Working Paper Series
Department of Economics
University of Verona

Inference in a similarity-based spatial autoregressive model

Offer Lieberman, Francesca Rossi

WP Number: 1 January 2022

ISSN: 2036-2919 (paper), 2036-4679 (online)

Inference in a similarity-based spatial autoregressive model

Offer Lieberman* Francesca Rossi†

January 13, 2022

Abstract

In this paper we develop asymptotic theory for a similarity-based spatial autoregressive (SAR) model. The model is hybrid in the terminology of Gilboa *et al.* (2006), with the data generating process for a dependent variable y_i containing a *rule-based* linear component, such as $\beta_0' z_i$ for some exogenous observables z_i, and a *case-based* term with a similarity structure. The weight of the similarity structure is allowed to vary in the unit interval and to be estimated explicitly. We prove consistency of the quasi-maximum-likelihood estimator and derive its limit distribution. This paper contributes to the literature on SAR and empirical similarity by incorporating a regression-type component in the data generating process, by allowing the similarity structure to accommodate non-ordered data and by estimating explicitly the weight of the similarity, allowing it to be equal to unity. The model we consider is formally similar to a standard SAR model with exogenous regressors and a data-driven weight matrix which depends on a finite set of parameters that have to be estimated. Our setup accommodates strong forms of cross-sectional correlation that are normally ruled out in the standard literature on spatial autoregressions, and also includes as special cases the random walk with a drift model, the local to unit root model (LUR) with a drift and the model for moderate integration with a drift.

Keywords: Spatial Autoregression; Similarity Function; Weight Matrix; Quasi-Maximum-Likelihood.

Paper's JEL Classification: C21, C22; *Francesca Rossi's JEL Classification:* C13, C21; *Offer Lieberman's JEL Classification:* C13, C22.

1 Introduction

We consider the model

$$y_1 = \beta_0' z_1 + \varepsilon_1, \tag{1.1}$$

*Bar-Ilan University. Support from Israel Science Foundation grant No. 1443-20 is gratefully acknowledged. Correspondence to: Department of Economics, Bar-Ilan University, Ramat Gan 52900, Israel. E-mail: offer.lieberman@gmail.com

†University of Verona . Correspondence to: Department of Economics, University of Verona, Via Cantarane 24, 37129, Verona, Italy . E-mail: francesca.rossi_02@univr.it.

$$y_i = \beta_0' z_i + \lambda_0 \sum_{\substack{j=1 \\ j \neq i}}^{n} h_{i,j} y_j + \varepsilon_i,\ i = 2, ..., n, \tag{1.2}$$

where ε_i, $i = 1, ..., n$, are i.i.d. random variables with zero mean and variance σ_0^2, z_i is the transpose of the $i-$th row of an exogenous $n \times m$ matrix Z of standard covariates, which may include a column of ones,

$$h_{i,j} = \frac{s\left(x_i, x_j; w_0\right)}{\sum_{j \neq i} s\left(x_i, x_j; w_0\right)}, \tag{1.3}$$

with $s\left(x_i, x_j; w_0\right)$ being a similarity function which belongs to $\mathbb{R}_+$ and x_i, x_j being the transpose of $i - th$ and $j-$th rows, respectively, of an $n \times k$ matrix X of fixed explanatory variables. The model is spatial except that, unlike the way it is formulated in the vast literature, the weights, $h_{i,j}$, are driven by some explanatory variables and parameterized by w_0 that needs to be estimated alongside the other parameters of the model. Moreover, the weights are similarity based.

Examples of well-defined similarity functions are given by the exponential and inverse similarity functions, viz.,

$$s\left(x_i, x_j; w_0\right) = \exp\left(-\sum_{t=1}^{k} w_{0t}\left(x_{it} - x_{jt}\right)^2\right) \tag{1.4}$$

and

$$s\left(x_i, x_j; w_0\right) = \frac{1}{1 + \sum_{t=1}^{k} w_{0t}\left(x_{it} - x_{jt}\right)^2}, \tag{1.5}$$

respectively. In both formulations as well as in others, the closer are the ith and the jth cases, through the x_i and x_j values, the larger will be the the value of h_{ij} and as a consequence, the larger will be the weight assigned to y_j in (1.2). It is a similarity model in this sense then - more similar cases result in larger weights attributed to y_j in the construction of y_i. In contrast, in most of the literature on spatial autoregression the weights are determined *a priori* and are fixed.

The unknown parameters of the full model in (1.2) are the scalar $\lambda_0 \in [-1, 1]$, the $k \times 1$ vector $w_0 = (w_{10}, ..., w_{k0})'$, which is assumed to belong to a subset of $\mathbb{R}_+^k$, the $m \times 1$ vector β_0 and σ_0^2, assumed to belong to suitable subsets of $\mathbb{R}_+$ and $\mathbb{R}^m$, respectively. Note that the possibility that $\lambda_0 = 1$ is not negated. The "initial" condition in (1.1) is analogous to the requirement that a process starts from the origin in the time series literature, when $\beta_0 = 0$.

The model (1.2) contains two parts. In the literature on similarity based modeling, originally axiomatized by Gilboa *et. al.* (2006), the model is hybrid, with a 'rule based' component, $\beta_0' z_i$, and a 'case-based' counterpart, $\sum_{j=1, j \neq i}^{n} h_{i,j} y$. When $\lambda_0 = 1$ and $\beta_0 = 0$ *a priori*, model (1.2) represents an extension to the spatial setting of the similarity process, whose asymptotic properties have been established in Lieberman (2010), in the case where the data is ordered, so that the sum in (1.2) extends over $j < i$.

The literature on models such as (1.2) has propagated along two separate paths over the years. The much larger body of literature on SAR modeling includes, just to mention a few contributions, Lee (2004), who established asymptotic theory for the (quasi-) maximum likelihood estimator (QMLE, henceforth), two-stage least squares theory, by Kelejian and Prucha (1998), generalized method of moments theory, by Kelejian and Prucha (1999), higher order SAR, by Gupta and Robinson (2015, 2018), and many more. Most of the theoretical work on standard SAR models rely on a conventional number of technical assumptions, including the spatial parameter lying typically in $(-1, 1)$ or, equivalently, in the interior of a compact subset that depends on the eigenvalues of the weight matrix (e.g. Kelejian and Prucha (2010)), and a suitably normalized weight matrix which is known *a priori*. Also, although several definitions of weak/strong spatial dependence are given in the literature (e.g. Robinson (2011), Chudik and Pesaran (2015) and Bailey et al. (2016)), standard SAR assumptions imply that the largest eigenvalue of the variance-covariance matrix of the dependent variable is bounded, such that every form of strong dependence is automatically ruled out. Related to the purpose of this project, Lee and Yu (2013) offered some insight on asymptotic theory for QMLE in SAR models with a spatial parameter that is local-to unity, under the condition that the weight matrix is diagonalizable, which rules out the LUR model of Phillips (1987) and Chan and Wei (1987). In this line of literature, Baltagi et. al. (2013) derived asymptotic theory for ordinary least squares and generalized least squares estimators for a cross-sectional model with SAR errors with spatial parameter that tends to unity as sample size increases.

On the other hand, the literature on similarity based models, include, *inter alia*, Gilboa *et. al.* (2010, 2011), Gayer *et. al.* (2007), Lieberman (2012), Lieberman and Phillips (2014) Gayer *et. al.* (2019), Kapetanios *et. al.* (2013), and Teitelbaum (2013). Recently, Rossi and Lieberman (2021, henceforth, RL) made the first attempt to bridge the two streams of literature, when they considered a special case of (1.2) with $\beta_0 = 0$. The more general setup with $\beta_0 \neq 0$ corresponds to a hybrid model that includes a rule-based component, as discussed above, it poses some interesting technical challenges and the results in this case are very different from the $\beta_0 = 0$ case.

In this paper we focus on developing the asymptotic theory for inference on

$$\theta_0 = (\beta_0', \sigma_0^2, \lambda_0, w_{10}, ..., w_{k0})'$$

in model (1.2). The setup is sufficiently general to include as special cases the random walk with a drift model, the local to unit root model (Chan and Wei (1987), Phillips (1987), henceforth, LUR), moderate deviations from a unit root model (Phillips and Magdalinos, (2007), henceforth, MI), and standard SAR models, as in Lee (2004). As the norming rates for the asymptotic theory are very different across the special cases, we employ random norming that treats all scenarios in a uniform manner. For instance, our random norming collapses to the well known $n^{3/2}$-rate for the QMLE of λ_0

in the random walk with a drift model (see, for instance, Hamilton (1994, equation (17.4.47)).

The plan for the rest of the paper is as follows. In Section 2 we provide the setup, assumptions, and identification and consistency of the model parameters. The limit distribution follows in Section 3 and discussion follows in Section 4. Concluding remarks are given in Section 5. Supplementary lemmas and all proofs are provided in the Appendix.

2 Setup, Assumptions, Identification and Consistency

For any generic $p \times q$ matrix A, we denote by a_{ij} its $(i,j)-$th element and by a_i the transpose of its $i-$th row. Also b^{ij} denotes the $(i,j)-$th elements of B^{-1} for any generic, square, invertible matrix B. Furthermore, $||\cdot||$, $||\cdot||_\infty$, and $||\cdot||_F$ represent spectral, uniform absolute row sum and Frobenius norms, respectively, A' is the transpose of A, and $K > 0$ is an arbitrary finite constant whose value may change in each location. For a generic square matrix, $\eta_{\min}(B)$ and $\eta_{\max}(B)$ denote minimum and maximum eigenvalues of B, respectively, while $|B|$ indicates the determinant of B. Throughout, the subscript $(\cdot)_0$ indicates true values, or quantities evaluated at the true parameters' values, while the absence of such subscript denotes parameters that are free to vary within the parameters' space or quantities evaluated at generic values of the parameters.

Model (1.2) can be written in matrix form as

$$S_{n0} y_n = Z_n \beta_0 + \varepsilon_n, \tag{2.1}$$

where

$$\begin{aligned} S_{n0} &= S_n(\lambda_0, w_0) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\lambda_0 h_{2,1} & 1 & \cdots & -\lambda_0 h_{2,n} \\ \cdots & & \cdots & \\ -\lambda_0 h_{n,1} & -\lambda_0 h_{n,2} & \cdots & 1 \end{pmatrix} \\ &= I - \lambda_0 C_n(\lambda_0, w_0) = I - \lambda_0 C_{n0}. \end{aligned} \tag{2.2}$$

In (2.1), as well as in (1.2), $y = y_n$, $\varepsilon = \varepsilon_n$, $X = X_n$, $Z = Z_n$, $C_0 = C_{n0}$ and $S_0 = S_{n0}$ are, in general, triangular arrays, but we omit the subscript n in the sequel for brevity. This means, in particular, that $h_{i,j} = h_{i,j,n}$, for $i, j = 1, ..., n$.

The reduced form of the model (2.1) is

$$y = S_0^{-1} \left(Z\beta_0 + \varepsilon \right), \tag{2.3}$$

provided that S_0^{-1} exists. For $|\lambda_0| < 1$ and for given w_0, under the well-known condition known as “weak dependence”, e.g. Kelejian and Prucha (1998),

$$\sup_{\Theta}\left(||S^{-1}||_\infty + ||S^{-1\prime}||_\infty\right) < K, \tag{2.4}$$

model (2.1) formally corresponds to a SAR model with exogenous regressors, and the theory for developing inference on λ_0 is well established under some suitable additional conditions.

We introduce the following Assumptions.

Assumption 1 *For all n and for $i = 1,\dots,n$, the $\{\epsilon_i\}$ are a set of independent random variables, with mean zero and unknown variances $\sigma^2 > 0$. In addition, for some $\delta > 0$,*

$$\mathbb{E}|\epsilon_i|^{4+\delta} \leq K \quad for \quad i = 1, ..., n.$$

Assumption 2 *There exists $\sigma_L^2 > 0$, $\sigma_H^2 < \infty$ and $w_H < \infty$ such that $\sigma_L^2 < \sigma_0^2 < \sigma_H^2$ and, for all $i = 1, ..., k$, $0 \leq w_{i0} < w_H$. Also, $-1 \leq \lambda_0 \leq 1$ and $\lambda_0 \neq 0$*[1]*.*

Assumption 3 *The matrix X is allowed to lie in the set of all $n \times k$ non-random, real matrices such that for all sufficiently large n*

$$S'S \neq S_0'S_0 \quad for \quad \theta \neq \theta_0. \tag{2.5}$$

Assumption 4 *For all n, S_0 is non singular and $0 < |(S'S)^{-1}| < K$ for all $\theta \in \Theta$.*

Assumption 5 *For all n, $S'S$ has bounded and continuous derivatives, uniformly in $\theta_2 \in \Theta_2$.*

Let $C_r = C_r(w_1, ..., w_k) = \frac{\partial C(w_1,....,w_k)}{\partial w_r}$ for $r = 1, ..., k$.

Assumption 6

a) $\sup_{\theta\in\Theta}(||C(\theta)||_\infty + ||C'(\theta)||_\infty) \leq K.$

b) $\sup_{\theta\in\Theta}(||C_r(\theta)||_\infty + ||C_r'(\theta)||_\infty) \leq K$

Assumptions 1-6 have been discussed extensively in RL in the context of a simpler model that does not include $\beta_0' z_i$ in (1.2). In particular, the first part of Assumption 4 guarantees that the reduced form in (2.3), while the second part ensures that the log-likelihood function remains well defined for all $\theta \in \Theta$. In the following we impose a condition on the covariates Z.

Assumption 7 *For all n, each element z_{ij} of Z $(n \times m)$ is non-random and $|z_{ij}| < K$. Also, for all*

[1] In case $\lambda_0 = 0$ we are not able to identify $w_{10}, ..., w_{k0}$.

sufficiently large n,

$$0 < c < \eta_{\min}\left(\frac{Z'Z}{n}\right), \tag{2.6}$$

where c is any arbitrarily small constant.

Assumption 7 could be relaxed to strictly exogenous z_{ij} with fairly minor modifications. We also impose an asymptotic no-collinearity condition similar to that of, e.g., Lee (2004).

Assumption 8

$$\lim_{n\to\infty}\frac{1}{n||S_0^{-1}||_\infty^2}\beta_0'Z'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}Z\beta_0 > 0,$$
$$\lim_{n\to\infty}\frac{1}{n||S_0^{-1}||_\infty^2}\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0 > 0 \quad \textit{for } S \neq \pm S_0,$$
$$\lim_{n\to\infty}\frac{1}{n||S_0^{-1}||_\infty^2}\beta_0'Z'S_0^{-1\prime}S_0^{-1}Z\beta_0 > 0. \tag{2.7}$$

By Lemmas 1 and 4(a), $\beta_0'Z'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}Z\beta_0 = O(n||S_0^{-1}||_\infty^2)$, $\beta_0'Z'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}Z\beta_0 = O(n||S_0^{-1}||_\infty^2)$ and $\beta_0'Z'S_0^{-1\prime}S_0^{-1}Z\beta_0 = O(n||S_0^{-1}||_\infty^2)$ and are non-negative. Our Assumption 8 is similar to Assumption 8 of Lee (2004) and implies that the aforementioned rates are exact.

As in RL, we aim to consistently estimate θ via a quasi-maximum-likelihood (QML) function that allows us to accommodate

$$\sup_\Theta ||S^{-1}||_\infty = O(n^\gamma), \quad \gamma \in [0,1] \tag{2.8}$$

within a unified framework. The case $\gamma = 0$ corresponds to the standard SAR setup, while in case $\gamma > 0$, the condition in (2.4) does not hold and standard limit theory for SAR models is not available. We furthermore assume $||S^{-1\prime}||_\infty = O(||S^{-1}||_\infty)$ such that, in case $= (||S^{-1}||_\infty) = O(n^\gamma)$ with $\gamma > 0$, $||S^{-1\prime}||_\infty$ could be bounded or increasing without bound. By allowing $\gamma > 0$ we relax the standard assumption of weak dependence across y and we are also allowing y_i, for $i = 1, ..., n$, to have a variance that increases with sample size, as in unit root models, since it is straightforward to see that $Var(y_i) = O(||S^{-1}||_\infty)$.

Let $\theta = (\beta', \sigma^2, \lambda, w')' = (\theta_1', \theta_2')'$, with $\theta_1 = \left(\beta', \sigma^2\right)'$ and $\theta_2 = (\lambda, w')'$. Given y, and letting $S = S(\theta_2)$ we define the shifted, normalized and negative pseudo-log-likelihood function as

$$\mathcal{L}(\theta) = \log\left(\sigma^2\right) - \frac{2}{n}\log|S| + \frac{(Sy - Z\beta)'(Sy - Z\beta)}{n\sigma^2} - \log\left(\frac{y'y}{n}\right) \tag{2.9}$$

and $\hat\theta = \arg\min_{\theta\in\Theta}\mathcal{L}(\theta)$. The shifting term $-\log(y'y/n)$ is introduced to allow us to accommodate both $\gamma = 0$ and $\gamma > 0$ cases, without affecting $\arg\min \mathcal{L}(\theta)$, where γ is defined in (2.8).

Given θ_2, we obtain

$$\hat\beta(\theta_2) = \hat\beta = \left(Z'Z\right)^{-1}Z'Sy \tag{2.10}$$

and as

$$\left(Sy - Z\hat{\beta}\right)'\left(Sy - Z\hat{\beta}\right) = y'S'M_ZSy,$$

with

$$M_Z = I - Z\left(Z'Z\right)^{-1}Z',$$

we have

$$\hat{\sigma}^2 = \hat{\sigma}^2\left(\theta_2\right) = \frac{y'S'M_ZSy}{n}. \tag{2.11}$$

We remark that $\hat{\sigma}^2$ is the estimator used in equation (2.6) of Lee (2004). Let

$$\hat{\sigma}^{*2} = \hat{\sigma}^{*2}\left(\theta_2\right) = \frac{y'S'M_ZSy}{y'y}. \tag{2.12}$$

Plugging (2.10) and (2.11) into (2.9), the profile, shifted, quasi-log-likelihood is equal to

$$\mathcal{L}^p(\theta_2) = \log\left(\hat{\sigma}^2\right) - \frac{2}{n}\log|S| + \frac{y'S'M_ZSy}{n\hat{\sigma}^2} - \log\left(\frac{y'y}{n}\right), \tag{2.13}$$

which, up to constant terms becomes

$$\mathcal{L}^p(\theta_2) = \log\left(\frac{y'S'M_ZSy}{y'y}\right) - \frac{2}{n}\log|S| = \log\left(\hat{\sigma}^{*2}\right) - \frac{2}{n}\log|S|. \tag{2.14}$$

The QML estimator of θ_{20} is defined to be $\hat{\theta}_2 = \underset{\theta_2\in\Theta_2}{\arg\min}\mathcal{L}^p(\theta_2)$.

