphi B_{\varsigma,i}'\equiv\underset{p\times r}{\xi_{i3}},\ \frac{1}{T^2}F'F\Rightarrow\int_0^1 B_\varphi B_\varphi'\equiv\underset{r\times r}{\xi_4},\ \frac{1}{T^2}V_i'V_i\Rightarrow\int_0^1 B_{\varsigma,i}B_{\varsigma,i}'\equiv\underset{p\times p}{\xi_{i5}}$$

where ξ_{i3},ξ_{i4} and ξ_{i5} are Gaussian processes. The proof of Proposition 3 above shows, $\frac{1}{T^2}\underline{X}_i(k_1^0,k_2^0)'M_{\mathbb{F}(k_3^0)}\underline{X}_i(k_1^0,k_2^0)\Rightarrow G_i$. According to the definitions of $\xi_{i1},\xi_{i2},\xi_{i3},\xi_4$, and let $\lambda=\begin{bmatrix}\lambda_1^0 & 0\\ \lambda_2^0-\lambda_1^0 & 0\\ \lambda_3^0-\lambda_2^0 & 1-\lambda_3^0\end{bmatrix}$, we obtain

$$\begin{aligned}
G_i=&\,diag(\lambda_1^0,\lambda_2^0-\lambda_1^0,1-\lambda_2^0)\otimes(\Gamma_i'\xi_4\Gamma_i+\xi_{i3}\Gamma_i+\Gamma_i'\xi_{i3}'+\xi_{i5})\\
&-\begin{pmatrix}\Gamma_i'\lambda_1^0\xi_4+\lambda_1^0\xi_{i3} & 0_{p\times r}\\ \Gamma_i'(\lambda_2^0-\lambda_1^0)\xi_4+(\lambda_2^0-\lambda_1^0)\xi_{i3} & 0_{p\times r}\\ \Gamma_i'(\lambda_3^0-\lambda_2^0)\xi_4+(\lambda_3^0-\lambda_2^0)\xi_{i3} & \Gamma_i'(1-\lambda_3^0)\xi_4+(1-\lambda_3^0)\xi_{i3}\end{pmatrix}\begin{pmatrix}\lambda_3^0\xi_4 & \\ & (1-\lambda_3^0)\xi_4\end{pmatrix}^{-1}\\
&\times\begin{pmatrix}\Gamma_i'\lambda_1^0\xi_4+\lambda_1^0\xi_{i3} & 0_{p\times r}\\ \Gamma_i'(\lambda_2^0-\lambda_1^0)\xi_4+(\lambda_2^0-\lambda_1^0)\xi_{i3} & 0_{p\times r}\\ \Gamma_i'(\lambda_3^0-\lambda_2^0)\xi_4+(\lambda_3^0-\lambda_2^0)\xi_{i3} & \Gamma_i'(1-\lambda_3^0)\xi_4+(1-\lambda_3^0)\xi_{i3}\end{pmatrix}'\\
=&\,diag(\lambda_1^0,\lambda_2^0-\lambda_1^0,1-\lambda_2^0)\otimes(\Gamma_i'\xi_4\Gamma_i+\xi_{i3}\Gamma_i+\Gamma_i'\xi_{i3}'+\xi_{i5})\\
&-\left[\lambda\otimes(\Gamma_i'\xi_4+\xi_{i3})\right]diag((\lambda_3^0\xi_4)^{-1},((1-\lambda_3^0)\xi_4)^{-1})\left[\lambda\otimes(\Gamma_i'\xi_4+\xi_{i3})\right]'.
\end{aligned}$$

Similarly, since $\frac{1}{T}\underline{X}_i(k_1^0,k_2^0)'M_{\overline{X}(k_3^0)}\varepsilon_i\Rightarrow H_i$ and

$$H_i=\begin{bmatrix}\lambda_1^0\Gamma_i'\xi_{i1}+\lambda_1^0\xi_{i2}\\ (\lambda_2^0-\lambda_1^0)\Gamma_i'\xi_{i1}+(\lambda_2^0-\lambda_1^0)\xi_{i2}\\ (1-\lambda_2^0)\Gamma_i'\xi_{i1}+(1-\lambda_2^0)\xi_{i2}\end{bmatrix}-\left[\lambda\otimes(\Gamma_i'\xi_4+\xi_{i3})\right]diag((\lambda_3^0\xi_4)^{-1},((1-\lambda_3^0)\xi_4)^{-1})\xi_{i1}$$

Therefore,

$$\Sigma_{MG} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E[(T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)' M_{\underline{\overline{X}}(\hat{k}_3)} \underline{X}_i(\hat{k}_1,\hat{k}_2))^{-1} (T^{-1}\underline{X}_i(\hat{k}_1,\hat{k}_2)' M_{\underline{\overline{X}}(\hat{k}_3)} \varepsilon_i) \tag{53}$$

$$\times (T^{-1}\varepsilon_i' M_{\underline{\overline{X}}(\hat{k}_3)} \underline{X}_i(\hat{k}_1,\hat{k}_2)')(T^{-2}\underline{X}_i(\hat{k}_1,\hat{k}_2)' M_{\underline{\overline{X}}(\hat{k}_3)} \underline{X}_i(\hat{k}_1,\hat{k}_2))^{-1}]$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E(G_i^{-1} H_i H_i' G_i^{-1})$$

Following Phillips and Moon (1999, p. 1081), we can estimate consistently Σ_{MG} by plugging the residuals $\hat{\varepsilon}_i$ into equation (53) above,

$$\hat{\Sigma}_{MG} = \frac{1}{N} \sum_{i=1}^{N} \left[\left(\frac{1}{T^2} \underline{X}_i(\hat{k}_1,\hat{k}_2)' M_{\underline{\overline{X}}(\hat{k}_3)} \underline{X}_i(\hat{k}_1,\hat{k}_2) \right)^{-1} \left(\frac{1}{T} \underline{X}_i(\hat{k}_1,\hat{k}_2)' M_{\underline{\overline{X}}(\hat{k}_3)} \hat{\varepsilon}_i \right) \right.$$

$$\left. \times \left(\frac{1}{T} \hat{\varepsilon}_i' M_{\underline{\overline{X}}(\hat{k}_3)} \underline{X}_i(\hat{k}_1,\hat{k}_2) \right) \left(\frac{1}{T^2} \underline{X}_i(\hat{k}_1,\hat{k}_2)' M_{\underline{\overline{X}}(\hat{k}_3)} \underline{X}_i(\hat{k}_1,\hat{k}_2) \right)^{-1} \right],$$

where $\hat{\varepsilon}_i = Y_i - \underline{X}_i(\hat{k}_1,\hat{k}_2)\hat{b}_{MG}$. In this special case of $b_i = b$, we use the efficient estimator $\hat{b}_{MG}$ instead of $\hat{b}_i$. In addition, the term $\bar{X}\gamma_i^*(k_3^0)$ in equation (39) will be partialled out by $M_{\bar{X}}$ in the expression above.

Appendix B: Proofs of Lemmas

In this appendix, we provide detailed proofs of technical lemmas used in Appendix A. Lemma 1 is used to prove Lemma 2 and Theorem 1. Lemma 3 is used to prove Lemma 4 and Theorem 2.

We remind readers of the nontrivial notation that for the interval $t \in [k_1+1, k_2]$, $Y_{i\bigstar} = (0,...,0, y_{i,k_1+1},...,y_{i,k_2},0,...,0)'$, $Z_{i\bigstar} = (0,...,0, z_{i,k_1+1},...,z_{i,k_2},0,...,0)'$, $\varepsilon^*_{i\bigstar} = (0,...,0,\varepsilon^*_{i,k_1+1},...,\varepsilon^*_{i,k_2},0,...,0)'$ and $\bar{V}_{\bigstar} = (0,\cdots,0,\bar{v}_{k_1+1},...,\bar{v}_{k_2},0,\cdots,0)'$. In addition, let $Z_{i\Delta} = (0,...,0,z_{i,k_1+1},...,z_{i,k_2^0},0,...,0)'$, $Z_{0i\bigstar} = (0,...,0,z_{i,k_1^0+1},...,z_{i,k_2},0,...,0)'$. Similarly, for the interval $t \in [1, k_1^0]$, $Z_{i\lozenge} = (z_{i,1},...,z_{i,k_1^0},0,...,0)'$, $Y_{i\lozenge} = (y_{i,1},...,y_{i,k_1^0},0,...,0)'$ and $\varepsilon^*_{i\lozenge} = (\varepsilon^*_{i,1},...,\varepsilon^*_{i,k_1^0},0,...,0)'$. As in BFK (2016), here we assume that $\left|\hat{k}_1 - k_1^0\right|$, $\left|\hat{k}_2 - k_2^0\right|$ and $\left|\hat{k}_3 - k_3^0\right|$ are bounded for simplicity. Under Assumption 1 and that the estimators of break fractions are consistent, we only consider the set $K(C_k) = \left\{(k_1,k_2,k_3) : 1 \leq \left|k_j - k_j^0\right|, \left|k_j - k_j^0\right| \leq C_k, aT \leq k_j \leq (1-a)T, j = (1,2,3)\right\}$ for a finite constant C_k and $a > 0$.

Lemma 1. Under Assumptions 1-5, 7,8, and uniformly over $K(C_k)$, as $(N,T) \to \infty$, for $i = 1,...,N$,

(i) $\frac{1}{T} Z'_{i\Delta} Z_{i\Delta} = O_p(1)$, $\frac{1}{T^2} Z'_{i\bigstar} Z_{i\bigstar} = O_p(1)$;

(ii) $\frac{1}{\sqrt{T}} Z'_{i\Delta}\varepsilon_{i\bigstar} = \frac{1}{\sqrt{T}} Z'_{i\Delta}\varepsilon_{i\Delta} = O_p(1)$, $\frac{1}{T} Z'_{i\bigstar}\varepsilon_{i\bigstar} = O_p(1)$;

(iii) $\frac{1}{\sqrt{T}} Z'_{i\Delta}\varepsilon_{i\lozenge} = \frac{1}{\sqrt{T}} Z'_{i\Delta}\varepsilon_{i\Delta} = O_p(1)$, $\frac{1}{T} Z'_{i\lozenge}\varepsilon_{i\lozenge} = O_p(1)$;

(iv) $\frac{1}{T}\bar{V}'_{\bigstar}\bar{V}_{\bigstar} = O_p\left(\frac{1}{N}\right)$, $\frac{1}{\sqrt{T}} Z'_{i\Delta}\bar{V}_{\bigstar} = O_p\left(\frac{1}{\sqrt{N}}\right)$, $\frac{1}{T} Z'_{i\bigstar}\bar{V}_{\bigstar} = O_p\left(\frac{1}{\sqrt{N}}\right)$.

Proof of Lemma 1. (i) According to Lemma 1(b) in the supplementary appendix of Baltagi, Kao and Liu (2017), for any $0 \leq \tau_1 < \tau_2 \leq 1$ and under Assumption 5,

$$\frac{1}{T^2}\sum_{t=[\tau_1 T]}^{[\tau_2 T]} f_t f'_t = O_p(\tau_2 - \tau_1). \tag{54}$$

In addition, according to Lemma 1(a) in the supplementary appendix of Baltagi, Kao and Liu (2017), for any $0 \leq \tau_1 < \tau_2 \leq 1$ and under Assumption 9,

$$\frac{1}{T}\sum_{t=[\tau_1 T]}^{[\tau_2 T]} v_{it} v'_{it} = O_p(\tau_2 - \tau_1). \tag{55}$$

Plugging $x_{it} = \Gamma_i' f_t + v_{it}$ into $X_{i\Delta}$ gives

$$\begin{aligned}\frac{1}{T}X_{i\Delta}'X_{i\Delta} &= \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} x_{it}x_{it}' = \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \left(\Gamma_i' f_t + v_{it}\right)\left(\Gamma_i' f_t + v_{it}\right)' \\ &= \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \Gamma_i + \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it}v_{it}' \\ &+ \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t v_{it}' + \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f_t' \Gamma_i.\end{aligned}$$

Since $|k_1 - k_1^0|$ is bounded, $|\lambda_1 - \lambda_1^0| = O_p\left(\frac{1}{T}\right)$. Let $F_\Delta = \left(0 \cdots 0, f_{k_1+1}, \cdots, f_{k_1^0}, 0 \cdots 0\right)'$. According to equation (49) and $|\lambda_1 - \lambda_1^0| = O_p\left(\frac{1}{T}\right)$, $\frac{1}{T}F_\Delta' F_\Delta = O_p(1)$ under Assumption 1. Thus, $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \Gamma_i = O_p(1)$, under Assumption 3 that Γ_i is bounded. From equation (50) and $|\lambda_1 - \lambda_1^0| = O_p\left(\frac{1}{T}\right)$, $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it}v_{it}' = O_p\left(\frac{1}{T}\right)$. Let $V_{i\Delta} = \left(0 \cdots 0, v_{i,k_1+1}, \cdots, v_{i,k_1^0}, 0 \cdots 0\right)$, $Var\left(\frac{1}{T}F_\Delta' V_{i\Delta}\right) = O(1)$ under Assumption 9. Thus, $\frac{1}{\sqrt{T}}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t v_{it}' = O_p(1)$ and $\frac{1}{\sqrt{T}}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f_t' \Gamma_i = O_p(1)$. Thus, $\frac{1}{T}X_{i\Delta}'X_{i\Delta} = O_p(1) + O_p\left(\frac{1}{T}\right) = O_p(1)$.