From (2.3), the numerator and denominator of (2.12) can be written respectively as

$$y'S'M_ZSy = \epsilon'S_0^{-1\prime}S'M_ZSS_0^{-1}\epsilon + \beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0 + 2\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}\epsilon \tag{2.15}$$

and

$$y'y = \epsilon'S_0^{-1\prime}S_0^{-1}\epsilon + \beta_0'Z'S_0^{-1\prime}S_0^{-1}Z\beta_0 + 2\beta_0'Z'S_0^{-1\prime}S_0^{-1}\epsilon. \tag{2.16}$$

From Lemma 1, $\|S'M_ZS\|_\infty < K$ and thus, by Lemma 2(b) the first term on the rhs of (2.15) is $O_p\left(n\left\|S^{-1}\right\|_\infty\right)$ and by Lemma 4 the second and third terms on the rhs of (2.15) are $O_p\left(n\left\|S^{-1}\right\|_\infty^2\right)$ and $O_p\left(\sqrt{n}\left\|S^{-1}\right\|_\infty^2\right)$, respectively. Therefore, (2.15) becomes

$$y'S'M_ZSy = \epsilon'S_0^{-1\prime}S'M_ZSS_0^{-1}\epsilon + \beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0 + O_p\left(\sqrt{n}\right) = O_p\left(n\right), \text{ if } \gamma = 0 \tag{2.17}$$

and if $0 < \gamma \le 1$,

$$y'S'M_ZSy = \beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0 + O_p\left(\max\left(n\left\|S^{-1}\right\|_\infty, \sqrt{n}\left\|S^{-1}\right\|_\infty^2\right)\right) = O_p\left(n\left\|S^{-1}\right\|_\infty^2\right). \tag{2.18}$$

Similarly, by Lemmas 2(b) and 4, (2.16) satisfy

$$y'y = \epsilon' S_0^{-1\prime} S_0^{-1}\epsilon + \beta_0' Z' S_0^{-1\prime} S_0^{-1} Z\beta_0 + O_p\left(\sqrt{n}\right) = O_p\left(n\right), \text{ if } \gamma = 0 \tag{2.19}$$

and

$$y'y = \beta_0' Z' S_0^{-1\prime} S_0^{-1} Z\beta_0 + O_p\left(\max\left(n\left\|S^{-1}\right\|_\infty, \sqrt{n}\left\|S^{-1}\right\|_\infty^2\right)\right) = O_p\left(n\left\|S^{-1}\right\|_\infty^2\right), \text{ if } 0 < \gamma \leq 1. \tag{2.20}$$

More concisely,

$$y'S'M_Z S y = O_p\left(n\left\|S^{-1}\right\|_\infty^2\right) = O_p(n^{1+2\gamma}) \text{ and } y'y = O_p\left(n\left\|S^{-1}\right\|_\infty^2\right) = O_p(n^{1+2\gamma}),\ \forall \gamma \in [0,1]. \tag{2.21}$$

Moreover, in view of (2.15) and as $M_Z Z = 0$,

$$\begin{aligned} y'S_0' M_Z S_0 y &= \epsilon' S_0^{-1\prime} S_0' M_Z S_0 S_0^{-1}\epsilon + \beta_0' Z' S_0^{-1\prime} S_0' M_Z S_0 S_0^{-1} Z\beta_0 + 2\beta_0' Z' S_0^{-1\prime} S_0' M_Z S_0 S_0^{-1}\epsilon \\ &= \epsilon' M_Z \epsilon \end{aligned}$$

and we have,

$$\epsilon' M_Z \epsilon \leq \epsilon'\epsilon \left\|M_Z\right\| = \epsilon'\epsilon = O_p\left(n\right).$$

It follows that

$$y'S_0' M_Z S_0 y = \epsilon' M_Z \epsilon = O_p\left(n\right), \tag{2.22}$$

and it is emphasized that the rate holds in (2.22) for $0 \leq \gamma \leq 1$, whereas for $S \neq S_0$ it follows from (2.21) that $y'S'M_Z S y = O_p\left(n^{1+2\gamma}\right)$.

In view of (2.12) and (2.21), $\hat{\sigma}^{*2}(\theta_2) = O_p\left(1\right)$, $\forall \gamma \in [0,1]$. We further define

$$\tilde{\sigma}^{*2}(\theta_2) = p\lim_{n\to\infty}\left(\hat{\sigma}^{*2}(\theta_2)\right). \tag{2.23}$$

In order to ensure existence of the limit objective function and to be able to establish consistency of $\hat{\theta}_2$, we introduce the following assumption.

Assumption 9

$$\begin{aligned} &\tilde{\sigma}^{*}(\theta_2) = p\lim_{n\to\infty} \hat{\sigma}^{*2}(\theta_2) \quad \textit{exists for all} \quad \theta_2 \in \Theta_2, \\ &p\lim_{n\to\infty}\frac{\partial}{\partial\lambda}\hat{\sigma}^{*2}(\theta_2) \quad \text{and} \quad p\lim_{n\to\infty}\frac{\partial}{\partial w_j}\hat{\sigma}^{*2}(\theta_2), \quad \textit{for } j = 1, ..., k, \textit{ exist for all} \quad \theta_2 \in \Theta_2. \end{aligned} \tag{2.24}$$

We stress that $\tilde{\sigma}^{*}(\theta_2)$ is strictly positive under Assumption 8, while its existence is guaranteed under

Assumption 9.
The limit objective function is given by

$$\tilde{\mathcal{L}}^p(\theta_2) = \log\left(\tilde{\sigma}^{*2}\right) - \frac{2}{n}\log|S| = \log\left(\frac{y'S'M_ZSy}{y'y}\right) - \frac{2}{n}\log|S| + o_p(1), \tag{2.25}$$

with $\theta_{20} = \underset{\theta_2\in\Theta_2}{\arg\min}\tilde{\mathcal{L}}^p(\theta_2)$.

Remark 1 *We emphasize that $\|S'M_ZS\|_\infty < K$ by Lemma 1. Using (2.15), in the $\gamma = 0$ case, by Lemma 3 with the generic matrix A replaced by $S'M_ZS$ and $\left\|S^{-1}\right\|_\infty = O(1)$,*

$$\underset{n\to\infty}{p\lim}\left(\hat{\sigma}^2(\theta_2)\right) = \frac{\sigma_0^2 tr\left(S_0^{-1\prime}S'M_ZSS_0^{-1}\right)}{n} + \frac{\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{n} + o_p(1).$$

Now,

$$\frac{tr\left(S_0^{-1\prime}S'M_ZSS_0^{-1}\right)}{n} = \frac{tr\left(S_0^{-1\prime}S'SS_0^{-1}\right)}{n} - \frac{tr\left(S_0^{-1\prime}S'P_ZSS_0^{-1}\right)}{n},$$

where

$$P_Z = Z\left(Z'Z\right)^{-1}Z'.$$

We notice that

$$tr\left(S_0^{-1\prime}S'SS_0^{-1}\right) \leq Kn$$

but

$$tr\left(S_0^{-1\prime}S'P_ZSS_0^{-1}\right) = \left\|P_ZSS_0^{-1}\right\|_F^2 \leq \|P_Z\|_F^2\left\|SS_0^{-1}\right\|^2 \leq m\left\|SS_0^{-1}\right\|_\infty^2 \leq Km.$$

Hence, in the $\gamma = 0$ case

$$\hat{\sigma}^2(\theta_2) = \frac{\sigma_0^2 tr\left(S_0^{-1\prime}S'SS_0^{-1}\right) + \beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{n} + o_p\left(1\right), \tag{2.26}$$

in line with (3.2) of Lee (2004). The last displayed expression is non-singular under Assumptions 3 and 8.

In Appendix A we shall prove the following.

Theorem 1. *Assume that model (2.1) and Assumptions 1-9 hold. Under the condition (2.8) with $0 \leq \gamma \leq 1$, θ_{20} is identified and $\hat{\theta}_2 \xrightarrow{p} \theta_{20}$.*

It is emphasized that identification and consistency of $\hat{\theta}_{20}$ hold under $0 \leq \gamma \leq 1$, so that the "weak dependence" condition given in (2.4) and used in the literature (e.g., e.g. Kelejian and Prucha (1998)) is not needed. Consistency of $\hat{\beta}$ and of $\hat{\sigma}^2$ follow from (2.10) and (2.11), respectively. Theorem 1, in addition to contributing to SAR literature by relaxing the usual constraint on the parameter space and, even more importantly, by establishing consistent estimation allowing forms of strong dependence across spatial units, extends results in Lieberman (2010) to a bilateral hybrid model where there is no natural ordering of observations and the strength of the similarity structure, embedded in λ_0, can be estimated explicitly similarly to what established in RL in the context of a simpler model.

In the next section we shall derive the asymptotic distribution of $\hat{\theta}_2$. The distribution of $\hat{\theta}$ and $\hat{\theta}_1$ will be deduced from it by standard arguments.

3 Limit Distribution

In Theorem 2 below we will show that central limit theorem holds, with rates depending on $O\left(\left\|S_0^{-1}\right\|_\infty\right)$.

For any matrix A, let $\underline{A} = A + A'$,

$$C_{r,0} = \frac{\partial C(w_1, ..., w_k)}{\partial w_r}|_{\theta_0}, \text{ for } r = 1, ..., k$$

and

$$C_{rs,0} = \frac{\partial^2 C(w_1, ..., w_k)}{\partial w_r \partial w_s}|_{\theta_0}, \text{ for } r, s = 1, ..., k.$$

and a similar notation is used for $C_{rst}(\theta)$.

Assumption 10

a) $\sup_{\theta\in\Theta} (||C_{rs}(\theta)||_\infty + ||C'_{rs}(\theta)||_\infty) \leq K \quad \text{for} \quad r, s = 1, ..., k.$

b) $\sup_{\theta\in\Theta} (||C_{rst}(\theta)||_\infty + ||C'_{rst}(\theta)||_\infty) \leq K \quad \text{for} \quad r, s, t = 1, ..., k.$

Assumption 10 extends Assumption 6 to uniform boundedness in row and column sums of the second- and third-order derivatives of $C(\cdot)$, as in Assumption 9 of RL.

We let

$$V_0 = \Sigma_{10} + \Sigma_{20} + \Sigma_{30} + \Sigma_{40}, \tag{3.1}$$

where Σ_{10}, Σ_{20}, Σ_{30} and Σ_{40} are defined in (A.8), (A.9), (A.10) and (A.11), respectively. Under Assumption 6, the elements of Σ_{20} are $O(1/||S_0^{-1}||_\infty^2)$ by Lemma 4(e) and (B.8), each element of Σ_{10} is $O(1/||S_0^{-1}||_\infty)$ by Lemma 4(f) while each element of Σ_{40} is $O(1/||S_0^{-1}||_\infty)$ by Lemma 4(e) and (B.8)[2]. The elements in Σ_{30} are $O(1)$ by Lemma 4. Also, let D_0 be the $(k+1)\times(k+1)$ matrix with elements given in (A.12), (A.13) and (A.14). For $\gamma = 0$, all elements of D_0 are $O(1)$ from Lemma 4, while

[2]Details are provided in the proof of Theorem 2 in Appendix A.

for $\gamma > 0$, D_0 reduces to $\tilde{D}_0$ with elements defined in (A.15). We stress that elements of D_0 are the probability limits of the elements of the normalized Hessian

$$\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^2 \mathcal{L}(\theta_{20})}{\partial\theta_2\partial\theta_2'}, \tag{3.2}$$

which are well defined from Lemma 9. Finally, let F_0 to be defined as in (A.16).

We introduce an additional condition to ensure that the variance-covariance matrix of the suitably normalized $\hat{\theta}_2$ exists and it is non singular in the limit.

Assumption 11 *The limits in Σ_{10}, Σ_{20}, Σ_{30}, Σ_{40} and D_0 exist. Furthermore $\eta_{\min}(\Sigma_{30}) > 0$ and $\eta_{\min}(\tilde{D}_0) > 0$.*

We establish the following.

Theorem 2. *Assume that model (2.1) and Assumptions 1-11 hold. For each $\gamma \in [0,1]$,*

$$\left(n||S_0^{-1}||_\infty^2\right)^{1/2}(\hat{\theta}_2 - \theta_{20}) \xrightarrow{d} \mathcal{N}(0, \mathcal{V}_0), \tag{3.3}$$

where $\mathcal{V}_0 = D_0^{-1}F_0V_0F_0D_0^{-1}$.

Thus, unlike Theorem 2 in RL, the presence of exogenous regressors allows a unified approach to establish the limit distribution even for the boundary case of $\gamma = 1$ in (2.8). We stress that when $\beta_0 = 0$ *a priori*, Assumption 11 is violated since $\Sigma_{30} = 0$ and Theorem 2 is well defined only for the trivial case with $\gamma = 0$. In the latter case, the limit distribution given in Theorem 2 is identical to that derived in RL. The result of Theorem 2 will be further discussed in the following section with illustrations through some key special cases.

We conclude this section by briefly focussing on inference on β_0, given results in Theorem 2. Let $\bar{\theta}_{2j}$ an intermediate point such that $|\bar{\theta}_{2j} - \theta_{20j}| < |\hat{\theta}_{2j} - \theta_{20j}|$ for $j = 1, ..., k+1$. By the MVT we can write,

$$\begin{aligned}
\hat{\beta} - \beta_0 &= \left(\frac{1}{n}Z'Z\right)^{-1}\frac{1}{n}Z'\epsilon - \left(\frac{1}{n}Z'Z\right)^{-1}\frac{1}{n}Z'\bar{C}y(\hat{\lambda} - \lambda_0) - \bar{\lambda}\sum_{j=1}^{k}\left(\frac{1}{n}Z'Z\right)^{-1}\frac{1}{n}Z'\bar{C}_j y(\hat{w}_j - w_{0j}) \\
&= \left(\frac{1}{n}Z'Z\right)^{-1}\frac{1}{n}Z'\epsilon - \left(\frac{1}{n}Z'Z\right)^{-1}\frac{1}{n}Z'\bar{C}S_0^{-1}Z\beta_0(\hat{\lambda} - \lambda_0) \\
&\quad -\bar{\lambda}\sum_{j=1}^{k}\left(\frac{1}{n}Z'Z\right)^{-1}\frac{1}{n}Z'\bar{C}_jS_0^{-1}Z\beta_0(\hat{w}_j - w_{0j}) + O_p\left(\frac{1}{n}\right),
\end{aligned} \tag{3.4}$$

where the second equality follows from the rates in Theorem 2, Lemma 2 and Lemma 4. The leading terms in (3.4) are of order $O_p(1/\sqrt{n})$, again from standard arguments and Lemma 4. Thus, from (3.4) it is clear that the rate of convergence of $\hat{\beta}$ remains the standard $\sqrt{n}$ for any $0 \le \gamma \le 1$ in (2.8).

In what follows we omit much of the technical details to avoid repetition. We can derive the joint distribution of the suitably normalized $\hat{\theta}$ by writing

$$\begin{pmatrix} \left(n||S_0||_\infty^2\right)^{1/2}(\hat{\lambda}-\lambda_0) \\ \left(n||S_0||_\infty^2\right)^{1/2}(\hat{w}_1-w_{01}) \\ ..... \\ \left(n||S_0||_\infty^2\right)^{1/2}(\hat{w}_k-w_{0k}) \\ n^{1/2}(\hat{\beta}-\beta_0) \end{pmatrix} = R_0 \begin{pmatrix} \frac{\sqrt{n}}{||S_0^{-1}||_\infty}\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} \\ \frac{\sqrt{n}}{||S_0^{-1}||_\infty}\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_1} \\ ..... \\ \frac{\sqrt{n}}{||S_0^{-1}||_\infty}\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_k} \\ \frac{1}{\sqrt{n}}Z'\epsilon \end{pmatrix} + o_p(1),$$

where R_0 is a $(k+m+1)\times(k+m+1)$ matrix given in (A.17).

Let F_0^β and V_0^β be $(k+m+1)\times(k+m+1)$ matrix defined as (A.19) and

$$V_0^\beta = \Sigma_{10}^\beta + \Sigma_{20}^\beta + \Sigma_{30}^\beta + \Sigma_{40}^\beta, \tag{3.5}$$

with Σ_{10}^β, Σ_{20}^β, Σ_{30}^β and Σ_{40}^β reported in (A.20), (A.21), (A.22) and (A.23), respectively.

To complement Assumption 11, we impose the additional

Assumption 12 *The limits of elements of R_0, Σ_{30}^β and Σ_{40}^β exist. Furthermore $\eta_{\min}(\Sigma_{30}^\beta) > 0$.*

Similarly to Theorem 2, we can prove

Theorem 3. *Assume that model (2.1) and Assumptions 1-12 hold. For each $\gamma \in [0,1]$,*

$$\begin{pmatrix} \left(n||S_0||_\infty^2\right)^{1/2}(\hat{\lambda}-\lambda_0) \\ \left(n||S_0||_\infty^2\right)^{1/2}(\hat{w}_1-w_{01}) \\ ..... \\ \left(n||S_0||_\infty^2\right)^{1/2}(\hat{w}_k-w_{0k}) \\ n^{1/2}(\hat{\beta}-\beta_0) \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, \mathcal{V}_0^\beta), \tag{3.6}$$

where $\mathcal{V}_0^\beta = R_0 F_0^\beta V_0^\beta F_0^\beta R_0'$.

4 Discussion

The rate of convergence in Theorem 2 collapses to $\sqrt{n}$ in standard SAR models in which $|\lambda_0| < 1$ and $||S_0^{-1}||_\infty = O(1)$, and it agrees with that derived in Lee (2004). Theorem 2 represents a novel contribution to the SAR literature since we derive the asymptotic distribution of QML estimators without the usual requirements on admissible values for λ_0 and by allowing forms of strong cross sectional dependence.

We can further discuss the generality of our results in view of the time series literature. In the random walk with a drift model, the convergence rate is $n^{3/2}$, because in this case

$$C_0 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

from which it follows that

$$S_0^{-1} = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ 1 & 1 & 0 & \cdots & \\ 1 & 1 & 1 & & \cdots \\ & \cdots & & & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix},$$

so that $||S_0^{-1}||_\infty = n$. This agrees with a well known result in, for instance, Hamilton (1994, equation (17.4.47)). In addition, the time series model

$$y_t = \beta_0 + \lambda_n y_{t-1} + \varepsilon_t,\ t = 2, ..., n,\ \lambda_n = 1 - c/k_n, \tag{4.1}$$

with $c > 0$, is called local to unit root (LUR) when $k_n = n$ and a moderate integration (MI) model when $k_n = n^\alpha$, and $\alpha \in (0, 1)$. See, for instance, Phillips (1987) and Phillips and Magdalinos (2007), respectively. Here,

$$S_0^{-1} = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ \lambda_n & 1 & 0 & \cdots & \\ \lambda_n^2 & \lambda_n & 1 & & \cdots \\ & \cdots & & & 0 \\ \lambda_n^{n-1} & \lambda_n^{n-2} & \cdots & \lambda_n & 1 \end{pmatrix}, \tag{4.2}$$

implying that

$$||S^{-1}(\theta)||_\infty = \frac{1}{1-\lambda_n} = \frac{k_n}{c}. \tag{4.3}$$

It follows that in the LUR case, $||S^{-1}(\theta)||_\infty = O(n)$ whereas in the MI case $||S^{-1}(\theta)||_\infty = O(n^a)$, $\alpha \in (0, 1)$, and the norming rates in (3.6) are $n^{3/2}$ and $n^{\alpha+1/2}$ for the two models, respectively, when a drift term is included in (4.1). The result for the LUR model is discussed in Phillips (1987, Section 6), noting that the limit distribution in this case is non-Gaussian when a drift does not exist and is Gaussian otherwise. We are not aware of similar discussion for the MI model in the literature.

The conclusion is that our setup is indeed very general, covering many special cases, from the SAR, similarity and nonstationary time series literature, with rates of convergence characterized by the order of magnitude of $\left\|S_0^{-1}\right\|_\infty$.

5 Final Remarks

We established in this paper asymptotic theory for the similarity based SAR model with exogenous regressors (1.2) under weak conditions. In particular, unlike the standard literature hitherto which has been done under the assumptions that $\left\|S_0^{-1}\right\|_\infty < K$ and $|\lambda_0| < 1$, our work allows for $\left\|S_0^{-1}\right\|_\infty = O\left(n^\gamma\right)$, with $\gamma \in [0,1]$ and $\lambda_0 = 1$. The result is a framework consisting of a very large class of models, with special cases including models behaving as a random walk with a drift, or even LUR or MI models with drifts, and, of course, standard SAR models with or without additional exogenous regressors. All cases are treated in a unified manner, with rates of convergence depending on the order of magnitude of $\left\|S_0^{-1}\right\|_\infty$, that is, on the value of γ. Extensions of our study to models including heteroscedastic errors seem challenging but highly desirable.

Appendix A

Proof or Theorem 1. To prove the identification condition, we write, for $\theta_2 \neq \theta_{20}$,

$$\tilde{\mathcal{L}}^p(\theta_2) - \tilde{\mathcal{L}}^p(\theta_{20}) = \log\left(\frac{y'S'M_ZSy}{y'S_0'M_ZS_0y}\right) - \frac{2}{n}\log|S| + \frac{2}{n}\log|S_0| + o_p(1) \tag{A.1}$$

$$= \log\left(\frac{y'S'M_ZSy}{y'S_0'M_ZS_0y}\right) + \frac{1}{n}\log|S^{-1\prime}S_0S_0'S^{-1\prime}| + o_p(1). \tag{A.2}$$

For identification it is required that (A.1) (or (A.2)) is strictly positive. We shall deal with the cases $\gamma = 0$ and $0 < \gamma \leq 1$ separately.

Case 1: $0 < \gamma \leq 1$**.** For this case, it follows from (2.18) and (2.22) that

$$\frac{y'S'M_ZSy}{y'S_0'M_ZS_0y} = O_p\left(\left\|S^{-1}\right\|_\infty^2\right)$$

which tends to $+\infty$ in this case. The second term on the rhs of (A.1) is bounded for all $\theta_2 \in \Theta_2$ since, by the geometric-arithmetic mean inequality and under Assumption 4,

$$1 = \frac{1}{n}tr(S) \geq |S|^{1/n} = |S'S|^{1/2n} = |\Omega^{-1}|^{1/2n} \geq (\eta_{\min}(\Omega^{-1}))^{1/2} = \frac{1}{(\eta_{\max}(\Omega))^{1/2}} > 0, \tag{A.3}$$

where $\Omega = (S'S)^{-1}$. The third term at the rhs of (A.1), in turn cannot diverge to $-\infty$ under A4, as

S_0 is non-singular.

Case 2: $\gamma = 0$**.** In view of (2.17), (2.22) and (2.26),

$$\begin{aligned}\frac{n^{-1}y'S'M_ZSy}{n^{-1}y'S_0'M_ZS_0y} &= \frac{n^{-1}\left(\sigma_0^2 tr\left(S_0^{-1\prime}S'SS_0^{-1}\right)+\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0\right)}{n^{-1}\left(\epsilon'M_Z\epsilon\right)}+o_p\left(1\right) \\ &= \frac{\sigma_0^2 tr\left(S_0^{-1\prime}S'SS_0^{-1}\right)+\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\sigma_0^2\left(n-m\right)}+o_p\left(1\right) \\ &= \frac{tr\left(S_0^{-1\prime}S'SS_0^{-1}\right)}{n}\left(1+\frac{\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\sigma_0^2 tr\left(S_0^{-1\prime}S'SS_0^{-1}\right)}\right)+o_p\left(1\right).\end{aligned} \tag{A.4}$$

Let $\Delta = S_0^{-1\prime}S'SS_0^{-1}$. It follows from (A.2) and the last line that

$$\tilde{\mathcal{L}}^p(\theta_2)-\tilde{\mathcal{L}}^p(\theta_{20}) = \log\left(\frac{1}{n}tr\left(\Delta\right)\left|\Delta\right|^{-1/n}\left(1+\frac{\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\sigma_0^2 tr\left(S_0^{-1\prime}S'SS_0^{-1}\right)}\right)\right)+o_p\left(1\right).$$

From the proof of Lemma 4 of RL,

$$\frac{1}{n}tr\left(\Delta\right)\left|\Delta\right|^{-1/n}\geq 1,$$

with equality iff $S'S = S_0'S_0$. Thus, $tr\left(\Delta\right)/n\left|\Delta\right|^{-1/n} > 1$ under Assumptions 3 and 4. Furthermore, under Assumption 8,

$$\lim_{n\to\infty}\frac{\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\sigma_0^2 tr\left(S_0^{-1\prime}S'SS_0^{-1}\right)}\geq 0,$$

with equality iff $\theta_2 = \theta_{20}$, under Assumption 3. Hence, for large enough n, $\tilde{\mathcal{L}}^p(\theta_2)-\tilde{\mathcal{L}}^p(\theta_{20})\geq 0$, with equality iff $\theta_2 = \theta_{20}$.

In order to show consistency of $\hat{\theta}$ we proceed along the lines of the proof of Theorem 1 of Delgado and Robinson (2015). Let $N_\delta = \{\theta : ||\theta_2-\theta_{20}|| < \delta\}$ for some $\delta > 0$, and $\bar{\mathcal{N}_\delta}$ its complement. We have,

$$\mathbb{P}(\hat{\theta}\in\bar{\mathcal{N}_\theta})\leq\mathbb{P}\big(\inf_{\bar{\mathcal{N}_\delta}}\mathcal{L}^p(\theta_2) < \mathcal{L}^p(\theta_{20})\big)\leq\mathbb{P}\big(\sup_{\Theta_2}|\mathcal{L}^p(\theta_2)-\tilde{\mathcal{L}}^p(\theta_2)|\geq\inf_{\bar{\mathcal{N}_\delta}}|\tilde{\mathcal{L}}^p(\theta_2)-\tilde{\mathcal{L}}^p(\theta_{20})|\big). \tag{A.5}$$

For consistency of $\hat{\theta}$ we need to establish the following statements:

$$\inf_{\bar{\mathcal{N}_\delta}}\left(\tilde{\mathcal{L}}^p(\theta_2)-\tilde{\mathcal{L}}^p(\theta_{20})\right) > \epsilon, \text{ for all sufficiently large } n \text{ and for some } \epsilon > 0, \tag{A.6}$$

$$\sup_{\Theta_2}|\mathcal{L}^p(\theta_2)-\tilde{\mathcal{L}}^p(\theta_2)|\xrightarrow{p}0, \quad \text{as} \quad n\to\infty. \tag{A.7}$$

The proofs of (A.6) and (A.7) are given in Lemmas 6 and 7, respectively. ■

Relevant quantities and matrices for Theorems 2 and 3.

We report here some matrices and lengthy expressions to avoid cumbersome notation in the body of the paper and in the proofs of Theorems 2 and 3.

We define

$$\Sigma_{10} = \lim_{n\to\infty} \frac{\sigma_0^4}{2n||S_0^{-1}||_\infty^2} \begin{pmatrix} tr((\underline{C_0 S_0^{-1}})^2) & tr((\underline{C_0 S_0^{-1}})(\underline{C_{1,0} S_0^{-1}})) & ... & tr((\underline{C_0 S_0^{-1}})(\underline{C_{k,0} S_0^{-1}})) \\ tr((\underline{C_{1,0} S^{-1}})(\underline{C_0 S_0^{-1}})) & tr((\underline{C_{1,0} S_0^{-1}})^2) & .... & tr((\underline{C_{1,0} S_0^{-1}})(\underline{C_{k,0} S_0^{-1}})) \\ .. & .. & tr((\underline{C_{2,0} S_0^{-1}})^2) & ... \\ .. & .. & .... & ... \\ ... & .. & ... & tr((\underline{C_{k,0} S_0^{-1}})^2) \end{pmatrix}, \tag{A.8}$$

$$\begin{aligned} \Sigma_{20} = &\lim_{n\to\infty} \frac{(\mu_0^{(4)} - 3\sigma_0^4)}{4n||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i ((\underline{C_0 S_0^{-1}}))_{ii}^2 & \sum_i (\underline{C_0 S_0^{-1}})_{ii} (\underline{C_{1,0} S_0^{-1}})_{ii} & ... & \sum_i (\underline{C_0 S_0^{-1}})_{ii} (\underline{C_{k,0} S_0^{-1}})_{ii} \\ \sum_i (\underline{C_{1,0} S_0^{-1}})_{ii} (\underline{C_0 S_0^{-1}})_{ii} & \sum_i ((\underline{C_{1,0} S_0^{-1}}))_{ii}^2 & .... & \sum_i (\underline{C_{1,0} S_0^{-1}})_{ii} (\underline{C_{k,0} S_0^{-1}})_{ii} \\ .. & .. & .. & ... \\ ... & .. & ... & \sum_i ((\underline{C_{k,0} S_0^{-1}}))_{ii}^2 \end{pmatrix} \\ &- \lim_{n\to\infty} \frac{(\mu_0^{(4)} - \sigma_0^4)}{4n^2||S_0^{-1}||_\infty^2} \begin{pmatrix} tr^2(\underline{C_0 S_0^{-1}}) & tr(\underline{C_0 S_0^{-1}}) tr(\underline{C_{1,0} S_0^{-1}}) & ... & tr(\underline{C_0 S_0^{-1}}) tr(\underline{C_{k,0} S_0^{-1}}) \\ tr(\underline{C_{1,0} S_0^{-1}}) tr(\underline{C_0 S_0^{-1}}) & tr^2(\underline{C_{1,0} S_0^{-1}}) & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & tr^2(\underline{C_{k,0} S_0^{-1}}) \end{pmatrix}, \end{aligned} \tag{A.9}$$

$$\Sigma_{30} = \lim_{n\to\infty} \frac{1}{n||S_0^{-1}||_\infty^2} \begin{pmatrix} \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S^{-1} Z \beta_0 & \beta_0' Z' S_0^{-1\prime} C_0 M_Z C_{1,0} S_0^{-1} Z \beta_0 & ... & ... \\ \beta_0' Z' S_0^{-1\prime} C_{1,0}' M_Z C_0 S^{-1} Z \beta_0 & \beta_0' Z' S_0^{-1\prime} C_{1,0}' M_Z C_{1,0} S^{-1} Z \beta_0 & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & \beta_0' Z' S_0^{-1\prime} C_{k,0}' M_Z C_{k,0} S^{-1} Z \beta_0 \end{pmatrix} \tag{A.10}$$

and

$$\Sigma_{40} = \lim_{n\to\infty} \frac{\mu_0^{(3)}}{n||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})_{ii}^d & \sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})_{ii}^d & ... \\ \sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})_{ii}^d & \sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})_{ii}^d & .... \\ .. & .. & .... \\ .. & ... & \sum_i (M_Z C_{k,0}' S_0^{-1} Z\beta_0)_i (C_{k,0} S_0^{-1})_{ii}^d \end{pmatrix}$$

$$+ \lim_{n\to\infty} \frac{\mu_0^{(3)}}{n||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})_{ii}^d & \sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})_{ii}^d & ... \\ \sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})_{ii}^d & \sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})_{ii}^d & .... \\ .. & .. & .... \\ .. & ... & \sum_i (M_Z C_{k,0}' S_0^{-1} Z\beta_0)_i (C_{k,0} S_0^{-1})_{ii}^d \end{pmatrix}, \tag{A.11}$$

where $\mu^{(3)} = E(\epsilon_i^3)$, $\mu^{(4)} = E(\epsilon_i^4)$ and $A_{ii}^d = a_{ii} - tr(A)/n$ for any $n \times n$ matrix A.

Also, we let

$$d_{11,0} = \lim_{n\to\infty} \frac{1}{n||S_0^{-1}||_\infty^2} \left(tr\left(\left(S_0^{-1\prime} C_0' + C_0 S_0^{-1} - tr(C_0 S_0^{-1}) \frac{2I}{n} \right)^2 \right) + \frac{2\beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S_0^{-1} Z\beta_0}{\sigma_0^2} \right), \tag{A.12}$$

$$\begin{aligned} d_{ij,0} = d_{ji,0} & \lim_{n\to\infty} \frac{2\lambda_0}{n||S_0^{-1}||_\infty^2} \left(2tr(C_{ij,0} S_0^{-1}) - \lambda_0 tr(S_0^{-1\prime} C_{j,0}' C_{i,0} S_0^{-1}) + tr(S_0^{-1} C_{i,0} S_0^{-1} C_{j,0}) \right) \\ & - \lim_{n\to\infty} \frac{4\lambda_0^2}{n^2||S_0^{-1}||_\infty^2} tr(C_{j,0} S_0^{-1}) tr(C_{i,0} S_0^{-1}) + \lim_{n\to\infty} \frac{2\lambda_0^2}{\sigma_0^2 n||S_0^{-1}||_\infty^2} \beta_0' Z' S_0^{-1\prime} C_{j,0}' M_Z C_{i,0} S_0^{-1} Z\beta_0, \\ & i,j = 2, ..., k+1, \end{aligned} \tag{A.13}$$

and

$$\begin{aligned} d_{1i,0} = & d_{i1,0} = \lim_{n\to\infty} \frac{2}{n||S_0^{-1}||_\infty^2} \left(\lambda_0 tr(S_0^{-1\prime} C_0' C_{i,0} S_0^{-1}) + tr(S_0^{-1} C_{i,0} S_0^{-1} C_0) \right) \\ & - \lim_{n\to\infty} \frac{4\lambda_0}{n^2||S_0^{-1}||_\infty^2} tr(C_0 S_0^{-1}) tr(C_{i,0} S_0^{-1}) + \lim_{n\to\infty} \frac{2\lambda_0}{\sigma_0^2 n||S_0^{-1}||_\infty^2} \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_{i,0} S_0^{-1} Z\beta_0, \\ & i = 2, ..., k+1. \end{aligned} \tag{A.14}$$

Furthermore,

$$\tilde{d}_{11,0} = \lim_{n\to\infty} \frac{2\beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S_0^{-1} Z\beta_0}{\sigma_0^2 n ||S_0^{-1}||_\infty^2}, \quad \tilde{d}_{ij,0} = \tilde{d}_{ji,0} = \lim_{n\to\infty} \frac{2\lambda_0^2 \beta_0' Z' S_0^{-1\prime} C_{j,0}' M_Z C_{i,0} S_0^{-1} Z\beta_0}{\sigma_0^2 n ||S_0^{-1}||_\infty^2},$$
$$\tilde{d}_{1i,0} = \tilde{d}_{i1,0} = \lim_{n\to\infty} \frac{2\lambda_0 \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_{i,0} S_0^{-1} Z\beta_0}{\sigma_0^2 n ||S_0^{-1}||_\infty^2} \tag{A.15}$$

for $i, j = 2, ...k+1$, and

$$F_0 = \begin{pmatrix} -\frac{2}{\sigma_0^2} & 0 & ... & 0 \\ 0 & -\frac{2\lambda_0}{\sigma_0^2} & 0 & .... \\ .. & .. & .. & ... \\ 0 & 0 & ... & -\frac{2\lambda_0}{\sigma_0^2} \end{pmatrix}. \tag{A.16}$$