Plugging $\bar{x}_t = \bar{\Gamma}' f_t + \bar{v}_t$ into $\bar{X}_\Delta$ gives,

$$\begin{aligned}\frac{1}{T}\bar{X}_\Delta' \bar{X}_\Delta &= \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{x}_t \bar{x}_t' = \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} (\bar{\Gamma}' f_t + \bar{v}_t)(\bar{\Gamma}' f_t + \bar{v}_t)' \\ &= \bar{\Gamma}'(\frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} f_t f_t')\bar{\Gamma} + \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{v}_t \bar{v}_t' \\ &+ \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{v}_t f_t' \Gamma_i + \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{\Gamma}' f_t \bar{v}_t'.\end{aligned}$$

According to equation (49), $\bar{\Gamma}'\left(\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} f_t f_t'\right)\bar{\Gamma} = O_p(1)$, under Assumption 3 that $\bar{\Gamma}$ is bounded and $|\lambda_1 - \lambda_1^0| = O_p\left(\frac{1}{T}\right)$. Under Assumption 9, we have $E\left\|\bar{v}_t\right\|^2 = \frac{1}{N^2}\sum_{i=1}^N E\left\|v_{it}\right\|^2 = O\left(\frac{1}{N}\right)$. Also, $E\left\|\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{v}_t \bar{v}_t'\right\| \leq \frac{1}{T}\sum_{t=k_1}^{k_1^0}\left(E\left\|\bar{v}_t\right\|^2\right) = O\left(\frac{1}{NT}\right)$ and then $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{v}_t \bar{v}_t' = O_p\left(\frac{1}{NT}\right)$. Similarly, $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{v}_t f_t' \Gamma_i = O_p\left(\frac{1}{\sqrt{NT}}\right)$ and $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{\Gamma}' f_t \bar{v}_t' = O_p\left(\frac{1}{\sqrt{NT}}\right)$. Thus, $\frac{1}{T}\bar{X}_\Delta' \bar{X}_\Delta = O_p(1) + O_p\left(\frac{1}{NT}\right) = O_p(1)$.

Similarly,

$$\begin{aligned}\frac{1}{T}X_i' \bar{X}_\Delta &= \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} x_{it} \bar{x}_t' = \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \left(\Gamma_i' f_t + v_{it}\right)\left(\bar{\Gamma}' f_t + \bar{v}_t\right)' \\ &= \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \bar{\Gamma} + \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} \bar{v}_t' \\ &+ \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t \bar{v}_t' + \frac{1}{T}\sum\nolimits_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f_t' \bar{\Gamma}.\end{aligned}$$

$\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \bar{\Gamma} = O_p(1)$ is obvious. For the term $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} \bar{v}_t' = \frac{1}{T}V_{i\Delta}' \bar{V}_\Delta$,

$$\frac{1}{T}V_{i\Delta}' \bar{V}_\Delta = \frac{1}{TN}V_{i\Delta}' V_{i\Delta} + \frac{1}{T}V_{i\Delta}' \bar{V}_{-i,\Delta},$$

where $\bar{V}_{-i,\Delta} = \frac{1}{N}\sum_{j=1,i\neq j}^{N} V_{j,\Delta}$. It's obvious that $V_{i\Delta}' V_{i\Delta} = O_p(1)$ uniformly over i and then $\frac{1}{TN} V_{i\Delta}' V_{i\Delta} = O_p\left(\frac{1}{TN}\right)$. Since $V_{i\Delta}'$ and $\bar{V}_{-i,\Delta}$ are independent under Assumption 3 and for the l^{th} row of $V_{i\Delta}'\bar{V}_{-i,\Delta}$denoted by $V_{il\Delta}'\bar{V}_{-i,\Delta}$

$$\begin{aligned}
\sup_i Var\left(\frac{1}{T} V_{il\Delta}' \bar{V}_{-i,\Delta}\right) &= \sup_i Var(\frac{1}{T}\sum_{t=k_1}^{k_1^0} V_{ilt}'\bar{V}_{-i,t}) \\
&= O(N^{-1}) \sup_i E(\frac{1}{T^2}\textstyle\sum_{t=k_1}^{k_1^0}\sum_{t'=k_1}^{k_1^0} V_{ilt}V_{ilt'}').
\end{aligned}$$

Under Assumption 9, $\sup_i E\left(\frac{1}{T}\sum_{t=k_1}^{k_1^0}\sum_{t'=k_1}^{k_1^0} V_{ilt}V_{ilt'}'\right) = O(\frac{1}{T^2})$. Thus,

$$\sup_i Var\left(\frac{1}{\sqrt{T}} V_{il\Delta}'\bar{V}_{-i,\Delta}\right) = O\left(\frac{1}{NT^2}\right)$$

and then $\frac{1}{T}V_{i\Delta}'\bar{V}_{-i,\Delta} = O_p\left(\frac{1}{T\sqrt{N}}\right)$. We obtain

$$\frac{1}{T}V_{i\Delta}'\bar{V}_{\bigstar} = O_p\left(\frac{1}{NT}\right) + O_p\left(\frac{1}{T\sqrt{N}}\right) = O_p\left(\frac{1}{T\sqrt{N}}\right),$$

$\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]}\Gamma_i' f_t \bar{v}_t' = O_p\left(\frac{1}{\sqrt{YN}}\right)$and $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f_t'\bar{\Gamma} = O_p\left(\frac{1}{\sqrt{T}}\right)$. Thus, $\frac{1}{T}X_i'\bar{X}_\Delta = O_p(1)$. Lastly, we obtain $\frac{1}{T}Z_{i\Delta}'Z_{i\Delta} = O_p(1)$.

Similarly,

$$\begin{aligned}
\frac{1}{T^2}X_{i\bigstar}'X_{i\bigstar} &= \frac{1}{T^2}\textstyle\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} x_{it}x_{it}' = \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]}\left(\Gamma_i' f_t + v_{it}\right)\left(\Gamma_i' f_t + v_{it}\right)' \\
&= \frac{1}{T^2}\textstyle\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]}\Gamma_i' f_t f_t'\Gamma_i + \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} v_{it}v_{it}' \\
&+ \frac{1}{T^2}\textstyle\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]}\Gamma_i' f_t v_{it}' + \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} v_{it} f_t'\Gamma_i.
\end{aligned}$$

Under equation (54), $\frac{1}{T^2}\sum_{t=1}^{T}\Gamma_i' f_t f_t'\Gamma_i = O_p(1)$. Under equation (55), $\frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} v_{it}v_{it}' = o_p(1)$. $\frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]}\Gamma_i' f_t v_{it}' = O_p\left(\frac{1}{T}\right)$ and $\frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} v_{it} f_t'\Gamma_i = O_p\left(\frac{1}{T}\right)$. Thus, $\frac{1}{T^2}X_{i\bigstar}'X_{i\bigstar} = O_p(1)$. $\frac{1}{T^2}\bar{X}_{\bigstar}'\bar{X}_{\bigstar} = O_p(1)$ and $\frac{1}{T^2}X_{i\bigstar}'\bar{X}_{\bigstar} = O_p(1)$ are shown similarly. Lastly, $\frac{1}{T^2}Z_{i\bigstar}'Z_{i\bigstar} = O_p(1)$.

(ii) Under Assumption 8(v) and Lemma 1(i), for large T, $Var\left(Z_{i\Delta}'\varepsilon_{i\bigstar}\right) = Z_{i\Delta}'\Sigma_{\varepsilon,i}Z_{i\Delta} = O(T)$. Thus, $\frac{1}{\sqrt{T}}Z_{i\Delta}'\varepsilon_{i\bigstar} = O_p(1)$. $\frac{1}{T}Z_{i\bigstar}'\varepsilon_{i\bigstar} = O_p(1)$ is shown similarly.

(iii) is proved as the arguments of (ii) and is omitted.

(iv)Under Assumption 9, we have $E\left\|\bar{v}_t\right\|^2 = \frac{1}{N^2}\sum_{i=1}^{N} E\left\|v_{it}\right\|^2 = O\left(\frac{1}{N}\right)$. Also, $E\left\|\frac{1}{T}\bar{V}_{\bigstar}'\bar{V}_{\bigstar}\right\| \leq \frac{1}{T}\sum_{t=k_1}^{k_2}\left(E\left\|\bar{v}_t\right\|^2\right)$. Thus, $E\left\|\frac{1}{T}\bar{V}_{\bigstar}'\bar{V}_{\bigstar}\right\| = O\left(\frac{1}{N}\right)$ and then $\frac{1}{T}\bar{V}_{\bigstar}'\bar{V}_{\bigstar} = O_p\left(\frac{1}{N}\right)$. Similarly, $E\left\|\frac{1}{T}\bar{V}_{\Delta}'\bar{V}_{\Delta}\right\| = O\left(\frac{1}{N}\right)$ and then $\bar{V}_{\Delta}'\bar{V}_{\Delta} = O_p\left(\frac{1}{N}\right)$.

Since $X_{i\Delta} = F_\Delta\Gamma_i + V_{i\Delta}$,

$$T^{-1/2}X_{i\Delta}'\bar{V}_{\bigstar} = T^{-1/2}\Gamma_i' F_\Delta'\bar{V}_{\bigstar} + T^{-1/2}V_{i\Delta}'\bar{V}_{\bigstar}.$$

For the first term, $T^{-1/2}F'_\Delta\bar{V}_\bigstar = T^{-1/2}\sum_{t=k_1}^{k_1^0} f_t\bar{v}'_t$. Consider the l^{th}row of matrix $T^{-1/2}F'_\Delta\bar{V}_\bigstar$, under the Assumption 9 that v_{it} are independent of common factor,

$$Var(T^{-1/2}\sum_{t=k_1}^{k_1^0} f_{lt}\bar{v}_t) = [T^{-1}E(\textstyle\sum_{t=k_1}^{k_1^0} f_{lt}f_{lt})]E(\sum_{t=k_1}^{k_1^0}\bar{v}_t\bar{v}'_t).$$

Since $E\left(\sum_{t=k_1}^{k_1^0}\bar{v}_t\bar{v}'_t\right) = O(\frac{1}{N})$ and $\frac{1}{T}F'_\Delta F_\Delta = O_p(1)$, $Var\left(T^{-1/2}\sum_{t=k_1}^{k_1^0} f_{lt}\bar{v}_t\right) = O(\frac{1}{N})$. Thus, $T^{-1/2}F'_\Delta\bar{V}_\bigstar = O_p\left(N^{-1/2}\right)$ and then $T^{-1/2}\Gamma'_iF'_\Delta\bar{V}_\bigstar = O_p\left(N^{-1/2}\right)$, under Assumption 3 that Γ_i is bounded.

For the second term $T^{-1/2}V'_{i\Delta}\bar{V}_\bigstar = T^{-1/2}V'_{i\Delta}\bar{V}_\Delta$,

$$T^{-1/2}V'_{i\Delta}\bar{V}_\Delta = T^{-1/2}N^{-1}V'_{i\Delta}V_{i\Delta} + T^{-1/2}V'_{i\Delta}\bar{V}_{-i,\Delta},$$

where $\bar{V}_{-i,\Delta} = \frac{1}{N}\sum_{j=1,i\neq j}^{N} V_{j,\Delta}$. It's obvious that $V'_{i\Delta}V_{i\Delta} = O_p(1)$ uniformly over i and then $T^{-1/2}N^{-1}V'_{i\Delta}V_{i\Delta} = O_p\left(T^{-1/2}N^{-1}\right)$. Since $V'_{i\Delta}$ and $\bar{V}_{-i,\Delta}$ are independent under Assumption 3 and for the l^{th} row of $V'_{i\Delta}\bar{V}_{-i,\Delta}$denoted by $V'_{il\Delta}\bar{V}_{-i,\Delta}$

$$\begin{aligned}\sup_i Var\left(T^{-1/2}V'_{il\Delta}\bar{V}_{-i,\Delta}\right) &= \sup_i Var(T^{-1/2}\textstyle\sum_{t=k_1}^{k_1^0} V'_{ilt}\bar{V}_{-i,t}) \\ &= O\left(N^{-1}\right)\sup_i E(T^{-1}\textstyle\sum_{t=k_1}^{k_1^0}\sum_{t'=k_1}^{k_1^0} V_{ilt}V'_{ilt'}).\end{aligned}$$

Under Assumption 9, $\sup_i E\left(\frac{1}{T}\sum_{t=k_1}^{k_1^0}\sum_{t'=k_1}^{k_1^0} V_{ilt}V'_{ilt'}\right) = O(\frac{1}{T})$. Thus,

$$\sup_i Var\left(\frac{1}{\sqrt{T}}V'_{il\Delta}\bar{V}_{-i,\Delta}\right) = O\left(\frac{1}{NT}\right)$$

and then $\frac{1}{\sqrt{T}}V'_{i\Delta}\bar{V}_{-i,\Delta} = O_p\left(\frac{1}{\sqrt{TN}}\right)$. We obtain

$$\frac{1}{\sqrt{T}}V'_{i\Delta}\bar{V}_\bigstar = O_p\left(\frac{1}{\sqrt{TN}}\right) + O_p\left(\frac{1}{\sqrt{TN}}\right) = O_p\left(\frac{1}{\sqrt{TN}}\right).$$

Lastly,

$$T^{-1/2}X'_{i\Delta}\bar{V}_\bigstar = O_p\left(N^{-1/2}\right) + O_p\left(T^{-1/2}N^{-1/2}\right) = O_p\left(N^{-1/2}\right).$$

According to $\bar{X}_\Delta = F_\Delta\bar{\Gamma} + \bar{V}_\Delta$,

$$T^{-1/2}\bar{X}'_\Delta\bar{V}_\bigstar = T^{-1/2}\bar{\Gamma}'F'_\Delta\bar{V}_\bigstar + T^{-1/2}\bar{V}'_\Delta\bar{V}_\bigstar.$$

Since $T^{-1/2}\bar{\Gamma}'F'_\Delta\bar{V}_\bigstar = O_p\left(N^{-1/2}\right)$ and $T^{-1/2}\bar{V}'_\Delta\bar{V}_\bigstar = O_p\left(T^{-1/2}N^{-1}\right)$, thus, $T^{-1/2}\bar{X}'_\Delta\bar{V}_\bigstar = O_p\left(N^{-1/2}\right)$.

We lastly conclude $T^{-1/2}Z'_{i\Delta}\bar{V}_\bigstar = O_p\left(N^{-1/2}\right)$.

$T^{-1}Z'_{i\bigstar}\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$ can be shown as the proof of $T^{-1/2}Z'_{i\Delta}\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$.

Lemma 2. Under Assumptions 1-8, uniformly on $K\left(C_k\right)$,
(i) $\sum_{i=1}^{N}\left(\delta_{i2}-\delta_{i1}\right)' Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar} = O_p\left(\sqrt{T\phi_{N,1}}\right)$;
(ii) $\sum_{i=1}^{N}\varepsilon^{*\prime}_{i\bigstar}M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z'_{i\Delta}M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar} = O_p\left(N\right)$;
(iii) $\sum_{i=1}^{N}\varepsilon^{*\prime}_{i\Diamond}M_{Z_{i\Diamond}}Z_{i\Delta}\left(Z'_{i\Delta}M_{Z_{i\Diamond}}Z_{i\Delta}\right)^{-1}Z'_{i\Delta}M_{Z_{i\Diamond}}\varepsilon^*_{i\Diamond} = O_p\left(N\right)$.

Proof of Lemma 2. (i) Since $\varepsilon^*_{it} = \varepsilon_{it} - \bar{v}'_t\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_i(k_1)$, $\varepsilon^*_{i\bigstar} = \varepsilon_{i\bigstar} - \bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}$ for the interval $[k_1^0+1, k_2]$. Plugging the expression of $\varepsilon^*_{i\bigstar}$ into $\sum_{i=1}^{N}2(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar}$ gives

$$
\begin{aligned}
& \textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}M_{Z_{i\bigstar}}\varepsilon^*_{i\bigstar} \\
= & \textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\varepsilon_{i\bigstar} - \sum_{i=1}^{N}\left(\delta_{i2}-\delta_{i1}\right)'Z'_{i\Delta}\bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i} \\
+ & \textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}Z_{i\bigstar}(Z'_{i\bigstar}Z_{i\bigstar})^{-1}Z'_{i\bigstar}\varepsilon_{i\bigstar} \\
- & \textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}Z_{i\bigstar}(Z'_{i\bigstar}Z_{i\bigstar})^{-1}Z'_{i\bigstar}\bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}.
\end{aligned}
$$

For the first term,

$$
\begin{aligned}
Var[\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\varepsilon_{i\bigstar}] &= \textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\Sigma_{\varepsilon,i}Z_{i\Delta}\left(\delta_{i2}-\delta_{i1}\right) \\
&= O\left(T\phi_{N,1}\right),
\end{aligned}
$$

Thus, $\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\varepsilon_{i\bigstar} = O_p\left(\sqrt{T\phi_{N,1}}\right)$. By Lemma 1(i) and Assumption 8, $\frac{1}{T}Z'_{i\Delta}\Sigma_{\varepsilon,i}Z_{i\Delta} = O_p(1)$. The second equality above is due to the fact that $\frac{1}{T}\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\Sigma_{\varepsilon,i}Z_{i\Delta}(\delta_{i2}-\delta_{i1})$ is of the same order of magnitude as $\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'(\delta_{i2}-\delta_{i1})$.

Consider the second term. Since $\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_i$ is bounded under Assumption 3,

$\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_i$ is of the same order of magnitude as $\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\bar{V}_{\bigstar}$. By Lemma 1(iv),

$$
\begin{aligned}
Var\left[\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}\bar{V}_{\bigstar}\right] &= \sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z'_{i\Delta}Var\left(\bar{V}_{\bigstar}\right)Z_{i\Delta}\left(\delta_{i2}-\delta_{i1}\right) \\
&= O\left(\frac{T}{N}\right)O\left(\phi_{N,1}\right) = O\left(\frac{T}{N}\phi_{N,1}\right).
\end{aligned}
$$

Thus, $\sum_{i=1}^{N}\left(\delta_{i2}-\delta_{i1}\right)'Z'_{i\Delta}\bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_i = O_p\left(\sqrt{\frac{T}{N}\phi_{N,1}}\right)$. By Lemma 1(i) and (iv) ,$\frac{N}{T}Z'_{i\Delta}Var\left(\bar{V}_{\bigstar}\right)Z_{i\Delta} = O_p(1)$. Thus $\sum_{i=1}^{N}\left(\delta_{i2}-\delta_{i1}\right)'Z'_{i\Delta}Var\left(\bar{V}_{\bigstar}\right)Z_{i\Delta}\left(\delta_{i2}-\delta_{i1}\right)$ is of the same order of magnitude of $O(TN^{-1})\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'(\delta_{i2}-\delta_{i1}) = O\left(TN^{-1}\phi_{N,1}\right)$.

For the third term, by Lemma 1(i) and (ii),

$$
Var\left(\textstyle\sum_{i=1}^{N}\left(\delta_{i2}-\delta_{i1}\right)'\left(T^{-1}Z'_{i\Delta}Z_{i\bigstar}\right)\left(T^{-1}Z'_{i\bigstar}Z_{i\bigstar}\right)^{-1}\left(T^{-1}Z'_{i\bigstar}\varepsilon_{i\bigstar}\right)\right) = O\left(\phi_{N,1}\right),
$$

Thus,
$$\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'\frac{Z_{i\Delta}'Z_{i\bigstar}}{T}\left(\frac{Z_{i\bigstar}'Z_{i\bigstar}}{T^2}\right)^{-1}\frac{Z_{i\bigstar}'\varepsilon_{i\bigstar}}{T}=O_p\left(\sqrt{\phi_{N,1}}\right).$$

For the fourth term, by Lemma 1(i) and (iv),

$$\begin{aligned}&Var\left(\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'(T^{-1}Z_{i\Delta}'Z_{i\bigstar})\left(T^{-1}Z_{i\bigstar}'Z_{i\bigstar}\right)^{-1}(T^{-1}Z_{i\bigstar}'\bar{V}_{\bigstar})\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\right)\\=&O(1)\,O(\phi_{N,1})\,O\left(\frac{1}{N}\right)=O\left(\frac{\phi_{N,1}}{N}\right),\end{aligned}$$

Thus, $\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'Z_{i\bigstar}\left(Z_{i\bigstar}'Z_{i\bigstar}\right)^{-1}Z_{i\bigstar}'\bar{V}_{\bigstar}\bar{\Gamma}'\left(\bar{\Gamma}\bar{\Gamma}'\right)^{-1}\gamma_{1i}=O_p\left(\sqrt{\frac{\phi_{N,1}}{N}}\right)$.

Lastly, $\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}^{*}=O_p\left(\sqrt{T\phi_{N,1}}\right)$.

(ii) Since $\varepsilon_{it}^{*}=\varepsilon_{it}-\bar{v}_t'\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}$ in interval $[k_1^0+1,k_2]$,

$$\begin{aligned}&\varepsilon_{i\bigstar}^{*\prime}M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}^{*}\\=&\ \varepsilon_{i\bigstar}'M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}\\&-\varepsilon_{i\bigstar}'M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\bar{V}_{\bigstar}\bar{\Gamma}'\left(\bar{\Gamma}\bar{\Gamma}'\right)^{-1}\gamma_{1i}\\&-\gamma_{1i}'\left(\bar{\Gamma}\bar{\Gamma}'\right)^{-1}\bar{\Gamma}\bar{V}_{\bigstar}'M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}\\&+\gamma_{1i}'\left(\bar{\Gamma}\bar{\Gamma}'\right)^{-1}\bar{\Gamma}\bar{V}_{\bigstar}'M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\bar{V}_{\bigstar}\bar{\Gamma}'\left(\bar{\Gamma}\bar{\Gamma}'\right)^{-1}\gamma_{1i}.\end{aligned}$$

For the first term, by Lemma 1(i) and (ii),

$$\begin{aligned}\frac{1}{T}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}&=\frac{1}{T}Z_{i\Delta}'\varepsilon_{i\bigstar}-\frac{1}{T}(T^{-1}Z_{i\Delta}'Z_{i\bigstar})\left(T^{-1}Z_{i\bigstar}'Z_{i\bigstar}\right)^{-1}(T^{-1}Z_{i\bigstar}'\varepsilon_{i\bigstar})\\&=\frac{1}{T}Z_{i\Delta}'\varepsilon_{i\bigstar}+O_p\left(\frac{1}{T}\right),\end{aligned}$$

Thus, $\frac{1}{T}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}$ is of same order of magnitude as $\frac{1}{T}Z_{i\Delta}'\varepsilon_{i\bigstar}$, as $T\to\infty$. Similarly, $\frac{1}{T}Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}$ is of same order of magnitude as $\frac{1}{T}Z_{i\Delta}'Z_{i\Delta}$, under Lemma 1(i) as $(N,T)\to\infty$. In addition, since $M_{Z_{i\bigstar}}Z_{i\Delta}\left(\frac{1}{T}Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}$ is positive semidefinite,

$\varepsilon_{i\bigstar}'M_{Z_{i\bigstar}}Z_{i\Delta}(T^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta})^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}\geq 0$. In addition, $|k_1-k_1^0|$ is bounded on $K(C_k)$,

$$\frac{1}{N}\sum_{i=1}^{N}(T^{-1/2}\varepsilon_{i\bigstar}'M_{Z_{i\bigstar}}Z_{i\Delta})(T^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta})^{-1}(T^{-1/2}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar})=O_p(1).$$

The above equation is due to the fact that since both $\frac{1}{\sqrt{T}}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}$ and $\frac{1}{T}Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}$ is of order $O_p(1)$, the order of above term is same as the order of $\frac{1}{N}\sum_{i=1}^{N}O_p(1)=O_p(1)$.