We define

$$R_0 = \lim_{n\to\infty} \begin{pmatrix} d^{11,0} & 0 & .... & 0_{1\times m} \\ 0 & d^{22,0} & 0 & .. \\ ... & .... & ... & .... \\ 0 & .... & d^{(k+1)(k+1),0} & 0_{1\times m} \\ \left(\frac{Z'Z}{n}\right)^{-1} \frac{d^{11,0}}{n||S_0^{-1}||_\infty} Z' C_0 S_0^{-1} Z\beta_0 & \left(\frac{Z'Z}{n}\right)^{-1} \frac{\lambda_0 d^{22,0}}{n||S_0^{-1}||_\infty} Z' C_{1,0} S_0^{-1} Z\beta_0 & .... & \left(\frac{Z'Z}{n}\right)^{-1} \end{pmatrix}, \tag{A.17}$$

where we denoted by $d^{ij,0}$ the $(i,j)-$th element of the $(k+1)\times(k+1)$ matrix D_0 whose elements are defined in (A.12), (A.13) and (A.14), given

$$\hat{\theta}_2 - \theta_{20} = -D_0^{-1} \frac{1}{||S_0^{-1}||_\infty^2} \frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \theta_2} + o_p\left(\frac{1}{\sqrt{n}||S_0^{-1}||_\infty}\right), \tag{A.18}$$

as shown in (A.28). Furthermore, we let

$$F_0^\beta = \begin{pmatrix} -\frac{2}{\sigma_0^2} & 0 & ... & 0 & 0 \\ 0 & -\frac{2\lambda_0}{\sigma_0^2} & 0 & .... & ... \\ .. & .. & .. & ... & ... \\ 0 & 0 & ... & -\frac{2\lambda_0}{\sigma_0^2} & 0 \\ 0 & 0 & 0 & 0 & I_m \end{pmatrix}, \tag{A.19}$$

$$\Sigma_{10}^{\beta} = \lim_{n\to\infty} \frac{\sigma_0^4}{2n||S_0^{-1}||_\infty^2} \begin{pmatrix} tr((\underline{C_0 S_0^{-1}})^2) & tr((\underline{C_0 S_0^{-1}})(\underline{C_{1,0} S_0^{-1}})) & ... & tr((\underline{C_0 S_0^{-1}})(\underline{C_{k,0} S_0^{-1}})) & 0_{1\times m} \\ tr((\underline{C_{1,0} S^{-1}})(\underline{C_0 S_0^{-1}})) & tr((\underline{C_{1,0} S_0^{-1}})^2) & .... & tr((\underline{C_{1,0} S_0^{-1}})(\underline{C_{k,0} S_0^{-1}})) & 0_{1\times m} \\ .. & .. & tr((\underline{C_{2,0} S_0^{-1}})^2) & ... & 0_{1\times m} \\ .. & .. & .... & ... & 0_{1\times m} \\ ... & .. & ... & tr((\underline{C_{k,0} S_0^{-1}})^2) & 0_{1\times m} \\ 0_{m\times 1} & .... & .... & 0_{m\times 1} & 0_{m\times m} \end{pmatrix}, \tag{A.20}$$

$$\Sigma_{20}^{\beta} = \lim_{n\to\infty} \frac{(\mu_0^{(4)} - 3\sigma_0^4)}{4n||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i ((\underline{C_0 S_0^{-1}}))_{ii}^2 & \sum_i (\underline{C_0 S_0^{-1}})_{ii} (\underline{C_{1,0} S_0^{-1}})_{ii} & ... & \sum_i (\underline{C_0 S_0^{-1}})_{ii} (\underline{C_{k,0} S_0^{-1}})_{ii} & 0_{1\times m} \\ \sum_i (\underline{C_{1,0} S_0^{-1}})_{ii} (\underline{C_0 S_0^{-1}})_{ii} & \sum_i ((\underline{C_{1,0} S_0^{-1}}))_{ii}^2 & .... & \sum_i (\underline{C_{1,0} S_0^{-1}})_{ii} (\underline{C_{k,0} S_0^{-1}})_{ii} & 0_{1\times m} \\ .. & .. & .. & ... & 0_{1\times m} \\ ... & .. & ... & \sum_i ((\underline{C_{k,0} S_0^{-1}}))_{ii}^2 & 0_{1\times m} \\ 0_{m\times 1} & .... & .... & 0_{m\times 1} & 0_{m\times m} \end{pmatrix}$$
$$- \lim_{n\to\infty} \frac{(\mu_0^{(4)} - \sigma_0^4)}{4n^2||S_0^{-1}||_\infty^2} \begin{pmatrix} tr^2(\underline{C_0 S_0^{-1}}) & tr(\underline{C_0 S_0^{-1}}) tr(\underline{C_{1,0} S_0^{-1}}) & ... & tr(\underline{C_0 S_0^{-1}}) tr(\underline{C_{k,0} S_0^{-1}}) & 0_{m\times 1} \\ tr(\underline{C_{1,0} S_0^{-1}}) tr(\underline{C_0 S_0^{-1}}) & tr^2(\underline{C_{1,0} S_0^{-1}}) & .... & .... & 0_{m\times 1} \\ .. & .. & .. & ... & \\ ... & .. & ... & tr^2(\underline{C_{k,0} S_0^{-1}}) & 0_{m\times 1} \\ 0_{m\times 1} & .... & .... & 0_{m\times 1} & 0_{m\times m} \end{pmatrix}, \tag{A.21}$$

$$\Sigma_{30}^{\beta} = \lim_{n\to\infty} \sigma_0^2 \begin{pmatrix} \frac{\beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S^{-1} Z \beta_0}{n||S_0^{-1}||_\infty^2} & \frac{\beta_0' Z' S_0^{-1\prime} C_0 M_Z C_{1,0} S_0^{-1} Z \beta_0}{n||S_0^{-1}||_\infty^2} & ... & ... & 0_{1\times m} \\ \frac{\beta_0' Z' S_0^{-1\prime} C_{1,0}' M_Z C_0 S^{-1} Z \beta_0}{n||S_0^{-1}||_\infty^2} & \frac{\beta_0' Z' S_0^{-1\prime} C_{1,0}' M_Z C_{1,0} S^{-1} Z \beta_0}{n||S_0^{-1}||_\infty^2} & .... & .... & 0_{1\times m} \\ .. & .. & .. & ... & \\ ... & .. & ... & \frac{\beta_0' Z' S_0^{-1\prime} C_{k,0}' M_Z C_{k,0} S^{-1} Z \beta_0}{n||S_0^{-1}||_\infty^2} & 0_{1\times m} \\ 0_{m\times 1} & .... & .... & 0_{m\times 1} & \frac{Z'Z}{n} \end{pmatrix} \tag{A.22}$$

and

$$
\Sigma_{40}^{\beta} = \lim_{n\to\infty} \mu_0^{(3)}
\begin{pmatrix}
\frac{\sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \frac{\sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \dots & \frac{\sum_i (C_0 S_0^{-1})^d_{ii} z_i'}{n||S_0^{-1}||_\infty} \\
\frac{\sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \frac{\sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \dots. & \frac{\sum_i (C_{1,0} S_0^{-1})^d_{ii} z_i'}{n||S_0^{-1}||_\infty} \\
.. & .. & \dots. & \dots. \\
.. & \dots & \frac{\sum_i (M_Z C_{k,0}' S_0^{-1} Z\beta_0)_i (C_{k,0} S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \frac{\sum_i (C_{k,0} S_0^{-1})^d_{ii} z_i'}{n||S_0^{-1}||_\infty} \\
0_{m\times 1} & \dots & \dots & 0_{m\times m}
\end{pmatrix}
$$

$$
+ \lim_{n\to\infty} \mu_0^{(3)}
\begin{pmatrix}
\frac{\sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \frac{\sum_i (M_Z C_0' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \dots & 0_{1\times m} \\
\frac{\sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_0 S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \frac{\sum_i (M_Z C_{1,0}' S_0^{-1} Z\beta_0)_i (C_{1,0} S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & \dots. & 0_{1\times m} \\
.. & .. & \dots. & \\
.. & \dots & \frac{\sum_i (M_Z C_{k,0}' S_0^{-1} Z\beta_0)_i (C_{k,0} S_0^{-1})^d_{ii}}{n||S_0^{-1}||_\infty^2} & 0_{1\times m} \\
\frac{\sum_i (C_0 S_0^{-1})^d_{ii} z_i}{n||S_0^{-1}||_\infty} & \frac{\sum_i (C_{1,0} S_0^{-1})^d_{ii} z_i}{n||S_0^{-1}||_\infty} & \dots & 0_{m\times m}
\end{pmatrix}. \tag{A.23}
$$

Proof or Theorem 2.

We let $\partial\mathcal{L}^p(\hat\theta_2)/\partial\theta_2$ denote $\partial\mathcal{L}^p(\theta_2)/\partial\theta_2$ evaluated at $\hat\theta_2$, with a similar notation for analogous quantities. Let $\bar\theta_{2j}$ an intermediate point such that $|\bar\theta_{2j} - \theta_{20j}| < |\hat\theta_{2j} - \theta_{20j}|$ for $j = 1, ..., k+1$. By the MVT, for each $j = 1, ..., k+1$, we obtain

$$
\begin{aligned}
0 = \frac{\partial\mathcal{L}^p(\hat\theta_2)}{\partial\theta_{2j}} =& \frac{\partial\mathcal{L}^p(\theta_{20})}{\partial\theta_{2j}} + \sum_{l=1}^{k+1} \frac{\partial^2\mathcal{L}^p(\theta_{20})}{\partial\theta_{2j}\partial\theta_{2l}} (\hat\theta_{2l} - \theta_{20l}) \\
&+ \frac{1}{2}\sum_{l=1}^{k+1}\sum_{m=1}^{k+1} \frac{\partial^3\mathcal{L}^p(\bar\theta_2)}{\partial\theta_{2j}\partial\theta_{2l}\partial\theta_{2m}} (\hat\theta_{2l} - \theta_{20l})(\hat\theta_{2m} - \theta_{20m})
\end{aligned} \tag{A.24}
$$

For each $0 \le \gamma \le 1$ in (2.8), from Lemmas 8 and 9, respectively,

$$
\frac{\partial\mathcal{L}^p(\theta_{20})}{\partial\theta_2} = O_p\left(\frac{||S_0^{-1}||_\infty}{\sqrt{n}}\right) \tag{A.25}
$$

and

$$\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^2\mathcal{L}^p(\theta_{20})}{\partial\theta_2\partial\theta_2'}\xrightarrow{p} D_0, \tag{A.26}$$

where D_0 is nonsingular under Assumption 11. Also, by Lemma 10, for each $j, l, m = 1, ..., k+1$,

$$\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^3\mathcal{L}^p(\bar{\theta}_2)}{\partial\theta_{2j}\partial\theta_{2l}\partial\theta_{2m}} = O_p(1) \tag{A.27}$$

so that we can write in vector form

$$\begin{aligned}\hat{\theta}_2-\theta_{20} = &-\left(\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^2\mathcal{L}^p(\theta_{20})}{\partial\theta_2\partial\theta_2'}\right)^{-1}\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial\mathcal{L}^p(\theta_{20})}{\partial\theta_2}-\left(\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^2\mathcal{L}^p(\theta_{20})}{\partial\theta_2\partial\theta_2'}\right)^{-1}O_p(1/(\sqrt{n}||S_0^{-1}||_\infty)^2)\\ = &-D_0^{-1}\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial\mathcal{L}^p(\theta_{20})}{\partial\theta_2}+o_p\left(\frac{1}{\sqrt{n}||S_0^{-1}||_\infty}\right),\end{aligned} \tag{A.28}$$

where the first equality follows since the leading term is $O_p(1/(\sqrt{n}||S_0^{-1}||_\infty))$ from (A.25) and (A.26), and by replacing each component of $(\hat{\theta}_{2l}-\theta_{20l})(\hat{\theta}_{2m}-\theta_{20m})$ by $O_p(1/(\sqrt{n}||S_0^{-1}||_\infty)^2)$. The second equality follows from replacing (A.26).

We need to show that

$$\left(n||S_0^{-1}||_\infty^2\right)^{1/2}\begin{pmatrix}\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial\mathcal{L}^p(\theta_{20})}{\partial\lambda}\\ \frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial\mathcal{L}^p(\theta_{20})}{\partial w_1}\\ ..\\ ..\\ \frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial\mathcal{L}^p(\theta_{20})}{\partial w_k}\end{pmatrix}\underset{d}{\rightarrow}\mathcal{N}(0, F_0V_0F_0), \tag{A.29}$$

where V_0 is a positive definite variance-covariance matrix given by (3.1) and F_0 defined in (A.16). The proof of Theorem 2 will then follow by Crámer's theorem.

In order to show (A.29), from the derivation reported in the proof of Lemma 8, we define

$$U_n = U = \frac{1}{\left(n||S_0^{-1}||_\infty^2\right)^{1/2}}\begin{pmatrix}\epsilon'\left(S_0^{-1\prime}C_0'-\frac{I}{n}tr(C_0S_0^{-1})\right)\epsilon+\beta_0'Z'S_0^{-1}C_0'M_Z\epsilon\\ \epsilon'\left(S_0^{-1\prime}C_{1,0}'-\frac{I}{n}tr(C_{1,0}S_0^{-1})\right)\epsilon+\beta_0'Z'S_0^{-1}C_{1,0}'M_Z\epsilon\\ ..\\ ..\\ \epsilon'\left(S_0^{-1\prime}C_{k,0}'-\frac{I}{n}tr(C_{k,0}S_0^{-1})\right)\epsilon+\beta_0'Z'S_0^{-1}C_{k,0}'M_Z\epsilon\end{pmatrix},$$

so that (A.29) can be written as

$$\left(n||S_0^{-1}||_\infty^2\right)^{1/2}\begin{pmatrix} \frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} \\ \frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_1} \\ .. \\ .. \\ \frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_k} \end{pmatrix} = F_0 U + o_p(1). \tag{A.30}$$

Let $\underline{A} = A + A'$ for any generic matrix A. We define $\psi_{ij} = \psi_{ijn}$ and $\phi_{ij} = \phi_{ijn}$ the $(k+1)\times 1$ vectors $(\psi_{1ij}, ...\psi_{(k+1)ij})'$ and $(\phi_{1ij}, ...\phi_{(k+1)ij})'$ such that, for each $i, j = 1, ..., n$,

$$\psi_{ij} = \frac{1}{2}\begin{pmatrix} (\underline{C_0 S_0^{-1}})_{ij} \\ (\underline{C_{1,0} S_0^{-1}})_{ij} \\ .... \\ (\underline{C_{k,0} S_0^{-1}})_{ij} \end{pmatrix} \quad \text{and} \quad \phi_{ij} = \begin{pmatrix} (M_Z C_0 S_0^{-1})_{ij} \\ (M_Z C_{1,0} S_0^{-1})_{ij} \\ .... \\ (M_Z C_{k,0} S_0^{-1})_{ij} \end{pmatrix}, \tag{A.31}$$

respectively. Also, let Ψ_s and Φ_s be the $n \times n$ matrices with ψ_{sij} and ϕ_{sij} for $s = 1, ..., k+1$ as their respective$(i, j)-$th component. We can write $U = \sum_{i=1}^{n} u_i / \left(n||S_0^{-1}||_\infty^2\right)^{1/2}$, with

$$u_i = u_{in} = (\epsilon_i^2 - \sigma_0^2)\left(\psi_{ii} - \frac{1}{n}\sum_{j=1}^{n}\psi_{jj}\right) + 2\epsilon_i \sum_{j<i}\psi_{ij}\epsilon_j + \epsilon_i \sum_{j}^{n}\phi_{ij} z_j' \beta_0, \tag{A.32}$$

where z_j is the $m \times 1$ vector containing the $j-$th row of Z. So, $\{u_i, 1 \leq i \leq n, n = 1, 2, ..., ..\}$ is a triangular array of martingale differences with respect to the filtration formed by the σ-field generated by $\{\epsilon_j; j < i\}$.

In the sequel, all the summations will range from 1 to n, unless otherwise specified. Let

$$
\begin{aligned}
\Omega = \Omega_n =& Var(U) = \frac{1}{n||S_0^{-1}||_\infty^2}\sum_{i=1}^{n} Var(u_i) \\
=& \frac{\sigma_0^2}{n||S_0^{-1}||_\infty^2}\sum_i\sum_j\sum_t \phi_{ij} z_j' \beta_0 \beta_0' z_t \phi_{it}' + (\mu_0^{(4)} - \sigma_0^4)\frac{1}{n||S_0^{-1}||_\infty^2}\left(\sum_i \psi_{ii}\psi_{ii}' - \frac{1}{n}\sum_i\sum_j \psi_{ii}\psi_{jj}'\right) \\
&+ \frac{4}{n||S_0^{-1}||_\infty^2}\sigma_0^4 \sum_i\sum_{j<i}\psi_{ij}\psi_{ij}' + \frac{2\mu_0^{(3)}}{n||S_0^{-1}||_\infty^2}\sum_i\left(\psi_{ii} - \frac{1}{n}\sum_{j=1}^{n}\psi_{jj}\right)\sum_t \phi_{it}' z_t' \beta_0 \\
=& \frac{(\mu_0^{(4)} - 3\sigma_0^4)}{n||S_0^{-1}||_\infty^2}\sum_i \psi_{ii}\psi_{ii}' - \frac{(\mu_0^{(4)} - \sigma_0^4)}{n^2||S_0^{-1}||_\infty^2}\sum_i\sum_j \psi_{ii}\psi_{jj}' + \frac{2}{n||S_0^{-1}||_\infty^2}\sigma_0^4\sum_i\sum_j \psi_{ij}\psi_{ij}' \\
&+ \frac{\sigma_0^2}{n||S_0^{-1}||_\infty^2}\sum_i\sum_j\sum_t \phi_{ij} z_j' \beta_0 \beta_0' z_t \phi_{it}' + \frac{2\mu_0^{(3)}}{n||S_0^{-1}||_\infty^2}\sum_i\left(\psi_{ii} - \frac{1}{n}\sum_{j=1}^{n}\psi_{jj}\right)\sum_t \phi_{it}' z_t' \beta_0, \qquad (A.33)
\end{aligned}
$$

and $v_i = z_{in} = \zeta'\Omega^{-1/2}u_i / \left(n||S_0^{-1}||_\infty\right)^{1/2}$, with ζ being any deterministic $(k+1)\times 1$ vector that satisfies $\zeta'\zeta = 1$. By Theorem 2 of Scott (1973), $\sum_{i=1}^{n} v_i \underset{d}{\rightarrow} \mathcal{N}(0,1)$, as long as

$$\sum_{i=1}^{n}\mathbb{E}\left(v_i^2|\epsilon_j, j<i\right) \underset{p}{\rightarrow} 1 \qquad (A.34)$$

and

$$\sum_{i=1}^{n}\mathbb{E}\left(v_i^2 1(|v_i|>\delta)\right) \rightarrow 0, \quad \forall\delta>0, \qquad (A.35)$$

where $1(\cdot)$ is the indicator function.

We define

$$V_0 = \lim_{n\rightarrow\infty}\Omega \equiv \Sigma_{10} + \Sigma_{20} + \Sigma_{30} + \Sigma_{40}, \qquad (A.36)$$

with

$$\Sigma_{10} = \lim_{n\rightarrow\infty}\frac{2\sigma_0^4}{n||S_0^{-1}||_\infty^2}\begin{pmatrix} \sum_i\sum_j \psi_{1ij}\psi_{1ji} & \sum_i\sum_j \psi_{1ij}\psi_{2ji} & ... & ... \\ \sum_i\sum_j \psi_{2ij}\psi_{1ji} & \sum_i\sum_j \psi_{2ij}\psi_{2ji} & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & \sum_i\sum_j \psi_{(k+1)ij}\psi_{(k+1)ji} \end{pmatrix}, \qquad (A.37)$$

$$\Sigma_{20} = \lim_{n\to\infty} \frac{(\mu_0^{(4)} - 3\sigma_0^4)}{n||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i \psi_{1ii}^2 & \sum_i \psi_{1ii}\psi_{2ii} & ... & ... \\ \sum_i \psi_{2ii}\psi_{1ii} & \sum_i \psi_{2ii}^2 & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & \sum_i \psi_{(k+1)ii}^2 \end{pmatrix}$$

$$- \lim_{n\to\infty} \frac{(\mu_0^{(4)} - \sigma_0^4)}{n^2||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i\sum_j \psi_{1ii}\psi_{1jj} & \sum_i\sum_j \psi_{1ii}\psi_{2jj} & ... & ... \\ \sum_i\sum_j \psi_{2ii}\psi_{1jj} & \sum_i\sum_j \psi_{2ii}\psi_{2jj} & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & \sum_i\sum_j \psi_{(k+1)ii}\psi_{(k+1)jj} \end{pmatrix}, \qquad (A.38)$$

$$\Sigma_{30} = \lim_{n\to\infty} \frac{\sigma_0^2}{n||S_0^{-1}||_\infty^2} \begin{pmatrix} \beta_0' Z'\Phi_1'\Phi_1 Z\beta_0 & \beta_0' Z'\Phi_1'\Phi_2 Z\beta_0 & ... & ... \\ \beta_0' Z'\Phi_2'\Phi_1 Z\beta_0 & \beta_0' Z'\Phi_2'\Phi_2 Z\beta_0 & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & \beta_0' Z'\Phi_{k+1}'\Phi_{k+1} Z\beta_0 \end{pmatrix} \qquad (A.39)$$

and

$$\Sigma_{40} = \lim_{n\to\infty} \frac{\mu_0^{(3)}}{n||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i(\psi_{1ii} - \frac{tr(\Psi_1)}{n})(\Phi_1 Z\beta_0)_i & \sum_i(\psi_{1ii} - \frac{tr(\Psi_1)}{n})(\Phi_2 Z\beta_0)_i & ... & ... \\ \sum_i(\psi_{2ii} - \frac{tr(\Psi_2)}{n})(\Phi_1 Z\beta_0)_i & \sum_i(\psi_{2ii} - \frac{tr(\Psi_2)}{n})(\Phi_2 Z\beta_0)_i & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & \sum_i(\psi_{(k+1)ii} - \frac{tr(\Psi_{k+1})}{n})(\Phi_{k+1} Z\beta_0)_i \end{pmatrix}$$

$$+ \lim_{n\to\infty} \frac{\mu_0^{(3)}}{n||S_0^{-1}||_\infty^2} \begin{pmatrix} \sum_i(\psi_{1ii} - \frac{tr(\Psi_1)}{n})(\Phi_1 Z\beta_0)_i & \sum_i(\psi_{2ii} - \frac{tr(\Psi_2)}{n})(\Phi_1 Z\beta_0)_i & ... & ... \\ \sum_i(\psi_{1ii} - \frac{tr(\Psi_1)}{n})(\Phi_2 Z\beta_0)_i & \sum_i(\psi_{2ii} - \frac{tr(\Psi_2)}{n})(\Phi_2 Z\beta_0)_i & .... & .... \\ .. & .. & .. & ... \\ ... & .. & ... & \sum_i(\psi_{(k+1)ii} - \frac{tr(\Psi_{k+1})}{n})(\Phi_{k+1} Z\beta_0)_i \end{pmatrix} \qquad (A.40)$$

where the explicit forms of Σ_{10}, Σ_{20}, Σ_{30} and Σ_{40} are given in (A.8), (A.9), (A.10) and (A.11), respectively. Also, elements of Σ_{30} are $O(1)$ from Lemma 4(a) and it is nonsingular under Assumption 11, while each element of Σ_{10} is $O(1/||S_0^{-1}||_\infty)$ by Lemma 4(f) and each element of Σ_{20} is $O(1/||S_0^{-1}||_\infty^2)$ by Lemma 4(e). Elements of Σ_{40} are $O(1/||S_0^{-1}||_\infty)$ from Lemma 4(e), (B.8) and since, for any $n \times n$ generic matrix A such that $||A||_\infty + ||A'||_\infty < K$ and any bounded $n \times 1$ vector a,

$$|AS_0^{-1}a|_i \leq K \sup_j \sum_t |s^{jt}| \sup_i \sum_j |a_{ij}| = O(||S_0^{-1}||_\infty). \qquad (A.41)$$

The proof of (A.34) and (A.35) are reported at the end of Appendix B.

Therefore $\sum_{i=1}^{n} v_i \underset{d}{\to} \mathcal{N}(0,1)$ implies $U \underset{d}{\to} \mathcal{N}(0,V)$, and (A.29) follows from Cramér's theorem. The statement in Theorem 2 follows then from (A.26) and Cramér's theorem, with $\mathcal{V}_0 = D_0^{-1} F_0 V_0 F_0 D_0^{-1}$.

Proof or Theorem 3.

The proof of Theorem 3 follows closely from that of Theorem 2 and much of the details are omitted to avoid repetitions. We can write

$$\begin{pmatrix} \frac{\sqrt{n}}{||S_0^{-1}||_\infty} \frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} \\ \frac{\sqrt{n}}{||S_0^{-1}||_\infty} \frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_1} \\ .. \\ .. \\ \frac{\sqrt{n}}{||S_0^{-1}||_\infty} \frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_k} \\ \frac{1}{\sqrt{n}} Z'\epsilon \end{pmatrix} = F_0^\beta U^\beta + o_p(1), \tag{A.42}$$

with

$$U_n^\beta = U^\beta = \begin{pmatrix} \frac{1}{\left(n||S_0^{-1}||_\infty^2\right)^{1/2}} \epsilon' \left(S_0^{-1'} C_0' - \frac{I}{n} tr(C_0 S_0^{-1})\right) \epsilon + \frac{1}{\left(n||S_0^{-1}||_\infty^2\right)^{1/2}} \beta_0' Z' S_0^{-1} C_0' M_Z \epsilon \\ \frac{1}{\left(n||S_0^{-1}||_\infty^2\right)^{1/2}} \epsilon' \left(S_0^{-1'} C_{1,0}' - \frac{I}{n} tr(C_{1,0} S_0^{-1})\right) \epsilon + \frac{1}{\left(n||S_0^{-1}||_\infty^2\right)^{1/2}} \beta_0' Z' S_0^{-1} C_{1,0}' M_Z \epsilon \\ .. \\ .. \\ \frac{1}{\left(n||S_0^{-1}||_\infty^2\right)^{1/2}} \epsilon' \left(S_0^{-1'} C_{k,0}' - \frac{I}{n} tr(C_{k,0} S_0^{-1})\right) \epsilon + \frac{1}{\left(n||S_0^{-1}||_\infty^2\right)^{1/2}} \beta_0' Z' S_0^{-1} C_{k,0}' M_Z \epsilon \\ \frac{1}{\sqrt{n}} Z'\epsilon \end{pmatrix} \tag{A.43}$$

and F_0^j defined in (A.19). The proof of Theorem 3 will then follow by Crámer's theorem.

Similarly to the proof of Theorem 2, we define $\underline{A} = A + A'$ for any generic matrix A. We define $\psi_{ij}^\beta = \psi_{ijn}^\beta$, $\phi_{ij}^\beta = \phi_{ijn}^\beta$ and $\tau_i^\beta = \tau_{in}^\beta$ the $(k+m+1) \times 1$ vectors such that

$$\psi_{ij}^\beta = \frac{1}{2(n||S_0^{-1}||_\infty^2)^{1/2}} \begin{pmatrix} (\underline{C_0 S_0^{-1}})_{ij} \\ (\underline{C_{1,0} S_0^{-1}})_{ij} \\ .... \\ (\underline{C_{k,0} S_0^{-1}})_{ij} \\ 0_{m\times 1} \end{pmatrix}, \quad \phi_{ij}^\beta = \frac{1}{(n||S_0^{-1}||_\infty^2)^{1/2}} \begin{pmatrix} (M_Z C_0 S_0^{-1})_{ij} \\ (M_Z C_{1,0} S_0^{-1})_{ij} \\ .... \\ (M_Z C_{k,0} S_0^{-1})_{ij} \\ 0_{m\times 1} \end{pmatrix} \text{ and } \tau_i^\beta = \frac{1}{\sqrt{n}} \begin{pmatrix} 0_{(k+1)\times 1} \\ z_i \end{pmatrix} \tag{A.44}$$

respectively. Also, let Ψ_s^β and Φ_s^β be the $n \times n$ matrices with ψ_{sij}^β and ϕ_{sij}^β for $s = 1, ..., k+m+1$ as their respective$(i,j)-$th component[3]. We can write $U^\beta=\sum_{i=1}^{n} u_i^\beta$, with

$$u_i^\beta=u_{in}^\beta=(\epsilon_i^2-\sigma_0^2)\left(\psi_{ii}^\beta-\frac{1}{n}\sum_{j=1}^{n}\psi_{jj}^\beta\right)+2\epsilon_i\sum_{j<i}\psi_{ij}^\beta\epsilon_j+\epsilon_i\left(\tau_i^\beta+\sum_j^n\phi_{ij}^\beta z_j'\beta_0\right). \tag{A.45}$$

As in the proof of Theorem 2, we define

$$\Omega^\beta=\Omega_n^\beta=Var(U^\beta)=\sum_{i=1}^{n}Var(u_i^\beta)=(\mu_0^{(4)}-3\sigma_0^4)\sum_i\psi_{ii}^\beta\psi_{ii}^{\beta\prime}-(\mu_0^{(4)}-\sigma_0^4)\sum_i\sum_j\psi_{ii}^\beta\psi_{jj}^{\beta\prime}+2\sigma_0^4\sum_i\sum_j\psi_{ij}^\beta\psi_{ij}^{\beta\prime}$$
$$+\sigma_0^2\sum_i\left(\sum_j\phi_{ij}^\beta z_j'\beta_0+\tau_i^\beta\right)\left(\sum_t\beta_0'z_t\phi_{it}^{\beta\prime}+\tau_i^{\beta\prime}\right)+2\mu_0^{(3)}\sum_i\left(\psi_{ii}^\beta-\frac{1}{n}\sum_{j=1}^{n}\psi_{jj}^\beta\right)\left(\sum_t\phi_{it}^{\beta\prime}z_t'\beta_0+\tau_i^{\beta\prime}\right), \tag{A.46}$$

and $v_i^\beta=z_{in}=\zeta'\Omega^{\beta\ -1/2}u_i^\beta$, with ζ being any deterministic $(k+m+1)\times 1$ vector that satisfies $\zeta'\zeta=1$. The rest of the proof follows by routine arguments and is omitted, after defining

$$V_0^\beta=\lim_{n\to\infty}\Omega^\beta\equiv\Sigma_{10}^\beta+\Sigma_{20}^\beta+\Sigma_{30}^\beta+\Sigma_{40}^\beta, \tag{A.47}$$

with

$$\Sigma_{10}^\beta=\lim_{n\to\infty}2\sigma_0^4\begin{pmatrix}\sum_i\sum_j\psi_{1ij}^\beta\psi_{1ji}^\beta & \sum_i\sum_j\psi_{1ij}^\beta\psi_{2ji}^\beta & \cdots & \cdots & 0_{1\times m}\\ \sum_i\sum_j\psi_{2ij}^\beta\psi_{1ji}^\beta & \sum_i\sum_j\psi_{2ij}^\beta\psi_{2ji}^\beta & \cdots\cdot & \cdots\cdot & 0_{1\times m}\\ .. & .. & .. & \cdots & \cdots\\ \cdots & .. & \cdots & \sum_i\sum_j\psi_{(k+1)ij}^\beta\psi_{(k+1)ji}^\beta & 0_{1\times m}\\ 0_{m\times 1} & \cdots\cdot & \cdots & \cdots\cdot & 0_{m\times m}\end{pmatrix}, \tag{A.48}$$

[3]We note that for $s=k+2,....,k+m+1$ both Ψ_s^β and Φ_s^β are matrices of zeros.

$$\Sigma_{20}^{\beta} = \lim_{n\to\infty}(\mu_0^{(4)} - 3\sigma_0^4)\begin{pmatrix} \sum_i(\psi_{1ii}^{\beta})^2 & \sum_i \psi_{1ii}^{\beta}\psi_{2ii}^{\beta} & ... & ... & 0_{1\times m} \\ \sum_i \psi_{2ii}^{\beta}\psi_{1ii}^{\beta} & \sum_i(\psi_{2ii}^{\beta})^2 & .... & .... & 0_{1\times m} \\ .. & .. & .. & ... & .... \\ .. & .. & ... & \sum_i(\psi_{(k+1)ii}^{\beta})^2 & 0_{1\times m} \\ 0_{m\times 1} & ... & ... & ... & 0_{m\times m} \end{pmatrix}$$

$$-\lim_{n\to\infty}(\mu_0^{(4)} - \sigma_0^4)\begin{pmatrix} \sum_i\sum_j \psi_{1ii}^{\beta}\psi_{1jj}^{\beta} & \sum_i\sum_j \psi_{1ii}^{\beta}\psi_{2jj}^{\beta} & ... & ... & 0_{1\times m} \\ \sum_i\sum_j \psi_{2ii}^{\beta}\psi_{1jj}^{\beta} & \sum_i\sum_j \psi_{2ii}^{\beta}\psi_{2jj}^{\beta} & .... & .... & 0_{1\times m} \\ .. & .. & .. & ... & 0_{1\times m} \\ ... & .. & ... & \sum_i\sum_j \psi_{(k+1)ii}^{\beta}\psi_{(k+1)jj}^{\beta} & 0_{1\times m} \\ ... & ... & ... & .... & ... \\ 0_{m\times 1} & 0_{m\times 1} & ... & ... & 0_{m\times m} \end{pmatrix}, \quad \text{(A.49)}$$

$$\Sigma_{30}^{\beta} = \lim_{n\to\infty}\sigma_0^2\begin{pmatrix} \beta_0' Z'\Phi_1^{\beta\prime}\Phi_1^{\beta}Z\beta_0 & \beta_0' Z'\Phi_1^{\beta\prime}\Phi_2^{\beta}Z\beta_0 & ... & \beta_0' Z'\Phi_1^{\beta\prime}\Phi_{k+1}^{\beta}Z\beta_0 & \beta_0 Z'\Phi_1^{\beta\prime}Z/\sqrt{n} \\ \beta_0' Z'\Phi_2^{\beta\prime}\Phi_1^{\beta}Z\beta_0 & \beta_0' Z'\Phi_2^{\beta\prime}\Phi_2^{\beta}Z\beta_0 & .... & \beta_0' Z'\Phi_2^{\beta\prime}\Phi_{k+1}^{\beta}Z\beta_0 & \beta_0 Z'\Phi_2^{\beta\prime}Z/\sqrt{n} \\ .. & .. & .. & ... & \\ ... & .. & ... & \beta_0' Z'\Phi_{k+1}^{\beta\prime}\Phi_{k+1}^{\beta}Z\beta_0 & \beta_0 Z'\Phi_{k+1}^{\beta\prime}Z/\sqrt{n} \\ Z'\Phi_1^{\beta}Z\beta_0/\sqrt{n} & Z'\Phi_2^{\beta}Z\beta_0/\sqrt{n} & ... & Z'\Phi_{k+1}^{\beta}Z\beta_0/\sqrt{n} & Z'Z/n \end{pmatrix} \quad \text{(A.50)}$$

and, by letting $\psi_{sii}^{\beta d} = \psi_{sii}^{\beta} - tr(\Psi_s^{\beta})/n$ for each $s = 1, ..., k+m+1$,

$$\Sigma_{40}^{\beta} = \lim_{n\to\infty}\mu_0^{(3)}\begin{pmatrix} \sum_i \psi_{1ii}^{\beta d}(\Phi_1^{\beta}Z\beta_0)_i & \sum_i \psi_{1ii}^{\beta d}(\Phi_2^{\beta}Z\beta_0)_i & ... & \sum_i \psi_{1ii}^{\beta d}(\Phi_{k+1}^{\beta}Z\beta_0)_i & \sum_i \psi_{1ii}^{\beta d}z_i'/\sqrt{n} \\ \sum_i \psi_{2ii}^{\beta d}(\Phi_1^{\beta}Z\beta_0)_i & \sum_i \psi_{2ii}^{\beta d}(\Phi_2^{\beta}Z\beta_0)_i & .... & .\sum_i \psi_{2ii}^{\beta d}(\Phi_{k+1}^{\beta}Z\beta_0)_i & \sum_i \psi_{2ii}^{\beta d}z_i'/\sqrt{n} \\ ... & .... & .... & .... & .... \\ .. & .. & .. & ...\sum_i \psi_{(k+1)ii}^{\beta d}(\Phi_{k+1}^{\beta}Z\beta_0)_i & \sum_i \psi_{(k+1)ii}^{\beta d}z_i'/\sqrt{n} \\ 0_{m\times 1} & 0_{m\times 1} & ... & ... & 0_{m\times m} \end{pmatrix}$$

$$+\lim_{n\to\infty}\mu_0^{(3)}\begin{pmatrix} \sum_i \psi_{1ii}^{\beta d}(\Phi_1^{\beta}Z\beta_0)_i & \sum_i \psi_{2ii}^{\beta d}(\Phi_1^{\beta}Z\beta_0)_i & ... & \sum_i \psi_{(k+1)ii}^{\beta d}(\Phi_1^{\beta}Z\beta_0)_i & 0_{1\times m} \\ \sum_i \psi_{1ii}^{\beta d}(\Phi_2^{\beta}Z\beta_0)_i & \sum_i \psi_{2ii}^{\beta d}(\Phi_2^{\beta}Z\beta_0)_i & .... & .\sum_i \psi_{(k+1)ii}^{\beta d}(\Phi_2^{\beta}Z\beta_0)_i & 0_{1\times m} \\ ... & .... & .... & .... & .... \\ .. & .. & .. & ...\sum_i \psi_{(k+1)ii}^{\beta d}(\Phi_{k+1}^{\beta}Z\beta_0)_i & 0_{1\times m} \\ \sum_i \psi_{1ii}^{\beta d}z_i/\sqrt{n} & \sum_i \psi_{2ii}^{\beta d}z_i/\sqrt{n} & ... & \sum_i \psi_{(k+1)ii}^{\beta d}z_i/\sqrt{n} & 0_{m\times m} \end{pmatrix} \quad \text{(A.51)}$$

where the explicit forms of Σ_{10}^{β}, Σ_{20}^{β}, Σ_{30}^{β} and Σ_{40}^{β} are given in (A.20), (A.21), (A.22) and (A.23), respectively. We outline that each element containing $Z'\Phi_s^{\beta\prime}Z$, for each $s = 1, ...k+1$, is a null matrix from (A.44).

Appendix B.

Lemma 1. *For all n, each element z_{ij} of Z $(n \times m)$ is non-random and $|z_{ij}| < K$. Also, for all sufficiently large n,*

$$0 < c < \eta_{\min}\left(\frac{Z'Z}{n}\right), \tag{B.1}$$

where c is any arbitrarily small constant. It follows that

$$||M_Z||_\infty \leq K. \tag{B.2}$$

Proof of Lemma 1. We show that $||Z(Z'Z)^{-1}Z'||_\infty < K$, and thence the claim in (B.2) follows trivially. Let z_i' the $i-$th row of Z, in line with our usual notation. The arbitrary constant K can change its value from step to step, as usual. We have

$$\begin{aligned} ||Z(Z'Z)^{-1}Z'||_\infty =& \max_i \sum_{j=1}^{n} |z_i'(Z'Z)z_j| \leq \max_i \sum_{j=1}^{n} ||z_i|| ||(Z'Z)^{-1}|| ||z_j|| \\ \leq& \max_{i,j} ||z_i|| \left|\left|\left(\frac{Z'Z}{n}\right)^{-1}\right|\right| ||z_j|| \leq \frac{mK^2}{c} < K, \end{aligned} \tag{B.3}$$

since

$$\left|\left|\left(\frac{Z'Z}{n}\right)^{-1}\right|\right| = \frac{1}{\eta_{\min}(Z'Z/n)} \leq \frac{1}{c} \tag{B.4}$$

and

$$\max_i ||z_i|| = \max_i (z_i'z_i)^{1/2} \leq (mK^2)^{1/2} \leq K. \tag{B.5}$$

■

In order to prove the following Lemmas we introduce the following assumption.