Thus, by Lemma 1(i) and (ii),

$$\sum_{i=1}^{N}(T^{-1/2}\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta})(T^{-1}Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1}(T^{-1/2}Z_{i\Delta}' M_{Z_{i\star}}\varepsilon_{i\bigstar}) = O_p\left(N\right).$$

For the second term, by Lemma 1(i), (ii) and (iv),

$$\begin{aligned}&\sum_{i=1}^{N}(T^{-1/2}\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta})(T^{-1}Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1}(T^{-1/2}Z_{i\Delta}' M_{Z_{i\star}}\varepsilon_{i\bigstar})\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\\ =&O_p\left(1\right)O_P\left(N^{-1/2}\right)O_p\left(\sqrt{N}\right) = O_p\left(1\right).\end{aligned}$$

The first equality above is due to the fact that the term $\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_i$ is bounded under Assumption 1 and under the Lemma 1(i), (ii) and (iv), the order of the second term is same as sum of finite elements on $K(C_k)$, also according to the proof of Lemma 7(iv) in BFK(2016) and thus $\sum_{i=1}^{N} T^{-1/2}\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta} = O_p\left(\sqrt{N}\right)$.

For the third term, by Lemma 1(i), (iii) and (iv),

$$\sum_{i=1}^{N}\gamma_{1i}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}\bar{V}_{\bigstar}' M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1}Z_{i\Delta}' M_{Z_{i\star}}\varepsilon_{i\bigstar} = O_p\left(1\right),$$

which is showed similarly as the second term above.

For the fourth term, by Lemma 1(i) and (iv),

$$\begin{aligned}&\sum_{i=1}^{N}\gamma_{1i}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}(T^{-1/2}Z_{i\Delta}' M_{Z_{i\star}}\bar{V}_{\bigstar})(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta}T^{-1})^{-1}(T^{-1/2}Z_{i\Delta}' M_{Z_{i\star}}\bar{V}_{\bigstar})\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\\ &= O_p\left(1\right)O_P\left(N^{-1/2}\right)O_P\left(N^{-1/2}\right)O_p(N) = O_p\left(1\right).\end{aligned}$$

Combining these four terms together, we obtain

$$\sum_{i=1}^{N}\varepsilon_{i\bigstar}^{*\prime} M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1}Z_{i\Delta}' M_{Z_{i\star}}\varepsilon_{i\bigstar}^{*} = O_p(N).$$

(iii) can be proved in the same way as (ii) by Lemma 1.

Lemma 3. Under Assumptions 1-10 and 12, uniformly on $K(C_k)$ and for $i = 1, ..., N$, as $(N, T) \to \infty$,

(i) $\frac{1}{T}Z_{i\Delta}'Z_{i\Delta} = O_p\left(1\right)$, $\frac{1}{T^2}Z_{i\bigstar}'Z_{i\bigstar} = O_p\left(1\right)$;

(ii) $T^{-1/2}Z_{i\Delta}'\varepsilon_{i\bigstar} = T^{-1/2}Z_{i\Delta}'\varepsilon_{i\Delta} = O_p\left(1\right)$, $T^{-1}Z_{i\bigstar}'\varepsilon_{i\bigstar} = O_p\left(1\right)$;

(iii) $T^{-1/2}Z_{i\Delta}'\varepsilon_{i\diamondsuit} = T^{-1/2}Z_{i\Delta}'\varepsilon_{i\Delta} = O_p\left(1\right)$, $T^{-1}Z_{i\diamondsuit}'\varepsilon_{i\diamondsuit} = O_p\left(1\right)$;

(iv)$T^{-2}\bar{V}_{\bigstar}'\bar{V}_{\bigstar} = O_p\left(N^{-1}\right)$, $T^{-1}Z_{i\Delta}'\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$, $T^{-3/2}Z_{i\bigstar}'\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$.

Proof of Lemma 3. (i) According to Lemma 1(b) in the supplementary appendix of Baltagi, Kao and Liu (2017), for any $0 \le \tau_1 < \tau_2 \le 1$ and under Assumption 12,

$$\frac{1}{T^2}\sum_{t=[\tau_1 T]}^{[\tau_2 T]} v_{it}v_{it}' = O_p\left(\tau_2 - \tau_1\right). \tag{56}$$

Plugging $x_{it} = \Gamma_i' f_t + v_{it}$ into $X_{i\Delta}$ gives

$$\begin{aligned}\frac{1}{T} X_{i\Delta}' X_{i\Delta} &= \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} x_{it} x_{it}' = \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \left(\Gamma_i' f_t + v_{it}\right) \left(\Gamma_i' f_t + v_{it}\right)' \\ &= \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \Gamma_i + \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} v_{it}' \\ &+ \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t v_{it}' + \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f_t' \Gamma_i.\end{aligned}$$

Same as Lemma 1(i) , $\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \Gamma_i = O_p(1)$. From equation (56) and $|\lambda_1 - \lambda_1^0| = O_p\left(\frac{1}{T}\right)$, $\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} v_{it}' = O_p(1)$. According to Lemma 2(d) in the supplementary appendix of Baltagi, Kao and Liu (2017), for any $0 \leq \tau_1 < \tau_2 \leq 1$ and under Assumption 6 and 12, uniformly over i,

$$\frac{1}{T^2} \sum_{t=[\tau_1 T]}^{[\tau_2 T]} f_t v_{it} = O_p\left(\tau_2 - \tau_1\right), \tag{57}$$

From equation (57) and $|\lambda_1 - \lambda_1^0| = O_p\left(\frac{1}{T}\right)$, $\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t v_{it}' = O_p(1)$ and $\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f_t' \Gamma_i = O_p(1)$. Thus,

$$\frac{1}{T} X_{i\Delta}' X_{i\Delta} = O_p(1) + O_p(1) = O_p(1).$$

Plugging $\bar{x}_t = \bar{\Gamma}' f_t + \bar{v}_t$ into $\bar{X}_\Delta$ gives

$$\begin{aligned}\frac{1}{T} \bar{X}_\Delta' \bar{X}_\Delta &= \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{x}_t \bar{x}_t' = \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} (\bar{\Gamma}' f_t + \bar{v}_t)(\bar{\Gamma}' f_t + \bar{v}_t)' \\ &= \bar{\Gamma}' (\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} f_t f_t') \bar{\Gamma} + \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{v}_t \bar{v}_t'.\end{aligned}$$

Same as Lemma 1(i) , $\bar{\Gamma}' (\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} f_t f_t') \bar{\Gamma} = O_p(1)$. Under Assumption 12, we have $\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \bar{v}_t \bar{v}_t' = O_p\left(\frac{1}{N}\right)$. Thus, $\frac{1}{T} \bar{X}_\Delta' \bar{X}_\Delta = O_p(1) + O_p\left(\frac{1}{N}\right) = O_p(1)$.

Similarly,

$$\begin{aligned}\frac{1}{T} X_i' \bar{X}_\Delta &= \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} x_{it} \bar{x}_t' = \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \left(\Gamma_i' f_t + v_{it}\right) \left(\bar{\Gamma}' f_t + \bar{v}_t\right)' \\ &= \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \bar{\Gamma} + \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} \bar{v}_t' \\ &+ \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t \bar{v}_t' + \frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f_t' \bar{\Gamma}.\end{aligned}$$

$\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} \Gamma_i' f_t f_t' \bar{\Gamma} = O_p(1)$ is obvious. For the term $\frac{1}{T} \sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} \bar{v}_t' = \frac{1}{T} V_{i\Delta}' \bar{V}_\Delta$,

$$\frac{1}{T} V_{i\Delta}' \bar{V}_\Delta = \frac{1}{TN} V_{i\Delta}' V_{i\Delta} + \frac{1}{T} V_{i\Delta}' \bar{V}_{-i,\Delta},$$

where $\bar{V}_{-i,\Delta} = \frac{1}{N} \sum_{j=1, i \neq j}^{N} V_{j,\Delta}$. It's obvious that uniformly over i, $\frac{1}{TN} V_{i\Delta}' V_{i\Delta} = O_p\left(\frac{1}{N}\right)$. Since $V_{i\Delta}'$ and $\bar{V}_{-i,\Delta}$ are independent under Assumption 3 and for the l^{th}

row of $V'_{i\Delta}\bar{V}_{-i,\Delta}$denoted by $V'_{il\Delta}\bar{V}_{-i,\Delta}$

$$\begin{aligned}\sup_i Var(\frac{1}{T}V'_{il\Delta}\bar{V}_{-i,\Delta}) &= \sup_i Var(\frac{1}{T}\sum_{t=k_1}^{k_1^0} V'_{ilt}\bar{V}_{-i,t})\\ &= E(\frac{1}{T}\sum_{t=k_1}^{k_1^0}\sum_{t'=k_1}^{k_1^0}\bar{V}_{-i,t}\bar{V}'_{-i,t'})\sup_i E(\frac{1}{T}\sum_{t=k_1}^{k_1^0} V_{ilt}V'_{ilt}).\end{aligned}$$

Under Assumption 12, $E(\frac{1}{T}\sum_{t=k_1}^{k_1^0}\bar{V}_{-i,t}\bar{V}'_{-i,t}) = O(N^{-1})$ and $\sup_i E(\frac{1}{T}\sum_{t=k_1}^{k_1^0} V_{ilt}V'_{ilt}) = O(1)$. Thus, $\sup_i Var(\frac{1}{T}V'_{il\Delta}\bar{V}_{-i,\Delta}) = O(N^{-1})$ and then $\frac{1}{T}V'_{i\Delta}\bar{V}_{-i,\Delta} = O_p(N^{-1/2})$. We obtain $\frac{1}{T}V'_{i\Delta}\bar{V}_{\bigstar} = O_p(N^{-1}) + O_p(N^{-1/2}) = O_p(N^{-1/2})$. $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]}\Gamma'_i f_t \bar{v}'_t = O_p(N^{-1/2})$ is proved similarly. Since $\frac{1}{T}\sum_{t=[\lambda_1 T]}^{[\lambda_1^0 T]} v_{it} f'_t\bar{\Gamma} = O_p(1)$, thus, $\frac{1}{T}X'_i\bar{X}_\Delta = O_p(1) + O_p(N^{-1/2}) = O_p(1)$. Lastly, we obtain $\frac{1}{T}Z'_{i\Delta}Z_{i\Delta} = O_p(1)$.

Similarly,

$$\begin{aligned}\frac{1}{T^2}X'_{i\bigstar}X_{i\bigstar} &= \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} x_{it}x'_{it} = \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]}(\Gamma'_i f_t + v_{it})(\Gamma'_i f_t + v_{it})'\\ &= \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]}\Gamma'_i f_t f'_t\Gamma_i + \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} v_{it}v'_{it}\\ &+ \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]}\Gamma'_i f_t v'_{it} + \frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} v_{it} f'_t\Gamma_i.\end{aligned}$$

Under equation (56), $\frac{1}{T^2}\sum_{t=[\lambda_1 T]}^{[\lambda_2 T]} v_{it}v'_{it} = o_p(1)$. Thus, $\frac{1}{T^2}X'_{i\bigstar}X_{i\bigstar} = O_p(1)$. $\frac{1}{T^2}\bar{X}'_{\bigstar}\bar{X}_{\bigstar} = O_p(1)$ and $\frac{1}{T^2}\bar{X}'_{\bigstar}X_{i\bigstar} = O_p(1)$ are shown similarly. Lastly, $\frac{1}{T^2}Z'_{i\bigstar}Z_{i\bigstar} = O_p(1)$.

(ii) According to Lemma 2(d) in the supplementary appendix of Baltagi, Kao and Liu(2017), for any $0 \leq \tau_1 < \tau_2 \leq 1$ and under Assumptions 6,

$$T^{-2}\sum_{t=[\tau_1 T]}^{[\tau_2 T]} f_t\varepsilon_{it} = O_p(\tau_2 - \tau_1), \tag{58}$$

uniformly over i. Since $|\lambda_1 - \lambda_1^0| = O_p(T^{-1})$, $T^{-2}F'_\Delta\varepsilon_{i\Delta} = T^{-2}\sum_{t=k_1}^{k_1^0} f_t\varepsilon_{it} = O_p(T^{-1})$ and then $T^{-1}F'_\Delta\varepsilon_{i\Delta} = O_p(1)$. Similarly, $T^{-1}V'_{i\Delta}\varepsilon_{i\Delta} = O_p(1)$. Thus,

$$T^{-1}X'_{i\Delta}\varepsilon_{i\bigstar} = T^{-1}\Gamma'_i F'_\Delta\varepsilon_{i\Delta} + T^{-1}V'_{i\Delta}\varepsilon_{i\Delta} = O_p(1),$$

under the Assumption 3 that Γ_i is bounded.