Assumption A1 *Let ϵ be an $n \times 1$ vector of i.i.d. random variables, satisfying*

$$\mathbb{E}(\epsilon_i) = 0, \quad \mathbb{E}(\epsilon_i)^4 < K \quad \forall i = 1, ..., n.$$

Also, let $A = A(\theta_2)$ be an $n \times n$ generic matrix, such that $||A||_\infty + ||A'||_\infty < K$ for all $\theta_2 \in \Theta_2$. Thus, we also have, $|A_{ij}| < K$ for all $i, j = 1, ..., n$ and for all $\theta_2 \in \Theta_2$. The proofs of Lemmas 2, 3 and 5 are given in the Online Supplement of RL.

Lemma 2. *Under Assumption A1, for all* $\theta_2 \in \Theta_2$*:*

a) $\epsilon' S^{-1}(\theta_2)' A \epsilon = O_p(n)$.

b) $\epsilon' S^{-1}(\theta_2)' A S^{-1}(\theta_2) \epsilon = O_p(n||S^{-1}||_\infty)$.

Lemma 3. *Under Assumption A1, for all* $\theta_2 \in \Theta_2$*:*

$$\frac{1}{n||S^{-1}(\theta_2)||_\infty^2} \left(\epsilon' \left(S^{-1}(\theta_2)' A S^{-1}(\theta_2)\right) \epsilon - \sigma^2 tr \left(S^{-1}(\theta_2)' A S^{-1}(\theta_2)\right)\right) = O_p\left(\left(\frac{1}{n||S^{-1}(\theta_2)||_\infty}\right)^{1/2}\right). \tag{B.6}$$

Lemma 4. *Let* a *an* $n \times 1$ *vector such that* $|a_i| < K$ *for all* $i = 1, ...m$ *and* A *an* $n \times n$ *matrix such that* $||A||_\infty + ||A'||_\infty < K$*. Let* $B = B(\theta_2) = (S^{-1\prime} A + A' S^{-1})/2$*. For all* $\theta_2 \in \Theta_2$*:*

a) $a' S^{-1\prime} A S^{-1} a = O\left(n||S^{-1}||_\infty^2\right)$;

b) $\epsilon' S^{-1\prime} A S^{-1} a = O_p(\sqrt{n}||S^{-1}||_\infty^2)$;

c) $\epsilon' A S^{-1} a = O_p(\sqrt{n}||S^{-1}||_\infty)$;

d) $a' S^{-1\prime} A a = O\left(n||S^{-1}||_\infty;\right)$;

e) $tr(B) = O(n)$;

f) $tr(B^2) = O(n||S^{-1}||_\infty)$;

g) $tr(B^3) = O(n||S^{-1}||_\infty^2)$.

Proof of Lemma 4 We let $B = B(\theta_2) = (S^{-1\prime} A + A' S^{-1})/2$. We have

$$||B||_\infty = O(||S^{-1}||_\infty), \tag{B.7}$$

$$|b_{ij}| \leq K \sum_{t=1}^n |s^{ti}||a_{tj}| \leq K \sup_{i,t} |s^{ti}| \sup_j \sum_{t=1}^n |a_{tj}| = O(1) \quad \forall i, j \tag{B.8}$$

and

$$|(BS^{-1})_{ij}| \leq K \sum_{t=1}^n |b_{it}||s^{tj}| \leq K \sup_{i,t} |b_{it}| \sup_j \sum_{t=1}^n |s^{tj}| = O(||S^{-1}||_\infty) \quad \forall i, j. \tag{B.9}$$

By standard norm inequalities

$$||S^{-1\prime}AS^{-1}||_\infty = O(||S^{-1}||_\infty^2). \tag{B.10}$$

Proof of part (a). We have

$$\begin{aligned}|a'S^{-1\prime}AS^{-1}a| = &\left|\sum_{i=1}^n\sum_{j=1}^n a_ia_j(S^{-1\prime}AS^{-1})_{ij}\right| \leq \sum_{i=1}^n\sum_{j=1}^n |a_i||a_j||(S^{-1\prime}AS^{-1})_{ij}| \leq Kn||S^{-1\prime}AS^{-1}||_\infty \\ =&O\left(n||S^{-1}||_\infty^2\right), \end{aligned} \tag{B.11}$$

which concludes part (a).

Proof of part (b). $\epsilon'S^{-1\prime}AS^{-1}a$ has mean zero and variance bounded by

$$\begin{aligned}&K\sum_{i=1}^n\sum_{j=1}^n\sum_{v=1}^n |a_j||a_v||(S^{-1\prime}AS^{-1})_{ij}||(S^{-1\prime}AS^{-1})_{iv}| \leq K\max_i\sum_{v=1}^n|(S^{-1\prime}AS^{-1})_{iv}|\sum_{i=1}^n\sum_{j=1}^n|(S^{-1\prime}AS^{-1})_{ij}| \\ \leq &Kn\max_i\sum_{v=1}^n|(S^{-1\prime}AS^{-1})_{iv}|\max_i\sum_{j=1}^n|(S^{-1\prime}AS^{-1})_{ij}| = O\left(n||S^{-1}||_\infty^4\right). \end{aligned} \tag{B.12}$$

The claim in part b) follows by Markov inequality.

The proof of parts (c) and (d) follow from very similar arguments to those used to prove parts (b) and (a), respectively, and it is omitted to avoid repetitions.

Proof of part (e). We have

$$tr(B) \leq K\sum_{i=1}^n\sum_{j=1}^n |s^{ji}||a_{ji}| \leq K\ n\sup_{i,j}|s^{ji}|\sup_j\sum_{i=1}^n|a_{ji}| = O(n). \tag{B.13}$$

Proof of part (f). We have

$$tr(B^2) \leq \sum_{i=1}^n\sum_{j=1}^n |b_{ij}||b_{ij}| \leq Kn||S_0^{-1}||_\infty = O(n||S^{-1}||_\infty). \tag{B.14}$$

Proof of part (g). We have

$$tr(B^3) \leq \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t=1}^{n}|b_{ij}||b_{jt}||b_{ti}| \leq Kn\sup_i \sum_{j=1}^{n}\sum_{t=1}^{n}|b_{jt}||b_{ti}|$$
$$\leq Kn\sup_t \sum_{j=1}^{n}|b_{jt}|\sup_i \sum_{t=1}^{n}|b_{ti}| = O\left(n||S^{-1}||_\infty^2\right). \quad \text{(B.15)}$$

■

Lemma 5. *Under Assumption A1, for all $\theta_2 \in \Theta_2$:*

$$\frac{1}{n||S_0^{-1}||_\infty}\epsilon'\left(S^{-1}(\theta_2)'A - \frac{I}{n}tr\left(S^{-1}(\theta_2)'A\right)\right)\epsilon = O_p\left(\left(\frac{1}{n||S_0^{-1}||_\infty}\right)^{1/2}\right). \quad \text{(B.16)}$$

Lemma 6. *Under Assumptions 1-9,*

$$\inf_{\mathcal{N}_\delta}\left(\tilde{\mathcal{L}}^p(\theta_2) - \tilde{\mathcal{L}}^p(\theta_{20})\right) > \epsilon, \quad \text{(B.17)}$$

for all sufficiently large n and for some $\epsilon > 0$, with $\tilde{\mathcal{L}}^p(\cdot)$ defined in (2.25).

Proof of Lemma 6. We prove the Lemma by using the inequality

$$\inf_{\theta_2:\left\|\theta_2-\theta_2^\dagger\right\|<\eta;\ \theta_2\in\Theta}\left(\tilde{\mathcal{L}}^p(\theta_2) - \tilde{\mathcal{L}}^p(\theta_{20})\right) \geq \left(\tilde{\mathcal{L}}^p(\theta_2^\dagger) - \tilde{\mathcal{L}}^p(\theta_{20})\right)$$
$$- \sup_{\theta_2:\left\|\theta_2-\theta_2^\dagger\right\|<\eta;\ \theta_2\in\Theta}\left|\tilde{\mathcal{L}}^p(\theta_2) - \tilde{\mathcal{L}}^p(\theta_2^\dagger)\right|, \quad \text{(B.18)}$$

where η is a positive constant, $\theta_2^\dagger \in \Theta_2\backslash\theta_{20}$, and Θ_2 is compact under Assumption 2 and hence it has a finite subcover. We need to show that the RHS of (B.18) is strictly positive for large n. From the proof of Theorem 1, the first term on the RHS of (B.18) is strictly positive for $\gamma = 0$ and diverges to $+\infty$ for $0 < \gamma \leq 1$ as $n \to \infty$. We continue to analyze the second term on the RHS of (B.18). Consider first the case $0 < \gamma \leq 1$. Let $S^\dagger = S(\theta_2^\dagger)$. Since the first term of the RHS of (B.18) diverges to $+\infty$ as $n \to \infty$, we only need to ensure that the second term at the RHS of (B.18) remains bounded in the limit. We have

$$\tilde{\mathcal{L}}^p(\theta_2) - \tilde{\mathcal{L}}^p(\theta_2^\dagger) = \log\left(\frac{y'S'M_ZSy}{y'S^{\dagger\prime}M_ZS^\dagger y}\right) - \frac{2}{n}\log|S(\theta_2)| + \frac{2}{n}\log|S^\dagger| + o_p(1)$$
$$= \log\left(\frac{\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\beta_0'Z'S_0^{-1\prime}S^{\dagger\prime}M_ZS^\dagger S_0^{-1}Z\beta_0}\right) - \frac{2}{n}\log|S(\theta_2)| + \frac{2}{n}\log|S^\dagger| + o_p(1). \quad \text{(B.19)}$$

The first term on the RHS of (B.19) is bounded, since by Lemma 4(a), both numerator and denominator in the argument of the logarithm are $O_p(n||S_0^{-1}||_\infty^2)$, uniformly in θ_2, so that the first term is $O_p(1)$. Also, let $\Omega = (S'S)^{-1}$ and write, for each $\theta_2 \in \Theta_2$,

$$|S|^{2/n} = |S'S|^{1/n} = |\Omega^{-1}|^{1/n} \le \eta_{\max}(\Omega^{-1}) = \frac{1}{\eta_{\min}(\Omega)} < K \quad \text{(B.20)}$$

where the last displayed bound follows under Assumption 4. Similarly,

$$|S|^{2/n} \ge \eta_{\min}(\Omega^{-1}) = \frac{1}{\eta_{\max}(\Omega)} > 0, \quad \text{(B.21)}$$

again under Assumption 4, such that the second and third terms at the rhs of (B.19) remain bounded. We therefore conclude that the RHS of (B.18) increases without bound as $n \to \infty$, when $0 < \gamma \le 1$. Next, we prove that the second term on the RHS of (B.18) tends to zero for $\gamma = 0$. In this case,

$$\tilde{\mathcal{L}}^p(\theta_2) - \tilde{\mathcal{L}}^p(\theta_2^\dagger) = \log\left(\frac{\sigma_0^2 tr(S_0^{-1\prime}S'SS_0^{-1}) + \beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime}S^{\dagger\prime}S^\dagger S_0^{-1}) + \beta_0'Z'S_0^{-1\prime}S^{\dagger\prime}M_ZS^\dagger S_0^{-1}Z\beta_0}\right) + \frac{1}{n}\log\left|S^{\dagger\prime}S^\dagger\,(S'S)^{-1}\right| + o_p(1)$$
$$= \log\left(\frac{tr(S_0^{-1}S_0^{-1\prime}S'S)}{tr(S_0^{-1}S_0^{-1\prime}S^{\dagger\prime}S^\dagger)}\frac{\left(1+\frac{\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime}S'SS_0^{-1})}\right)}{\left(1+\frac{\beta_0'Z'S_0^{-1\prime}S^{\dagger\prime}M_ZS^\dagger S_0^{-1}Z\beta_0}{\sigma_0^2 tr(S_0^{-1}S_0^{-1\prime}S^{\dagger\prime}S^\dagger)}\right)}\right) + \frac{1}{n}\log\left|S^{\dagger\prime}S^\dagger\,(S'S)^{-1}\right|$$
$$= \left\{\log\left(\frac{tr\left(\Omega_0\Omega^{-1}\right)}{tr\left(\Omega_0\Omega^{\dagger-1}\right)}\right) + \frac{1}{n}\log\left|\Omega^{\dagger-1}\Omega\right|\right\} + \log\left(\frac{\left(1+\frac{\beta_0'Z'S_0^{-1\prime}S'M_ZSS_0^{-1}Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime}S'SS_0^{-1})}\right)}{\left(1+\frac{\beta_0'Z'S_0^{-1\prime}S^{\dagger\prime}M_ZS^\dagger S_0^{-1}Z\beta_0}{\sigma_0^2 tr(S_0^{-1}S_0^{-1\prime}S^{\dagger\prime}S^\dagger)}\right)}\right), \quad \text{(B.22)}$$

where $\Omega = (S'S)^{-1}$. The term in the curly brackets in the rhs of (B.22) was shown in Lemma 5 of RL to be as small as as desired. Specifically, fixing $\delta > 0$, there exists $\zeta > 0$ such that for large enough n

$$\sup_{\theta_2:\left\|\theta_2-\theta_2^\dagger\right\|<\eta}\left|\log\left(\frac{tr\left(\Omega_0\Omega^{-1}\right)}{tr\left(\Omega_0\Omega^{\dagger-1}\right)}\right) + \frac{1}{n}\log\left|\Omega^{\dagger-1}\Omega\right|\right| < \delta.$$

Let θ_2^* such that $|\lambda^*-\lambda^\dagger| < |\lambda-\lambda^\dagger|$ and $|w_j^*-w_j^\dagger| < |w_j-w_j^\dagger|$ for each $j = 1, ..., k$, and $S^* = S(\theta_2^*)$, with analogous notation for similar quantities. We consider the argument of the logarithm in the second term at the rhs of (B.22). By the MVT,

$$\frac{\beta_0' Z' S_0^{-1\prime} S' M_Z S S_0^{-1} Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime} S' S S_0^{-1})} = \frac{\beta_0' Z' S_0^{-1\prime} S^{\dagger\prime} M_Z S^{\dagger} S_0^{-1} Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime} S^{\dagger\prime} S^{\dagger} S_0^{-1})} - \frac{2\beta_0' Z' S_0^{-1\prime} C^{*\prime} M_Z S^{*} S_0^{-1} Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime} S^{*\prime} S^{*} S_0^{-1})}(\lambda - \lambda^{\dagger})$$
$$-\frac{2\lambda^*}{\sigma_0^2 tr(S_0^{-1\prime} S^{*\prime} S^{*} S_0^{-1})}\sum_{j=1}^{k}\beta_0' Z' S_0^{-1\prime} C_j^{*\prime} M_Z S^{*} S_0^{-1} Z\beta_0 (w_j - w_j^{\dagger})$$
$$+\frac{2\beta_0' Z' S_0^{-1\prime} S^{*\prime} M_Z S^{*} S_0^{-1} Z\beta_0}{\sigma_0^2 tr^2(S_0^{-1\prime} S^{*\prime} S^{*} S_0^{-1})} tr(S_0^{-1} S_0^{-1\prime}(C^{*\prime} S^{*} + S^{*\prime} C^{*}))(\lambda - \lambda\dagger)$$
$$+\frac{2\beta_0' Z' S_0^{-1\prime} S^{*\prime} M_Z S^{*} S_0^{-1} Z\beta_0}{\sigma_0^2 tr^2(S_0^{-1\prime} S^{*\prime} S^{*} S_0^{-1})}\lambda^* \sum_{j=1}^{k} tr(S_0^{-1} S_0^{-1\prime}(C_j^{*\prime} S^{*} + S^{*\prime} C_j^{*}))(w_j - w_j\dagger). \tag{B.23}$$

Under Assumption 6, in view of Lemmas 1, 4(a) and equation (S.26) of RL,

$$\frac{\beta_0' Z' S_0^{-1\prime} S' M_Z S S_0^{-1} Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime} S' S S_0^{-1})} = \frac{\beta_0' Z' S_0^{-1\prime} S^{\dagger\prime} M_Z S^{\dagger} S_0^{-1} Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime} S^{\dagger\prime} S^{\dagger} S_0^{-1})} + O(\eta) \tag{B.24}$$

over the support $||\theta_2 - \theta_2^{\dagger}|| < \eta$, with the first term on the rhs being $O(1)$, because $\beta_0' Z' S_0^{-1\prime} S' M_Z S S_0^{-1} Z\beta_0 = O(n)$ and $tr(S_0^{-1\prime} S' S S_0^{-1}) = O(n)$, uniformly in Θ_2, by Lemma 4(a) and equation (S.26) of RL. It follows that the second term on the rhs of (B.22) reduces to

$$\log(1 + O(\eta)) = O(\eta) \tag{B.25}$$

and thus, for each $\delta > 0$ we can choose $\eta > 0$ such that

$$\sup_{\theta_2 : \left\|\theta_2 - \theta_2^{\dagger}\right\| < \eta}\left|\log\left(\frac{\left(1 + \frac{\beta_0' Z' S_0^{-1\prime} S' M_Z S S_0^{-1} Z\beta_0}{\sigma_0^2 tr(S_0^{-1\prime} S' S S_0^{-1})}\right)}{\left(1 + \frac{\beta_0' Z' S_0^{-1\prime} S^{\dagger\prime} M_Z S^{\dagger} S_0^{-1} Z\beta_0}{\sigma_0^2 tr(S_0^{-1} S_0^{-1\prime} S^{\dagger\prime} S^{\dagger})}\right)}\right)\right| < \delta. \tag{B.26}$$

We conclude that in the $\gamma = 0$ case the rhs of (B.18) remains strictly positive as $n \to \infty$, as required. ■

Lemma 7. Under Assumptions 1-9,

$$\sup_{\Theta_2}|\mathcal{L}^p(\theta_2) - \tilde{\mathcal{L}}^p(\theta_2)| \xrightarrow{p} 0, \quad \text{as} \quad n \to \infty, \tag{B.27}$$

with $\mathcal{L}^p(\cdot)$ and $\tilde{\mathcal{L}}^p(\cdot)$ defined respectively in (2.14) and (2.25).

Proof of Lemma 7. The approach in the proof of Lemma 7 is similar to that of the proof of Lemma 6 in RL, with substantial differences in the orders of magnitude of the various terms due to the inclusion of the linear part in (1.2). Let $N(\theta_2, \delta)$ a δ-neighborhood of θ_2 such that

$$\mathcal{N}(\theta_2, \delta) = \{\theta_2^\sharp : |\lambda^\sharp - \lambda| < \delta/(k+1), |w_j^\sharp - w_j| < \delta/(k+1) \text{ for each } j = 1, ..., k\}. \tag{B.28}$$

Let $\bar{\theta}_2$ such that: $|\bar{\lambda} - \lambda| < |\lambda^\sharp - \lambda|$, $|\bar{w}_j - w_j| < |w_j^\sharp - w_j|$ for each j. Let $S^\sharp = S(\theta_2^\sharp)$ and $\bar{S} = S(\bar{\theta}_2)$, with analogous notation for $C(\cdot)$ and $C_r(\cdot)$ for $r = 1, ..., k$. Since Θ_2 is compact under Assumption 2, it has a finite sub-covering and we focus on

$$\begin{aligned}\sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} |\mathcal{L}^p(\theta_2^\sharp) - \tilde{\mathcal{L}}^p(\theta_2^\sharp)| \leq & \sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} |\mathcal{L}^p(\theta_2^\sharp) - \mathcal{L}^p(\theta_2)| + |\mathcal{L}^p(\theta_2) - \tilde{\mathcal{L}}(\theta_2)| \\ & + \sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} |\tilde{\mathcal{L}}^p(\theta_2^\sharp) - \tilde{\mathcal{L}}^p(\theta_2)|. \end{aligned} \tag{B.29}$$

Pointwise convergence in probability of $L^p(\theta_2)$ to $\tilde{\mathcal{L}}^p(\theta_2)$ holds by definition of $\tilde{\sigma}^{*2}(\theta_2)$ so that the second term at the RHS of (B.29) is $o_p(1)$.
We start with the first term at the RHS of (B.29). By the mean value theorem, we may write

$$\begin{aligned} y'S^{\sharp\prime}M_Z S^\sharp y =& y'S'M_Z S y + \frac{\partial y'\bar{S}'M_Z\bar{S}y}{\partial\lambda}(\lambda^\sharp - \lambda) + \sum_{j=1}^k \frac{\partial y'\bar{S}'M_Z\bar{S}y}{\partial w_j}(w_j^\sharp - w_j) \\ =& y'S'M_Z S y - 2y'\bar{S}'M_Z\bar{C}y(\lambda^\sharp - \lambda) - 2\bar{\lambda}\sum_{j=1}^k y'\bar{S}'M_Z\bar{C}_j y(w_j^\sharp - w_j), \end{aligned} \tag{B.30}$$

so that

$$\begin{aligned} \Delta_n\left(\theta_2, \theta_2^\sharp\right) \quad : \quad =& \frac{\left|y'S^{\sharp\prime}M_Z S^\sharp y - y'S'M_Z S y\right|}{y'y} = \frac{2}{y'y}|y'\bar{S}'M_Z\bar{C}y(\lambda^\sharp - \lambda) + \bar{\lambda}\sum_{j=1}^k y'\bar{S}'M_Z\bar{C}_j y(w_j^\sharp - w_j)| \\ \leq & \frac{K}{y'y}\left(|y'\bar{S}'M_Z\bar{C}y||(\lambda^\sharp - \lambda)| + \sum_{j=1}^k |y'\bar{S}'M_Z\bar{C}_j y||w_j^\sharp - w_j|\right). \end{aligned} \tag{B.31}$$

In the $0 < \gamma \leq 1$ case, by Lemmas 1,2 and 4, as $\left\|\bar{S}'M_Z\bar{C} + \bar{C}'M_Z\bar{S}\right\|_\infty < K$,

$$y'\bar{S}'M_Z\bar{C}y = \beta_0'Z'S_0^{-1\prime}\frac{\bar{S}'M_Z\bar{C} + \bar{C}'M_Z\bar{S}}{2}S_0^{-1}Z\beta_0 + O_p\left(\max\left(n\left\|S^{-1}\right\|_\infty, \sqrt{n}\left\|S^{-1}\right\|_\infty^2\right)\right) = O_p\left(n\left\|S^{-1}\right\|_\infty^2\right)$$

and similarly,

$$y'y = O_p\left(n\left\|S_0^{-1}\right\|_{\infty}^{2}\right) \text{ and } y'\bar{S}'M_Z\bar{C}_jy = O_p\left(n\|S_0^{-1}\|_{\infty}^2\right).$$

It follows that in this case, for each $\zeta > 0$ there exists a $\delta > 0$ such that

$$\sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} \Delta_n\left(\theta_2, \theta_2^\sharp\right) < \zeta.$$

Next, consider the case $\gamma = 0$. Here,

$$y'\bar{S}'M_Z\bar{C}y = \epsilon' S_0^{-1\prime}\frac{\bar{S}'M_Z\bar{C} + \bar{C}'M_Z\bar{S}}{2}S_0^{-1}\epsilon + \beta_0' Z' S_0^{-1\prime}\frac{\bar{S}'M_Z\bar{C} + \bar{C}'M_Z\bar{S}}{2}S_0^{-1}Z\beta_0 + O_p\left(\sqrt{n}\right) = O_p\left(n\right)$$

and similarly,

$$y'y = O_p\left(n\right) \text{ and } y'\bar{S}'M_Z\bar{C}_jy = O_p\left(n\right).$$

From the definition of stochastic equicontinuity (Andrews (1994)), (B.31) implies that for all $\zeta_1 > 0$ and $\zeta_2 > 0$ there exists a $\delta > 0$ such that

$$\limsup_{n\to\infty} \Pr\left(\sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} \Delta_n\left(\theta_2, \theta_2^\sharp\right) > \zeta_1\right) < \zeta_2, \tag{B.32}$$

where ζ_1, ζ_2 and δ do not depend on θ_2. This proves that the first term of $L^p(\theta_2)$ in (2.14) is stochastic equicontinuous.
Now we consider the second term of $\mathcal{L}^p(\theta_2)$ in (2.14). We have

$$-2\log|S^\sharp| = -2\log|S| + 2tr(\bar{S}^{-1}\bar{C})(\lambda^\sharp - \lambda) + 2\bar{\lambda}\sum_{j=1}^{k} tr(\bar{S}^{-1}\bar{C}_j)(w_j^\sharp - w_j) \tag{B.33}$$

Under Assumption 6, by Lemma 4(e)

$$\left|tr\left(\bar{S}^{-1}\bar{C}\right)\right| = O(n) \quad \text{and} \quad \left|tr\left(\bar{S}^{-1}\bar{C}_j\right)\right| = O(n) \quad \forall j = 1, ..., k. \tag{B.34}$$

Hence, for every $\nu > 0$ there exists a neighborhood $N\left(\theta_2, \delta\right)$ that does not depend on n such that for all $n > N$,

$$\sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} \left|\frac{2\log\left|S^\sharp\right|}{n} - \frac{2\log|S|}{n}\right| \le K\left((\lambda^\sharp - \lambda) + \sum_{j=1}^{k}(w_j^\sharp - w_j)\right) \le K\delta \le \nu. \tag{B.35}$$

Thus, the second term of of $\mathcal{L}^p(\theta_2)$ in (2.14) is uniformly equicontinuous. The implication is that we are done for the first term on the rhs of (B.29) for both the $\gamma = 0$ and $0 < \gamma \leq 1$ cases.

In order to conclude the proof we need to focus on the third term at the RHS of (B.29) and show stochastic equicontinuity of $\tilde{\sigma}^{2*}(\theta_2)$ in (2.25), as equicontinuity of the second term in $\tilde{\mathcal{L}}^p(\theta_2)$ follows as in (B.33) - (B.35). By the MVT, under Assumption 5 and since the module is a continuous function,

$$\begin{aligned}
&\sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} \left| p\lim_{n\to\infty} \left(\frac{-2y'\bar{S}'M_Z\bar{C}y(\lambda^\sharp - \lambda)}{y'y} - \frac{2\bar{\lambda}\sum_{j=1}^{k} y'\bar{S}'M_Z\bar{C}_j y(w_j^\sharp - w_j)}{y'y} \right) \right| \\
\leq & K \sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} p\lim_{n\to\infty} \left(\frac{|y'\bar{S}'M_Z\bar{C}y||\lambda^\sharp - \lambda|}{y'y} + \sum_{j=1}^{k} \frac{|y'\bar{S}'M_Z\bar{C}_j y||w_j^\sharp - w_j|}{y'y} \right) \\
\leq & K\delta \, p\lim_{n\to\infty} \left(\frac{|y'\bar{S}'M_Z\bar{C}y|}{y'y} + \sum_{j=1}^{k} \frac{|y'\bar{S}'M_Z\bar{C}_j y|}{y'y} \right),
\end{aligned} \tag{B.36}$$

where K, as usual, denotes a constant that can change value from step to step. Under Assumption 6, from Lemmas 1,2,4, for each ζ_2 there exists a Δ such that

$$Pr\left(p\lim_{n\to\infty} \left(\frac{|y'\bar{S}'M_Z\bar{C}y|}{y'y} + \sum_{j=1}^{k} \frac{|y'\bar{S}'M_Z\bar{C}_j y|}{y'y} \right) > \Delta \right) < \zeta_2. \tag{B.37}$$

Let $\zeta_1 = \delta K \Delta$. We have

$$\begin{aligned}
&Pr\left(\sup_{\theta_2^\sharp \in \mathcal{N}(\theta_2,\delta)} \left| p\lim_{n\to\infty} \left(\frac{-2y'\bar{S}'M_Z\bar{C}y(\lambda^\sharp - \lambda)}{y'y} - \frac{2\bar{\lambda}\sum_{j=1}^{k} y'\bar{S}'M_Z\bar{C}_j y(w_j^\sharp - w_j)}{y'y} \right) \right| > \zeta_1 \right) \\
\leq & Pr\left(p\lim_{n\to\infty} \left(\frac{|y'\bar{S}'M_Z\bar{C}y|}{y'y} + \sum_{j=1}^{k} \frac{|y'\bar{S}'M_Z\bar{C}_j y|}{y'y} \right) > \Delta \right) < \zeta_2,
\end{aligned} \tag{B.38}$$

concluding the proof. ■

Lemma 8. *Under Assumptions 1-11, we have*

$$\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} = O_p\left(\left(\frac{||S_0^{-1}||_\infty}{n^{1/2}} \right) \right) \quad and \quad \frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_j} = O_p\left(\left(\frac{||S_0^{-1}||_\infty}{n^{1/2}} \right) \right), \; j = 1, ...k \tag{B.39}$$

For $0 \leq \gamma \leq 1$ *in (2.8).*

Proof of Lemma 8. By standard algebra,

$$\begin{aligned}
\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} &= -2\frac{y'C_0'M_Z S_0 y}{y'S_0'M_Z S_0 y} + \frac{2}{n} tr\left(S_0^{-1}C_0\right) \\
&= -2\frac{\epsilon' S_0^{-1\prime} C_0'\epsilon + \beta_0' Z' S_0^{-1\prime} C_0' M_Z \epsilon - \epsilon' S_0^{-1\prime} C_0' Z(Z'Z)^{-1}Z'\epsilon}{\epsilon'\epsilon - \epsilon' Z(Z'Z)^{-1}Z'\epsilon} + \frac{2}{n} tr\left(S_0^{-1}C_0\right) \\
&= -2\frac{\epsilon' S_0^{-1\prime} C_0'\epsilon + \beta_0' Z' S_0^{-1\prime} C_0' M_Z \epsilon - \epsilon' S_0^{-1\prime} C_0' Z(Z'Z)^{-1}Z'\epsilon}{\epsilon'\epsilon - \epsilon' Z(Z'Z)^{-1}Z'\epsilon} + \frac{2}{n} tr\left(S_0^{-1}C_0\right) \\
&= -\left(\frac{\epsilon'\epsilon}{n} + O_p\left(\frac{1}{n}\right)\right)^{-1} \frac{2}{n}\left(\epsilon'\left(S_0^{-1\prime}C_0' - \frac{I}{n}tr(C_0 S_0^{-1})\right)\epsilon + \beta_0' Z' S_0^{-1} C_0' M_Z \epsilon + O_p\left(||S_0^{-1}||_\infty\right)\right),
\end{aligned} \tag{B.40}$$

where the last equality follows since $\epsilon' S_0^{-1} C_0' Z(Z'Z)^{-1}Z'\epsilon = O_p(||S_0^{-1}||_\infty)$ since

$$E\left(\epsilon' S_0^{-1} C_0' Z(Z'Z)^{-1}Z'\epsilon\right) = \sigma_0^2 tr\left(S_0^{-1}C_0' Z(Z'Z)^{-1}Z'\right) = tr\left(Z' S_0^{-1} C_0' Z(Z'Z)^{-1}\right) = O(||S_0^{-1}||_\infty) \tag{B.41}$$

from each component of $Z' S_0^{-1} C_0' Z = O(n||S_0^{-1}||_\infty)$ by Lemma 4(d) and $Z'Z \sim n$. Rearranging terms we obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} &= -2\frac{y'C_0'M_Z S_0 y}{y'S_0'M_Z S_0 y} + \frac{2}{n} tr\left(S_0^{-1}C_0\right) \\
&= -\frac{2}{n}\left(\frac{\epsilon'\epsilon}{n}\right)^{-1}\left(\epsilon'\left(S_0^{-1\prime}C_0' - \frac{I}{n}tr(C_0 S_0^{-1})\right)\epsilon + \beta_0' Z' S_0^{-1} C_0' M_Z \epsilon\right) + O_p\left(\frac{||S_0^{-1}||_\infty}{n}\right) \\
&= O_p\left(\max\left(\frac{||S_0^{-1}||_\infty^{1/2}}{n^{1/2}}, \frac{||S_0^{-1}||_\infty}{n^{1/2}}\right)\right) + O_p\left(\frac{||S_0^{-1}||_\infty}{n}\right),
\end{aligned} \tag{B.42}$$

where both terms in the $\max(\cdot,\cdot)$ contribute as long as $\gamma = 0$ and the second one dominates for $\gamma > 0$. The remainder $O_p\left(||S_0^{-1}||_\infty/n\right)$ vanishes as long as $\gamma < 1$ and it is dominated by the leading term, which is $O_p(||S_0^{-1}||_\infty/\sqrt{n})$, $\forall \gamma \in [0,1]$. Thus, using $(\epsilon'\epsilon/n) = \sigma_0^2 + O_p(1/\sqrt{n})$, for $\gamma = 0$, (B.42) becomes

$$\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} = -\frac{2}{n\sigma_0^2}\left(\epsilon'\left(S_0^{-1\prime}C_0' - \frac{I}{n}tr(C_0 S_0^{-1})\right)\epsilon + \beta_0' Z' S_0^{-1} C_0' M_Z \epsilon\right) + O_p\left(\frac{1}{n}\right), \tag{B.43}$$

while for $0 < \gamma \le 1$ we get

$$\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial \lambda} = -\frac{2}{n\sigma_0^2}\beta_0' Z' S_0^{-1} C_0' M_Z \epsilon + O_p\left(\left(\frac{||S_0^{-1}||_\infty}{n}\right)^{1/2}\right). \tag{B.44}$$

Similarly, for $j = 1, ...k$ and under Assumption 10

$$\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_j} = -2\lambda_0 \frac{y' C'_{j,0} M_Z S_0 y}{y' S'_0 S_0 y} + \frac{2\lambda_0}{n} tr\left(S_0^{-1} C_{j,0}\right)$$
$$= -\frac{2}{n}\left(\frac{\epsilon'\epsilon}{n}\right)^{-1}\left(\epsilon'\left(S_0^{-1\prime} C'_{j,0} - \frac{I}{n} tr(C_{j,0} S_0^{-1})\right)\epsilon + \beta'_0 Z' S_0^{-1} C'_{j,0} M_Z \epsilon\right) + O_p\left(\frac{||S_0||_\infty^{-1}}{n}\right)$$
$$= O_p\left(\max\left(\frac{||S_0^{-1}||_\infty^{1/2}}{n^{1/2}}, \frac{||S_0^{-1}||_\infty}{n^{1/2}}\right)\right) + O_p\left(\frac{||S_0^{-1}||_\infty}{n}\right), \tag{B.45}$$

which becomes

$$\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_j} = -\frac{2}{n\sigma_0^2}\left(\epsilon'\left(S_0^{-1\prime} C'_{j,0} - \frac{I}{n} tr(C_{j,0} S_0^{-1})\right)\epsilon + \beta'_0 Z' S_0^{-1} C'_{j,0} M_Z \epsilon\right) + O_p\left(\frac{1}{n}\right) \tag{B.46}$$

for $\gamma = 0$, and

$$\frac{\partial \mathcal{L}^p(\theta_{20})}{\partial w_j} = -\frac{2}{n\sigma_0^2}\beta'_0 Z' S_0^{-1} C'_{j,0} M_Z \epsilon + O_p\left(\left(\frac{||S_0^{-1}||_\infty}{n}\right)^{1/2}\right) \tag{B.47}$$

for $0 < \gamma \leq 1$. ■

Lemma 9. *Under Assumptions 1-11,*

$$\frac{1}{||S_0^{-1}||_\infty^2} \frac{\partial^2 \mathcal{L}^p(\theta_{20})}{\partial \theta_2 \partial \theta'_2} \xrightarrow{p} D_0 > 0, \tag{B.48}$$

where the elements of D_0 are given in (A.12), (A.13) and (A.14).