Plugging $\bar{x}_t = \bar{\Gamma}' f_t + \bar{v}_t$ gives

$$T^{-1}\bar{X}'_\Delta\varepsilon_{i\bigstar} = T^{-1}\bar{\Gamma}' F'_\Delta\varepsilon_{i\Delta} + T^{-1}\bar{V}'_\Delta\varepsilon_{i\Delta}.$$

Under the Assumption 3 that $\bar{\Gamma}$ is bounded, $T^{-1}\bar{\Gamma}' F'_\Delta\varepsilon_{i\Delta} = O_p(1)$. $T^{-1}\bar{V}'_\Delta\varepsilon_{i\Delta} = O_p(N^{-1/2})$ and then $T^{-1}\bar{X}'_\Delta\varepsilon_{i\bigstar} = O_p(1) + O_p(N^{-1/2}) = O_p(1)$.

Lastly, we obtain $T^{-1}Z'_{i\Delta}\varepsilon_{i\bigstar} = O_p(1)$. $T^{-2}Z'_{i\bigstar}\varepsilon_{i\bigstar} = O_p(1)$ is shown similarly and omitted.

(iii) is proved as the arguments of (ii).

(iv) Since

$$T^{-2}\bar{V}'_{\bigstar}\bar{V}_{\bigstar} = T^{-2}\sum_{t=k_1}^{k_2}(\bar{v}_t\bar{v}'_t) = T^{-2}\sum_{t=k_1}^{k_2}(\frac{1}{N}\sum_{i=1}^{N}v_{it})(\frac{1}{N}\sum_{i=1}^{N}v_{it})'.$$

under Assumption 12,

$$\begin{aligned} T^{-2}E(\bar{V}'_{\bigstar}\bar{V}_{\bigstar}) &= T^{-2}E(\sum_{t=k_1}^{k_2}(N^{-1}\sum_{i=1}^{N}v_{it})(\frac{1}{N}\sum_{i=1}^{N}v_{it})') \\ &\leq \sup_i N^{-1}T^{-2}E[\sum_{t=k_1}^{k_2}v_{it}v'_{it}] = O\left(N^{-1}\right). \end{aligned}$$

Thus, $\frac{1}{T^2}\bar{V}'_{\bigstar}\bar{V}_{\bigstar} = O_p\left(\frac{1}{N}\right)$. $\frac{1}{T}\bar{V}'_{\Delta}\bar{V}_{\Delta} = O_p\left(\frac{1}{N}\right)$ is proved similarly and omitted.

Since $X_{i\Delta} = F_\Delta\Gamma_i + V_{i\Delta}$,

$$\frac{1}{T}X'_{i\Delta}\bar{V}_{\bigstar} = \frac{1}{T}\Gamma'_iF'_\Delta\bar{V}_{\bigstar} + \frac{1}{T}V'_{i\Delta}\bar{V}_{\bigstar}.$$

For the first term, $\frac{1}{T}F'_\Delta\bar{V}_{\bigstar} = \frac{1}{T}\sum_{t=k_1}^{k_1^0}f_t\bar{v}'_t$. Consider the l^{th}row of matrix $\frac{1}{T}F'_\Delta\bar{V}_{\bigstar}$, under the Assumption 12 that v_{it} are independent of common factor,

$$Var(T^{-1}\sum_{t=k_1}^{k_1^0}f_{lt}\bar{v}_t) = E(T^{-1}\sum_{t=k_1}^{k_1^0}f_{lt}f_{lt})E(T^{-1}\sum_{t=k_1}^{k_1^0}\bar{v}_t\bar{v}'_t).$$

Since $E(\frac{1}{T}\sum_{t=k_1}^{k_1^0}\bar{v}_t\bar{v}'_t) = O(\frac{1}{N})$ and $\frac{1}{T}F'_\Delta F_\Delta = O_p(1)$, $Var(\frac{1}{T}\sum_{t=k_1}^{k_1^0}f_{lt}\bar{v}_t) = O(N^{-1})$. Thus, $\frac{1}{T}F'_\Delta\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$ and then $\frac{1}{T}\Gamma'_iF'_\Delta\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$, under Assumption 3 that Γ_i is bounded.

For the second term $T^{-1/2}V'_{i\Delta}\bar{V}_{\bigstar} = T^{-1/2}V'_{i\Delta}\bar{V}_\Delta$,

$$\frac{1}{T}V'_{i\Delta}\bar{V}_\Delta = \frac{1}{TN}V'_{i\Delta}V_{i\Delta} + \frac{1}{T}V'_{i\Delta}\bar{V}_{-i,\Delta},$$

where $\bar{V}_{-i,\Delta} = \frac{1}{N}\sum_{j=1,i\neq j}^{N}V_{j,\Delta}$. It's obvious that $\frac{1}{T}V'_{i\Delta}V_{i\Delta} = O_p(1)$ uniformly over i and then $\frac{1}{TN}V'_{i\Delta}V_{i\Delta} = O_p\left(\frac{1}{N}\right)$. Since $V'_{i\Delta}$ and $\bar{V}_{-i,\Delta}$ are independent under Assumption 3 and for the l^{th} row of $V'_{i\Delta}\bar{V}_{-i,\Delta}$denoted by $V'_{il\Delta}\bar{V}_{-i,\Delta}$

$$\begin{aligned} &\sup_i||Var(T^{-1}V'_{il\Delta}\bar{V}_{-i,\Delta})|| = \sup_i||Var(T^{-1}\sum_{t=k_1}^{k_1^0}V'_{ilt}\bar{V}_{-i,t})|| \\ \leq &E(||T^{-1}\sum_{t=k_1}^{k_1^0}\bar{v}_t\bar{v}'_t||)\sup_iE(||T^{-1}\sum_{t=k_1}^{k_1^0}V_{ilt}V'_{ilt}||). \end{aligned}$$

Under Assumption 9, $\sup_i E(||T^{-1}\sum_{t=k_1}^{k_1^0}V_{ilt}V'_{ilt}||) = O(1)$ and $E(||T^{-1}\sum_{t=k_1}^{k_1^0}\bar{v}_t\bar{v}'_t||) = O\left(N^{-1}\right)$. Thus, $\sup_i Var\left(\frac{1}{T}V'_{il\Delta}\bar{V}_{-i,\Delta}\right) = O\left(N^{-1}\right)$ and then $\frac{1}{T}V'_{i\Delta}\bar{V}_{-i,\Delta} = O_p\left(N^{-1/2}\right)$. We obtain

$$\frac{1}{T}V'_{i\Delta}\bar{V}_{\bigstar} = O_p\left(N^{-1}\right) + O_p\left(N^{-1/2}\right) = O_p\left(N^{-1/2}\right).$$

Lastly, $\frac{1}{T}X_{i\Delta}'\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right).$

According to $\bar{X}_{\Delta} = F_{\Delta}\bar{\Gamma} + \bar{V}_{\Delta}$,

$$\frac{1}{T}\bar{X}_{\Delta}'\bar{V}_{\bigstar} = \frac{1}{T}\bar{\Gamma}'F_{\Delta}'\bar{V}_{\bigstar} + \frac{1}{T}\bar{V}_{\Delta}'\bar{V}_{\bigstar}.$$

Since $\frac{1}{T}\bar{\Gamma}'F_{\Delta}'\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$ and $\frac{1}{T}\bar{V}_{\Delta}'\bar{V}_{\bigstar} = O_p\left(N^{-1}\right)$, thus, $\frac{1}{T}\bar{X}_{\Delta}'\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$.

We lastly conclude $\frac{1}{T}Z_{i\Delta}'\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$. $\frac{1}{T^2}Z_{i\bigstar}'\bar{V}_{\bigstar} = O_p\left(N^{-1/2}\right)$ is proved similarly.

Lemma 4. Under Assumptions 1-10 and 12, uniformly on $K\left(C_k\right)$,

(i) $\sum_{i=1}^{N}\left(\delta_{i2}-\delta_{i1}\right)' Z_{i\Delta}' M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}^{*} = O_p\left(\sqrt{T\phi_{N,1}}\right) + O_p\left(T\sqrt{\frac{\phi_{N,1}}{N}}\right)$;

(ii) $\sum_{i=1}^{N}\varepsilon_{i\bigstar}^{*\prime} M_{Z_{i\bigstar}} Z_{i\Delta}\left(Z_{i\Delta}' M_{Z_{i\bigstar}} Z_{i\Delta}\right)^{-1} Z_{i\Delta}' M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}^{*} = O_p\left(N\right) + O_p\left(T\right)$;

(iii) $\sum_{i=1}^{N}\varepsilon_{i\diamondsuit}^{*\prime} M_{Z_{i\diamondsuit}} Z_{i\Delta}\left(Z_{i\Delta}' M_{Z_{i\diamondsuit}} Z_{i\Delta}\right)^{-1} Z_{i\Delta}' M_{Z_{i\diamondsuit}}\varepsilon_{i\diamondsuit}^{*} = O_p\left(N\right) + O_p\left(T\right)$.

Proof of Lemma 4. (i) since $\varepsilon_{it}^{*} = \varepsilon_{it} - \bar{v}_t'\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_i(k_1)$,$\varepsilon_{i\bigstar}^{*} = \varepsilon_{i\bigstar} - \bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}$ for the interval $[k_1^0+1, k_2]$. Plugging the expression of $\varepsilon_{i\bigstar}^{*}$ into $\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}^{*}$ gives

$$\begin{aligned}
&\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}^{*}\\
=&\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'\varepsilon_{i\bigstar}\\
&-\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'\bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\\
&+\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'Z_{i\bigstar}(Z_{i\bigstar}'Z_{i\bigstar})^{-1}Z_{i\bigstar}'\varepsilon_{i\bigstar}\\
&-\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'Z_{i\bigstar}(Z_{i\bigstar}'Z_{i\bigstar})^{-1}Z_{i\bigstar}'\bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}.
\end{aligned}$$

For the first term, by Lemma 3(ii) Assumption 4,

$$\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'\varepsilon_{i\bigstar} = O_p\left(\sqrt{\phi_{N,1}}\right)O_p\left(\sqrt{T}\right) = O_p\left(\sqrt{T\phi_{N,1}}\right).$$

For the second term, by Lemma 3(iv) and Assumptions 3 and 4,

$$\begin{aligned}
\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'Z_{i\Delta}'\bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i} &= O_p\left(\sqrt{\phi_{N,1}}\right)O_p\left(TN^{-1/2}\right)\\
&= O_p\left(T\phi_{N,1}^{1/2}N^{-1/2}\right).
\end{aligned}$$

For the third term, by Lemma 3(i) (ii) and Assumption 4,

$$T\textstyle\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})'(T^{-1}Z_{i\Delta}'Z_{i\bigstar})(T^{-2}Z_{i\bigstar}'Z_{i\bigstar})^{-1}(T^{-1}Z_{i\bigstar}'\varepsilon_{i\bigstar}) = O_p\left(\sqrt{\phi_{N,1}}\right).$$

For the fourth term, by Lemma 3(i) (iv) and Assumptions 3 and 4,

$$\begin{aligned}
&\sqrt{T}\textstyle\sum_{i=1}^{N}\left(\delta_{i2}-\delta_{i1}\right)'(T^{-1}Z_{i\Delta}'Z_{i\bigstar})(T^{-2}Z_{i\bigstar}'Z_{i\bigstar})^{-1}(T^{-3/2}Z_{i\bigstar}'\bar{V}_{\bigstar})\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\\
=&T^{1/2}O_p\left(\phi_{N,1}^{1/2}\right)O_p\left(N^{-1/2}\right) = O_p\left(T^{1/2}\phi_{N,1}^{1/2}N^{-1/2}\right).
\end{aligned}$$

Thus,

$$\sum_{i=1}^{N}(\delta_{i2}-\delta_{i1})' Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar}^{*} = O_p\left(\sqrt{T\phi_{N,1}}\right) + O_p\left(T\phi_{N,1}^{1/2}N^{-1/2}\right) + O_p\left(\sqrt{\phi_{N,1}}\right) + O_p\left(T^{1/2}\phi_{N,1}^{1/2}N^{-1/2}\right) = O_p\left(T^{1/2}\phi_{N,1}^{1/2}\right) + O_p(T\phi_{N,1}^{1/2}N^{-1/2}).$$

(ii) Since $\varepsilon_{it}^{*}=\varepsilon_{it}-\bar{v}_t'\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}$ in interval $[k_1^0+1,k_2]$,

$$\begin{aligned}&\varepsilon_{i\bigstar}^{*\prime} M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar}^{*}\\
=&\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar}\\
-&\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\star}} \bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\\
-&\gamma_{1i}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}\bar{V}_{\bigstar}' M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar}\\
+&\gamma_{1i}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}\bar{V}_{\bigstar}' M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\star}} \bar{V}_{\bigstar}\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}.\end{aligned}$$

Since $M_{Z_{i\star}} Z_{i\Delta}(T^{-1}Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\star}}$ is positive semidefinite,

$$\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta}(Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1} Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar} \geq 0.$$

In addition, $|k_1-k_1^0|$ is bounded on $K\left(C_k\right)$, by Lemma 3(i) (ii),

$$\frac{1}{N}\sum_{i=1}^{N}(T^{-1/2}\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta})(T^{-1}Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1}(T^{-1/2}Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar}) = O_p\left(1\right).$$

Thus, for the first term,

$$\sum_{i=1}^{N}(T^{-1/2}\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta})(T^{-1}Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1}(T^{-1/2}Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar}) = O_p\left(N\right).$$

For the second term, by Lemma 3(i)(ii)(iv) and Assumption 3,

$$\begin{aligned}&\sqrt{T}\sum_{i=1}^{N}(T^{-1/2}\varepsilon_{i\bigstar}' M_{Z_{i\star}} Z_{i\Delta})T^{-1}Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})(T^{-1}Z_{i\Delta}' M_{Z_{i\star}} \bar{V}_{\bigstar})\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\\
=&\sqrt{T}O_p\left(1\right)O_p\left(1\right)O_P\left(N^{-1/2}\right)O_p\left(\sqrt{N}\right) = O_p\left(\sqrt{T}\right).\end{aligned}$$

For the third term, by Lemma 3(i)(ii)(iv) and Assumption 3,

$$\begin{aligned}&\sqrt{T}\sum_{i=1}^{N}\gamma_{1i}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}(\bar{V}_{\bigstar}' M_{Z_{i\star}} Z_{i\Delta}T^{-1})\\
&\times(T^{-1}Z_{i\Delta}' M_{Z_{i\star}} Z_{i\Delta})^{-1}(T^{-1/2}Z_{i\Delta}' M_{Z_{i\star}} \varepsilon_{i\bigstar})\\
&= O_p\left(\sqrt{T}\right)O_P\left(N^{-1/2}\right)O_p\left(1\right)O_p\left(1\right)O_p\left(\sqrt{N}\right) = O_p\left(\sqrt{T}\right).\end{aligned}$$

For the fourth term, by Lemma 3(i)(iv) and Assumption 3,

$$\begin{aligned}&T\textstyle\sum_{i=1}^{N}\gamma_{1i}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\bar{\Gamma}(T^{-1}\bar{V}_{\bigstar}'M_{Z_{i\bigstar}}Z_{i\Delta})(T^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta})^{-1}\\&\times(T^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\bar{V}_{\bigstar})\bar{\Gamma}'(\bar{\Gamma}\bar{\Gamma}')^{-1}\gamma_{1i}\\=&TO_p(N)\,O_P\left(N^{-1/2}\right)O_p(1)\,O_P\left(N^{-1/2}\right)=O_p(T).\end{aligned}$$

Thus,

$$\sum_{i=1}^{N}\varepsilon_{i\bigstar}^{*\prime}M_{Z_{i\bigstar}}Z_{i\Delta}\left(Z_{i\Delta}'M_{Z_{i\bigstar}}Z_{i\Delta}\right)^{-1}Z_{i\Delta}'M_{Z_{i\bigstar}}\varepsilon_{i\bigstar}^{*}=O_p(N)+O_p\left(\sqrt{T}\right)+O_p\left(\sqrt{T}\right)+O_p(T)$$

(iii) can be proved in the same way as (ii).

Lemma 5. Under Assumptions 1-5, 7, 8, and uniformly over $K(C_k)$ and for each $i=1,...,N$, as $(N,T)\to\infty$,

(i) $||\frac{1}{T}\bar{V}'(k_1^0,k_2^0)M_{\underline{X}(k_1^0,k_2^0)}\bar{V}(k_1^0,k_2^0)||=O_p(N^{-1})$, $||\frac{1}{T}\underline{V}_i'(k_1^0,k_2^0)M_{\underline{X}(k_1^0,k_2^0)}\underline{V}_i(k_1^0,k_2^0)||=O_p(1)$

(ii)$||\frac{1}{T}\mathbb{F}(k_1^0,k_2^0)'M_{\underline{X}(k_1^0,k_2^0)}\mathbb{F}(k_1^0,k_2^0)||=O_p\left(N^{-1}\right)$, $||\frac{1}{T}\underline{V}_i'(k_1^0,k_2^0)M_{\underline{X}(k_1^0,k_2^0)}\mathbb{F}(k_1^0,k_2^0)||=O_p(N^{-1/2})$;

(iii)(iii) $||\frac{1}{T}\bar{V}(k_1^0,k_2^0)'\varepsilon_i||=O_p(\frac{1}{N})+O_p(\frac{1}{\sqrt{NT}})$, $||\frac{1}{T}\bar{V}(k_1^0,k_2^0)'\mathbb{F}^0(k_1^0,k_2^0)||=\frac{1}{\sqrt{N}}$;

(iv) $||\frac{1}{T}\mathbb{F}^{0\prime}(k_1^0,k_2^0)\varepsilon_i||=O_p(1)$.

Proof of Lemma 5. (i) Since $M_{\underline{X}(k_1^0,k_2^0)}=M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)}$,

$$\begin{aligned}\frac{1}{T}\bar{V}'(k_1^0,k_2^0)M_{\underline{X}(k_1^0,k_2^0)}\bar{V}(k_1^0,k_2^0)=&\frac{1}{T}\bar{V}'(k_1^0,k_2^0)M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)}\bar{V}(k_1^0,k_2^0)\\=&\frac{1}{T}\bar{V}'(k_1^0,k_2^0)M_{\mathbb{F}^0(k_1^0,k_2^0)}\bar{V}(k_1^0,k_2^0)\\&+\frac{1}{T}\bar{V}'(k_1^0,k_2^0)\left[M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)}-M_{\mathbb{F}^0(k_1^0,k_2^0)}\right]\bar{V}(k_1^0,k_2^0).\end{aligned}$$

We first analysis the first term,

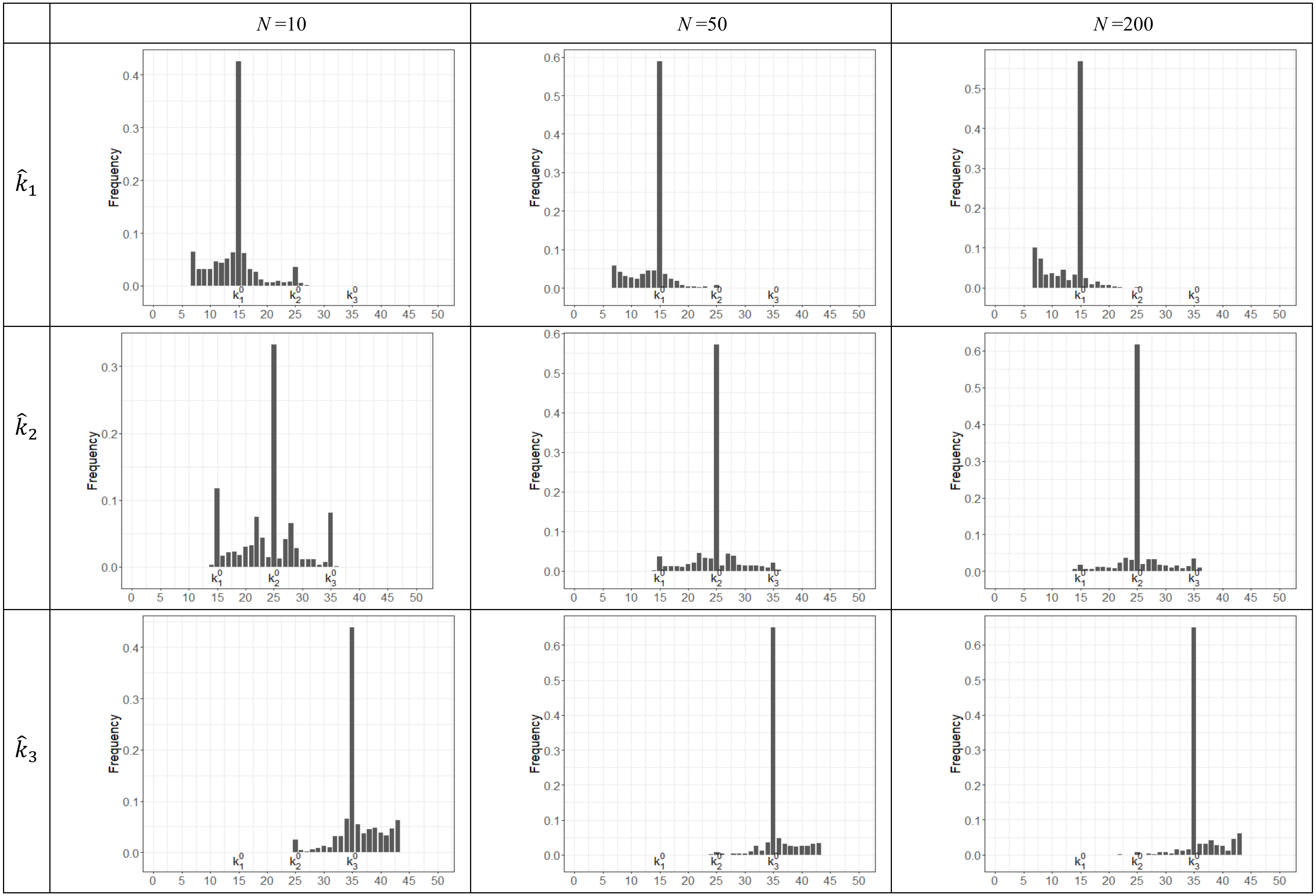
$$\begin{aligned}&||\frac{1}{T}\bar{V}'(k_1^0,k_2^0)M_{\mathbb{F}^0(k_1^0,k_2^0)}\bar{V}(k_1^0,k_2^0)||\\\le&||\frac{1}{T}\bar{V}'(k_1^0,k_2^0)\bar{V}(k_1^0,k_2^0)||\\+&||\frac{1}{T^2}\bar{V}'(k_1^0,k_2^0)\mathbb{F}^0(k_1^0,k_2^0)\left[\frac{1}{T^2}\mathbb{F}^0(k_1^0,k_2^0)'\mathbb{F}^0(k_1^0,k_2^0)\right]^{+}\frac{1}{T}\mathbb{F}^0(k_1^0,k_2^0)'\bar{V}(k_1^0,k_2^0)||\\&=O_p(\frac{1}{N})+O_p(\frac{1}{T})O_p(\frac{1}{N})=O_p(\frac{1}{N}).\end{aligned}$$

Next, for the second term, we can follow the P.12 and equation (S32) of the Appendix in Karabiyik et al.(2017) to show that

$$
\begin{aligned}
||\frac{1}{T}\bar{V}'(k_1^0,k_2^0)\left[M_{\hat{\mathbb{F}}^0(k_1^0,k_2^0)}-M_{\mathbb{F}^0(k_1^0,k_2^0)}\right]\bar{V}(k_1^0,k_2^0)|| &= ||\frac{1}{T}\bar{V}'(k_1^0,k_2^0)P_{\bar{V}^0_{-(m_1+1)q}}\bar{V}(k_1^0,k_2^0|| \\
&= O_P(\frac{1}{N})+O_p(N^{-3/2})+O_p(\frac{1}{\sqrt{T}N}) \\
&= O_P(\frac{1}{N}).
\end{aligned}
$$

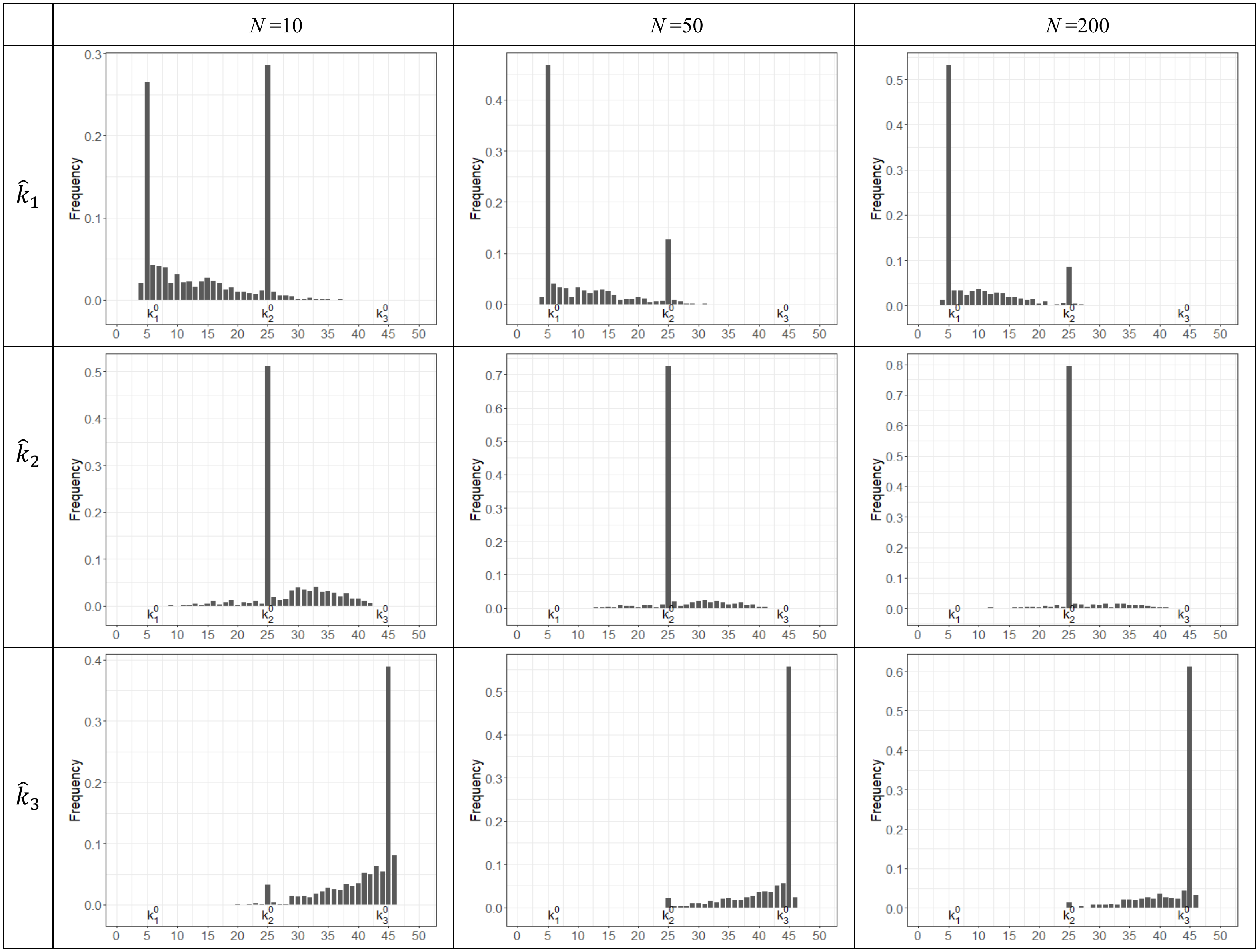
Thus, we obtain (i) and (ii) can be proved similarly. (iii) and (iv) are obvious and then omitted.

Figure 7: Histograms of Break Point Estimators in Case 1 with Nonstationary Factors ($T = 50$)



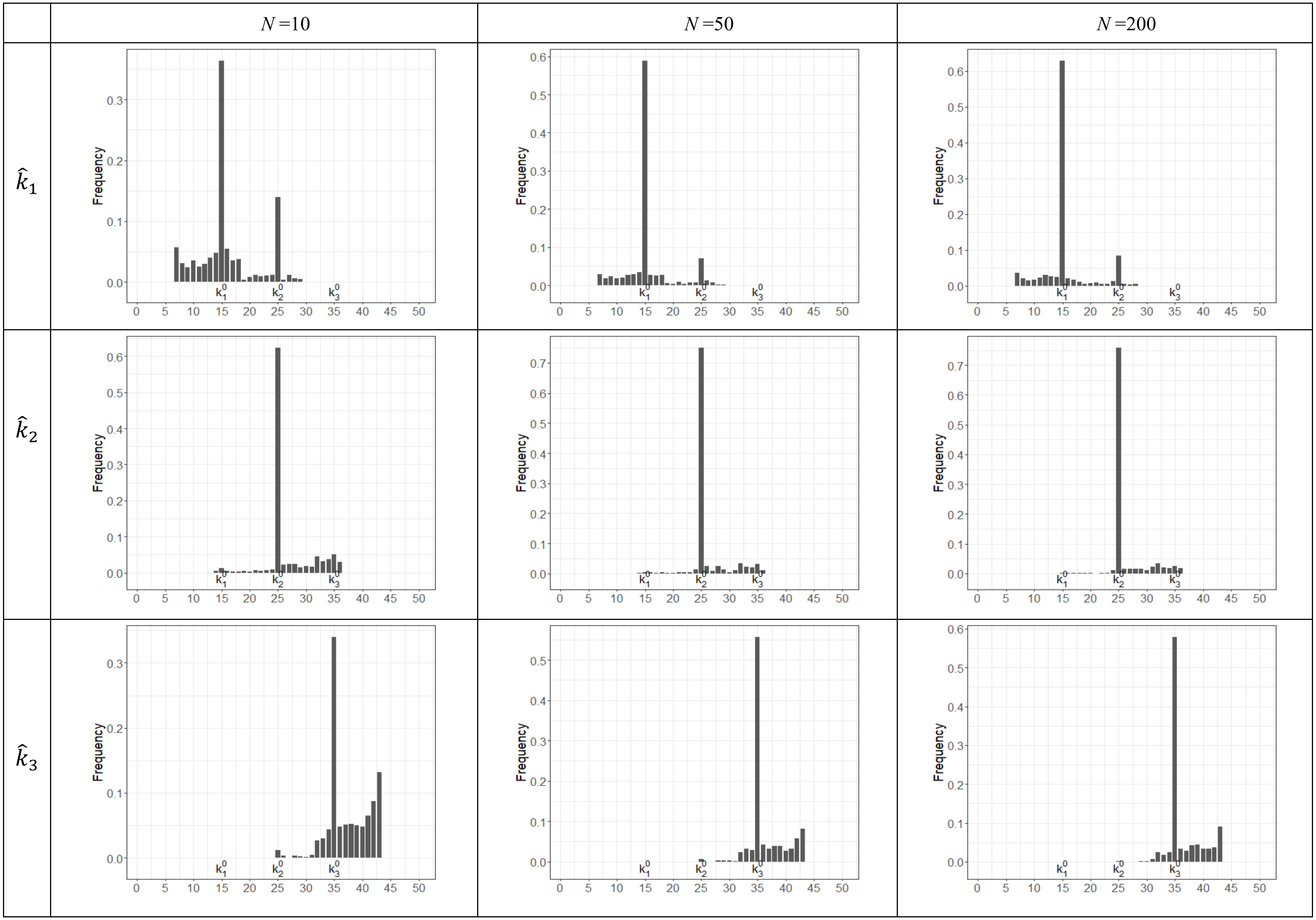
Note: The DGP is the same as that of Figure 1, except that we use (y_{it}, x_{it}) proxy common factors, as Pesaran (2006), KPY and others.

Figure 8: Histograms of Break Point Estimators in Case 1 with Nonstationary Factors and Boundary Breaks($T = 50$)



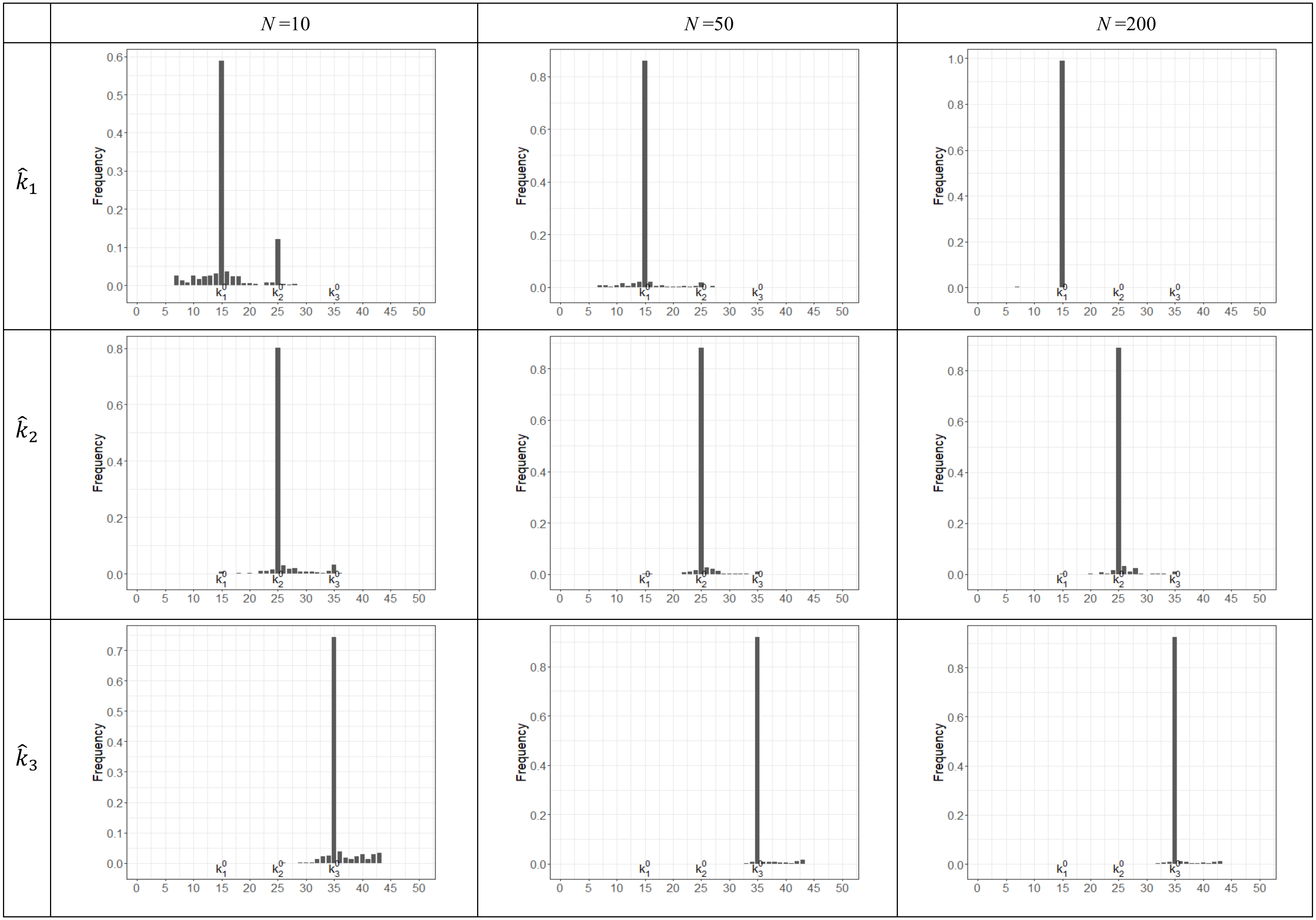
Note: The DGP is the same as that of Figure 1, except that the structural breaks k_1^0 and k_3^0 lie in the boundary, $k_1^0 = 0.1T, k_2^0 = 0.5T, k_3^0 = 0.9T$.

Figure 9: Histograms of Break Point Estimators in Case 1 with Individual Fixed Effects



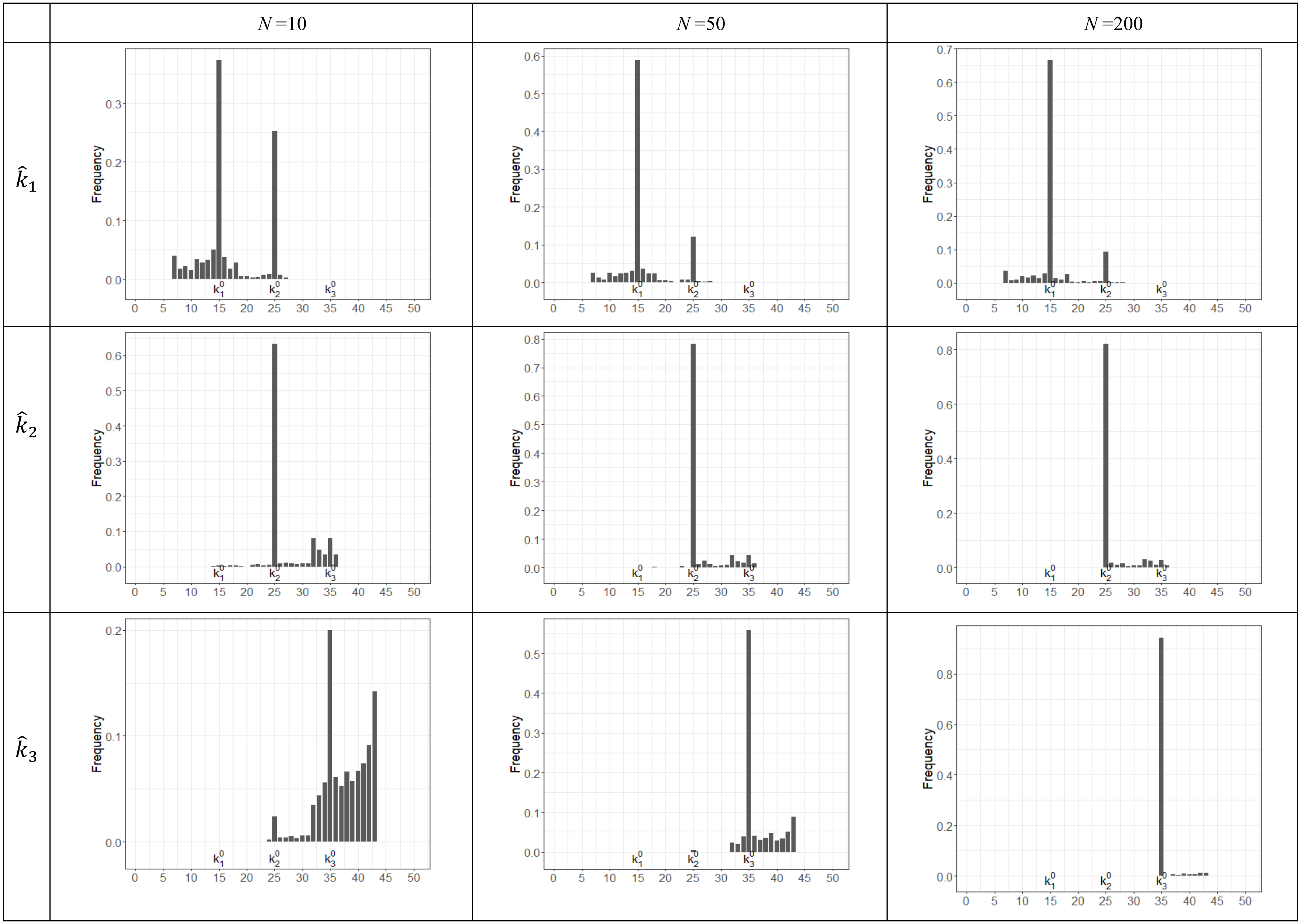
Note: The DGP is the same as that of Figure 1, except that α_i is the individual effects and $x_{1,it} = a_i + \alpha_i + \gamma_{2,i} f_t + v_{it}$.

Figure 10: Histograms of Break Point Estimators in Case 1 with Bigger Size of Breaks in Slopes



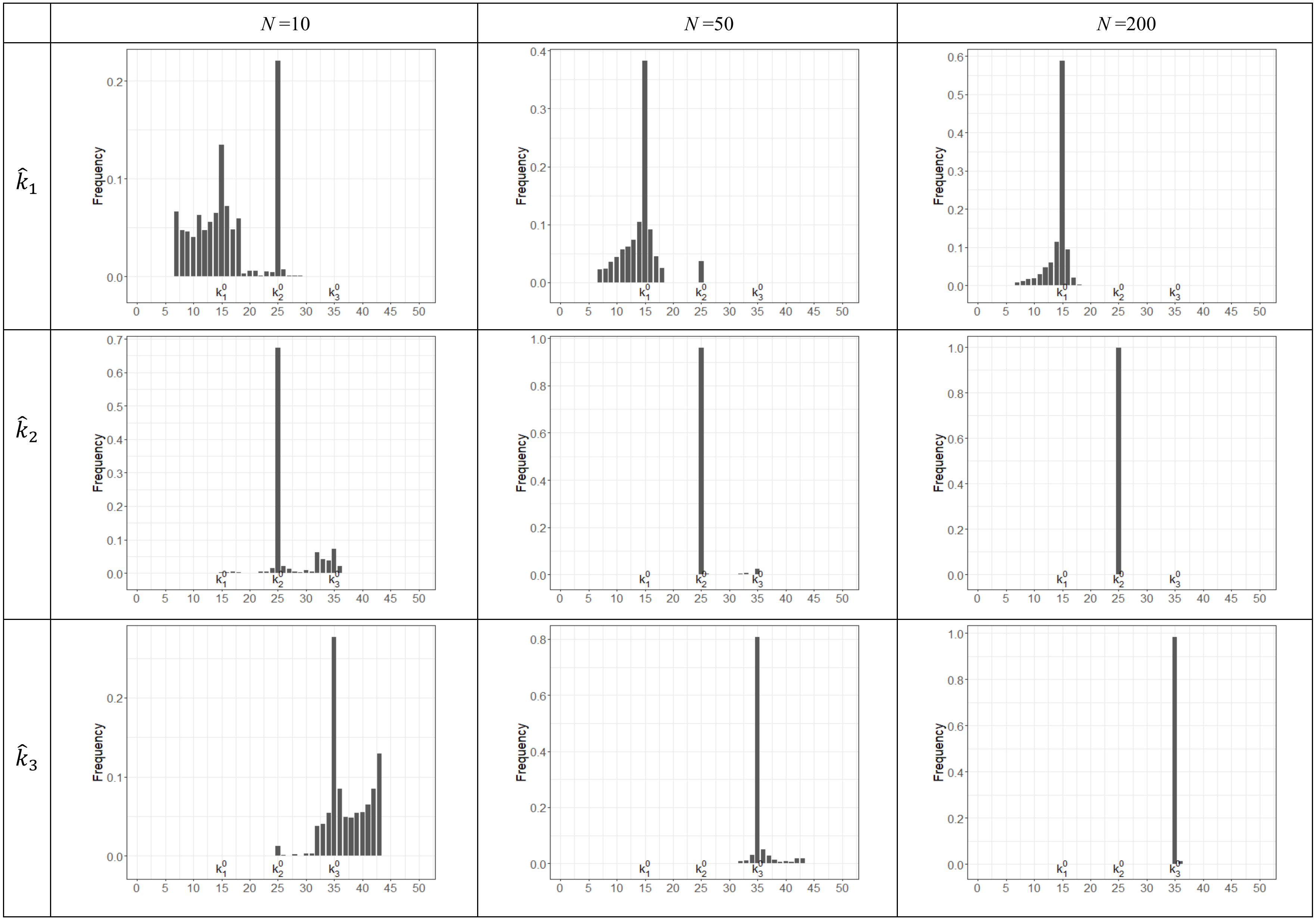
Note: The DGP is the same as that of Figure 1, except that except that $\Delta\beta_i \sim iidN(0,1.2), \Delta\gamma_i \sim iidN(0.5,0.5)$.

Figure 11: Histograms of Break Point Estimators in Case 1 with Bigger Size of Breaks in Loadings



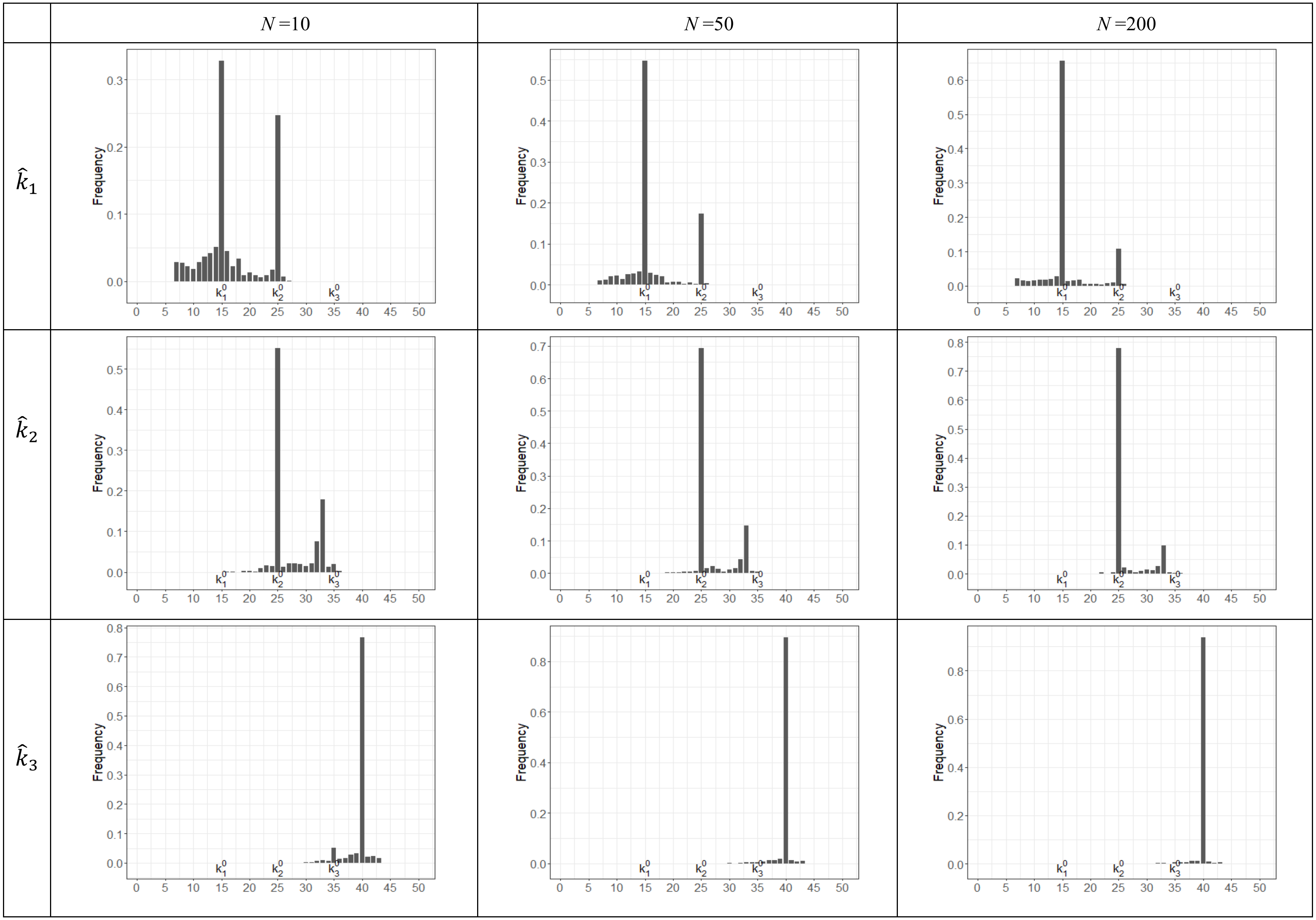
Note: The DGP is the same as that of Figure 1, except that $\Delta\beta_i \sim iidN(0, 0.5), \Delta\gamma_i \sim iidN(0.5, 1)$.

Figure 12: Histograms of Break Point Estimators in Case 1 with trend



Note: The DGP is the same as that of Figure 1, except that the common factor has a time trend $F_t = 0.15t + v_{ft}$.

Figure 13: Histograms of Break Point Estimators in Case 1 with Additional Break Point in Γ_i



Note: The DGP is the same as that of Figure 1, except that let $k_4^0 = 0.8T$, $\gamma_{2,it}\left(k_4^0\right) = \begin{cases} \gamma_{2,i}, & t = 1, \ldots, k_4^0 \\ \gamma_{2,i} + \Delta\gamma_i, & t = k_4^0 + 1, \ldots, T, \end{cases}$ with $\gamma_{2,i} \sim iidN(0.5, 0.5), \Delta\gamma_i \sim iidN(0.5, 0.5)$.