Proof of Lemma 9. Let, as usual, $S = S(\theta_2)$, with the same notation for similar quantities. Using standard algebra we derive

$$\frac{1}{||S_0^{-1}||_\infty^2} \frac{\partial^2 \mathcal{L}^p(\theta_{20})}{\partial \lambda^2} = \frac{2}{||S_0^{-1}||_\infty^2} \frac{y' C'_0 M_Z C_0 y/n}{y' S'_0 M_Z S_0 y/n} - \frac{4}{||S_0^{-1}||_\infty^2} \frac{(y' C'_0 M_Z S_0 y/n)^2}{(y' S'_0 M_Z S_0 y/n)^2} + \frac{2}{n||S_0^{-1}||_\infty^2} tr\left((S_0^{-1} C_0)^2\right) \tag{B.49}$$

$$\frac{1}{||S_0^{-1}||_\infty^2} \frac{\partial^2 \mathcal{L}^p(\theta_{20})}{\partial w_i \partial w_j} = -\frac{2\lambda_0}{||S_0^{-1}||_\infty^2} \frac{y' C'_{ij,0} M_Z S_0 y/n}{y' S'_0 M_Z S_0 y/n} + \frac{2\lambda_0^2}{||S_0^{-1}||_\infty^2} \frac{y' C'_{j,0} M_Z C_{i,0} y/n}{y' S'_0 M_Z S_0 y/n}$$
$$- \frac{4\lambda_0^2}{||S_0^{-1}||_\infty^2} \frac{y' C'_{j,0} M_Z S_0 y y' C'_{i,0} M_Z S_0 y/n}{(y' S'_0 M_Z S_0 y/n)^2}$$
$$+ \frac{2\lambda_0}{n||S_0^{-1}||_\infty^2} tr\left(S_0^{-1} C_{ij,0}\right) + \frac{2\lambda_0^2}{n||S_0^{-1}||_\infty^2} tr\left(S_0^{-1} C_{i,0} S_0^{-1} C_{j,0}\right) \tag{B.50}$$

and

$$\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^2\mathcal{L}^p(\theta_{20})}{\partial w_i\partial\lambda} = -\frac{2}{||S_0^{-1}||_\infty^2}\frac{y'C_{i,0}'M_ZS_0y/n}{y'S_0'M_ZS_0y/n} + \frac{2\lambda_0}{||S_0^{-1}||_\infty^2}\frac{y'C_0'M_ZC_{i,0}y/n}{y'S_0'M_ZS_0y/n}$$
$$+\frac{4\lambda_0}{||S_0^{-1}||_\infty^2}\frac{y'C_0'M_ZS_0yy'C_{i,0}'M_ZS_0y/n^2}{(y'S_0'M_ZS_0y/n)^2}$$
$$+\frac{2}{n||S_0^{-1}||_\infty^2}tr\left(S_0^{-1}C_{i,0}\right) + \frac{2\lambda_0}{n||S_0^{-1}||_\infty^2}tr\left(S_0^{-1}C_{i,0}S_0^{-1}C_0\right). \tag{B.51}$$

The denominator in (B.49), (B.50) and (B.51) is

$$\frac{1}{n}y'S_0'M_ZS_0y = \frac{1}{n}\epsilon'M_Z\epsilon = \sigma_0^2 + O_p\left(\frac{1}{n}\right) \tag{B.52}$$

We focus on (B.49), although the same argument can be applied to (B.50) and (B.51). The numerator of the first term in (B.49) is

$$\frac{2}{n||S_0^{-1}||_\infty^2}y'C_0'M_ZC_0y = \frac{2}{n||S_0^{-1}||_\infty^2}\left(\epsilon'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}\epsilon + \beta_0'Z'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}Z\beta_0\right)$$
$$+\frac{4}{n||S_0^{-1}||_\infty^2}\epsilon'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}Z\beta_0$$
$$= \frac{2}{n||S_0^{-1}||_\infty^2}\left(\epsilon'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}\epsilon + \beta_0'Z'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}Z\beta_0\right) + O_p\left(\frac{1}{\sqrt{n}}\right), \tag{B.53}$$

where the last equality follows from Lemma 2, Lemma 4(a) and Lemma 4(b). Furthermore, the first term at the RHS of the last displayed expression is

$$\frac{2}{n||S_0^{-1}||_\infty^2}\epsilon'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}\epsilon = \frac{2}{n||S_0^{-1}||_\infty^2}\epsilon'S_0^{-1\prime}C_0'C_0S_0^{-1}\epsilon - \frac{2}{n||S_0^{-1}||_\infty^2}\epsilon'S_0^{-1\prime}C_0'Z(Z'Z)^{-1}ZC_0S_0^{-1}\epsilon$$
$$= \frac{2}{n||S_0^{-1}||_\infty^2}\epsilon'S_0^{-1\prime}C_0'C_0S_0^{-1}\epsilon + O_p\left(\frac{1}{n}\right). \tag{B.54}$$

where the last equality follows from an argument similar to that applied to derive (B.41). Therefore, the numerator of first term in (B.49) can be written as

$$\frac{2}{n||S_0^{-1}||_\infty^2}y'C_0'M_ZC_0y = \frac{2}{n||S_0^{-1}||_\infty^2}\left(\epsilon'S_0^{-1\prime}C_0'C_0S_0^{-1}\epsilon + \beta_0'Z'S_0^{-1\prime}C_0'M_ZC_0S_0^{-1}Z\beta_0\right) + O_p\left(\frac{1}{\sqrt{n}}\right), \tag{B.55}$$

with the first term being $O_p(1/||S_0^{-1}||_\infty)$ by Lemma 2(b), and the second term being $O_p(1)$ by Lemma 4(a). Thus, it is only when $\gamma = 0$ that the first term in the last displayed equation does not vanish and, for $n \to \infty^4$, by Lemma 3,

$$\frac{2}{n||S_0^{-1}||_\infty^2} y' C_0' M_Z C_0 y \underset{p}{\to} \lim_{n\to\infty} \frac{2\sigma_0^2}{n||S_0^{-1}||_\infty^2} tr\left(S_0^{-1\prime} C_0' C_0 S_0^{-1}\right) + \lim_{n\to\infty} \frac{2}{n||S_0^{-1}||_\infty^2} \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S_0^{-1} Z \beta_0. \tag{B.56}$$

When $\gamma > 0$, as $n \to \infty$,

$$\frac{2}{n||S_0^{-1}||_\infty^2} y' C_0' M_Z C_0 y \underset{p}{\to} \lim_{n\to\infty} \frac{2}{n||S_0^{-1}||_\infty^2} \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S_0^{-1} Z \beta_0. \tag{B.57}$$

By combining (B.52) and (B.56)/(B.57), we conclude that the first term in (B.49) is $O_p(1)$ and it converges to

$$\lim_{n\to\infty} \frac{2}{n||S_0^{-1}||_\infty^2} tr\left(S_0^{-1\prime} C_0' C_0 S_0^{-1}\right) + \lim_{n\to\infty} \frac{2}{n||S_0^{-1}||_\infty^2 \sigma_0^2} \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S_0^{-1} Z \beta_0 \text{ for } \gamma = 0 \tag{B.58}$$

and

$$\lim_{n\to\infty} \frac{2}{n||S_0^{-1}||_\infty^2 \sigma_0^2} \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S_0^{-1} Z \beta_0 \text{ for } \gamma > 0. \tag{B.59}$$

Under Assumptions 6 and 10, the square root of the numerator of second term in (B.49) involves

$$\frac{1}{n||S_0^{-1}||_\infty} y' C_0' M_Z S_0 y = \frac{1}{n||S_0^{-1}||_\infty} \epsilon' S_0^{-1\prime} C_0' \epsilon - \frac{1}{n||S_0^{-1}||_\infty} \epsilon' S_0^{-1\prime} C_0' Z (Z'Z)^{-1} Z' \epsilon + O_p\left(\frac{1}{\sqrt{n}}\right),$$

where the last displayed equality follows since

$$\frac{1}{n||S_0^{-1}||_\infty} \epsilon' S_0^{-1\prime} C_0' \epsilon = O_p\left(\frac{1}{||S_0^{-1}||_\infty}\right) \text{ by Lemma 2(a),} \tag{B.60}$$

$$\frac{1}{n||S_0^{-1}||_\infty} \epsilon' S_0^{-1\prime} C_0' Z (Z'Z)^{-1} Z' \epsilon = O_p\left(\frac{1}{n}\right) \text{ by (B.41),} \tag{B.61}$$

and

$$\frac{1}{n||S_0^{-1}||_\infty} \beta_0' Z' S_0^{-1\prime} M_Z \epsilon = O_p\left(\frac{1}{\sqrt{n}}\right) \text{ by Lemma 4(c).}$$

Therefore, as $n \to \infty$, by Lemma 5,

$$\frac{1}{n||S_0^{-1}||_\infty} y' C_0' M_Z S_0 y \underset{p}{\to} \lim_{n\to\infty} \frac{\sigma_0^2}{n||S_0^{-1}||_\infty} tr\left(S_0^{-1\prime} C_0'\right) \text{ for } \gamma = 0 \tag{B.62}$$

[4] Note that the term in $||S_0^{-1}||_\infty^2 = O(1)$ for $\gamma = 0$, but it is retained to deal with both cases in a unified approach.

and

$$\frac{1}{n||S_0^{-1}||_\infty} y'C_0'M_Z S_0 y \underset{p}{\rightarrow} 0 \text{ for } \gamma > 0. \tag{B.63}$$

Collecting (B.49), (B.52) and (B.62)/(B.63), the second term in (B.49) is $O_p(1)$ for $\gamma = 0$ and it converges to

$$\lim_{n\to\infty} \frac{4}{n^2||S_0^{-1}||_\infty^2} tr^2(C_0 S_0^{-1}), \tag{B.64}$$

while it is $o_p(1)$ for $\gamma > 0$.

A similar argument follows for the third term of (B.49), since, by Lemma 4(f), $tr\left((S_0^{-1}C_0)^2\right) = O_p(n||S_0^{-1}||_\infty)$ and thus

$$\frac{1}{n||S_0^{-1}||_\infty^2} tr\left((S_0^{-1}C_0)^2\right) = O\left(\frac{1}{||S_0^{-1}||_\infty}\right), \tag{B.65}$$

which is $O_p(1)$ for $\gamma = 0$ and vanishes otherwise.

Thus, collecting (B.58)/(B.59), (B.64) and (B.65), and by standard algebra, we obtain,

$$\begin{aligned} \frac{1}{||S_0^{-1}||_\infty^2} \frac{\partial^2 \mathcal{L}^p(\bar{\theta}_2)}{\partial \lambda^2} \overset{p}{\rightarrow} & \lim_{n\to\infty} \frac{1}{n||S_0^{-1}||_\infty^2} tr\left(\left(S_0^{-1\prime}C_0' + C_0 S_0^{-1} - tr(C_0 S_0^{-1})\frac{2I}{n}\right)^2\right) \\ & + \lim_{n\to\infty} \frac{2\beta_0' Z' S_0^{-1\prime} C_0' M_Z C_0 S_0^{-1} Z\beta_0}{\sigma_0^2 n ||S_0^{-1}||_\infty^2}, \end{aligned} \tag{B.66}$$

with the first term vanishing for $\gamma > 0$.

In order to avoid repetition we omit a similar argument for (B.50) and (B.51) and by standard algebra of quadratic forms in i.i.d. random variables, we obtain

$$\begin{aligned} \frac{1}{||S_0^{-1}||_\infty^2} \frac{\partial^2 \mathcal{L}^p(\bar{\theta}_2)}{\partial w_i \partial w_j} \underset{p}{\rightarrow} & \lim_{n\to} \frac{2\lambda_0}{n||S_0^{-1}||_\infty^2} \left(2tr(C_{ij,0}S_0^{-1}) - \lambda_0 tr(S_0^{-1\prime}C_{j,0}' C_{i,0} S_0^{-1}) + tr(S_0^{-1}C_{i,0}S_0^{-1}C_{j,0})\right) \\ & - \lim_{n\to} \frac{4\lambda_0^2}{n^2||S_0^{-1}||_\infty^2} tr(C_{j,0}S_0^{-1}) tr(C_{i,0}S_0^{-1}) + \lim_{n\to} \frac{2\lambda_0^2}{n||S_0^{-1}||_\infty^2} \beta_0' Z' S_0^{-1\prime} C_{j,0}' M_Z C_{i,0} S_0^{-1} Z\beta_0, \end{aligned} \tag{B.67}$$

where all terms contribute for $\gamma = 0$ and only the last one does not vanish if $\gamma > 0$. Also,

$$\begin{aligned} \frac{1}{||S_0^{-1}||_\infty^2} \frac{\partial^2 \mathcal{L}^p(\bar{\theta}_2)}{\partial w_i \partial \lambda} \underset{p}{\rightarrow} & \lim \frac{2}{n||S_0^{-1}||_\infty^2} \left(\lambda_0 tr(S_0^{-1\prime}C_0' C_{i,0} S_0^{-1}) + tr(S_0^{-1}C_{i,0}S_0^{-1}C_0)\right) \\ & - \lim_{n\to} \frac{4\lambda_0}{n^2||S_0^{-1}||_\infty^2} tr(C_0 S_0^{-1}) tr(C_{i,0}S_0^{-1}) + \lim_{n\to} \frac{2\lambda_0}{n||S_0^{-1}||_\infty^2} \beta_0' Z' S_0^{-1\prime} C_0' M_Z C_{i,0} S_0^{-1} Z\beta_0, \end{aligned} \tag{B.68}$$

where again all terms contribute for $\gamma = 0$ and only the last one does not vanish if $\gamma > 0$.

Thus, the elements of D_0 reduces to $\tilde{D}_0$ in (A.15) if $\gamma > 0$ in (2.8). ■

Lemma 10. *Under Assumptions 1-11,*

$$\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^3 \mathcal{L}^p(\bar{\theta}_2)}{\partial \theta_2^3} = O_p(1) \tag{B.69}$$

Proof of Lemma 10. We present in detail the argument for $\partial^3\mathcal{L}(\bar{\theta}_2)/\partial\lambda^3$, although a similar argument follows for the remaining derivatives under Assumption 10. By standard algebra, we derive

$$\begin{aligned}\frac{1}{||S_0^{-1}||_\infty^2}\frac{\partial^3\mathcal{L}(\bar{\theta}_2)}{\partial\lambda^3} =& \frac{1}{||S_0^{-1}||_\infty^2}\frac{4\left(y'\bar{C}'M_Z\bar{C}y\right)^2}{\left(y'\bar{S}'M_Z\bar{S}y\right)^2} + \frac{1}{||S_0^{-1}||_\infty^2}\frac{8y'\bar{C}'M_Z\bar{C}yy'\bar{C}'M_Z\bar{S}y}{\left(y'\bar{S}'M_Z\bar{S}y\right)^2} - \frac{1}{||S_0^{-1}||_\infty^2}\frac{\left(y'\bar{C}'M_Z\bar{S}y\right)^3}{\left(y'\bar{S}'M_Z\bar{S}y\right)^3}\\ &+\frac{4}{n||S_0^{-1}||_\infty}tr\left((\bar{S}^{-1}\bar{C})^3\right)\end{aligned} \tag{B.70}$$

From (2.3) and from Lemmas 2(b), 4(a) and 4(b), the quadratic forms in y at numerator and denominator appearing in first three terms at the rhs of (B.70) are $O_p(n||S_0^{-1}||_\infty)$, such that the first three terms of (B.70) are $O_p(1/||S_0^{-1}||_\infty)$. The last term in (B.70) is $O_p(1)$ from Lemma 4(g).■

Proof of (A.34)

We start by showing (A.34) by proving, equivalently,

$$\sum_{i=1}^n \mathbb{E}\left(v_i^2|\epsilon_j, j<i\right) - \zeta'\Omega^{-1/2}\Omega\Omega^{-1/2}\zeta \underset{p}{\rightarrow} 0, \tag{B.71}$$

which is

$$\begin{aligned}&\zeta'\Omega^{-1/2}\left(\frac{1}{n||S_0^{-1}||_\infty^2}\sum_i \mathbb{E}(u_iu_i'|\epsilon_j, j<i) - \Omega\right)\Omega^{-1/2}\zeta\\
=&\frac{4}{n||S_0^{-1}||_\infty^2}\zeta'\Omega^{-1/2}\left(\sigma_0^2\sum_i\left(\sum_{j<i}\psi_{ij}\epsilon_j\right)\left(\sum_{j<i}\psi_{ij}\epsilon_j\right)' - \sigma_0^4\sum_i\sum_{j<i}\psi_{ij}\psi_{ij}'\right.\\
&\left.+\mu_0^{(3)}\sum_i\left(\psi_{ii}-\frac{1}{n}\sum_t\psi_{tt}\right)\sum_{j<i}\psi_{ij}'\epsilon_j\right)\Omega^{-1/2}\zeta\\
&+\frac{4}{n||S_0^{-1}||_\infty^2}\zeta'\Omega^{-1/2}\left(\sigma_0^2\sum_i\sum_j\sum_{t<i}\beta_0'z_j(\phi_{ij}\psi_{it}'+\psi_{it}\phi_{ij}')\epsilon_t\right)\Omega^{-1/2}\zeta \underset{p}{\rightarrow} 0.\end{aligned}$$

Since $\Omega = O(1)$ as n increases and it is non singular in the limit under Assumption 11, we need to show (for a typical element of the following matrices) that

$$\frac{1}{n||S_0^{-1}||_\infty^2}\left(\sigma_0^2\sum_i\left(\sum_{j<i}\psi_{ij}\epsilon_j\right)\left(\sum_{j<i}\psi_{ij}\epsilon_j\right)' - \sigma_0^4\sum_i\sum_{j<i}\psi_{ij}\psi_{ij}'\right)\underset{p}{\rightarrow} 0, \tag{B.72}$$

and

$$\frac{1}{n||S_0^{-1}||_\infty^2}\mu_0^{(3)}\sum_i\left(\psi_{ii}-\frac{1}{n}\sum_t\psi_{tt}\right)\sum_{j<i}\psi_{ij}'\epsilon_j\underset{p}{\rightarrow} 0 \tag{B.73}$$

and

$$\frac{\sigma_0^2}{n||S_0^{-1}||_\infty^2}\sum_i\sum_j\sum_{t<i}\beta_0' z_j(\phi_{ij}\psi_{it}'+\psi_{it}\phi_{ij}')\epsilon_t\underset{p}{\rightarrow} 0 \tag{B.74}$$

We begin by showing (B.72). We consider the typical elements of the lhs of (B.72)

$$\frac{\sigma_0^2}{n||S_0^{-1}||_\infty^2}\left(\sum_i\sum_{j<i}\psi_{sij}^2(\epsilon_j^2-\sigma_0^2)+\sum_i\sum_{\substack{j,k<i\\ j\neq k}}\psi_{sij}\psi_{sik}\epsilon_j\epsilon_k\right),\quad s=1,.....,k+1, \tag{B.75}$$

and

$$\frac{\sigma_0^2}{n||S_0^{-1}||_\infty^2}\left(\sum_i\sum_{j<i}\psi_{sij}\psi_{tij}(\epsilon_j^2-\sigma_0^2)+\sum_i\sum_{\substack{j,k<i\\ j\neq k}}\psi_{sij}\psi_{tik}\epsilon_j\epsilon_k\right),\quad s,t=1,...,k+1,\ s\neq t. \tag{B.76}$$

The first term in (B.75) has mean zero and variance bounded by

$$\begin{aligned}\frac{K}{n^2||S_0^{-1}||_\infty^4}\sum_i\sum_k\sum_{j<i,k}\psi_{sij}^2\psi_{skj}^2 &\leq \frac{K}{n^2||S_0^{-1}||_\infty^4}\sum_i\sum_k\sum_j\psi_{sij}^2\psi_{skj}^2\\ &\leq \frac{K}{n^2||S_0^{-1}||_\infty^4}\left(\max_j\sum_i\psi_{sij}^2\right)\sum_k\sum_j\psi_{skj}^2 = O\left(\frac{1}{n||S_0^{-1}||_\infty}\right),\end{aligned} \tag{B.77}$$

where the last equality follows from Lemma 4(f), since

$$\sum_k\sum_j\psi_{skj}^2=\frac{1}{4}tr((C_0S_0^{-1}+S_0^{-1\prime}C_0')^2),\quad \text{or}\quad \sum_k\sum_j\psi_{skj}^2=\frac{1}{4}tr((C_{j,0}S_0^{-1}+S_0^{-1\prime}C_{j,0}')^2)\quad j=1,....k, \tag{B.78}$$

and letting e_j to denote the $n \times 1$ vector with 1 in the $j-$th position and zero elsewhere,

$$\sum_i \psi_{sij}^2 = e_j' \Psi_s^2 e_j \le ||\Psi_s||^2 = O(||S_0^{-1}||_\infty^2). \tag{B.79}$$

By Markov's inequality, for each $0 \le \gamma \le 1$ in (2.8), the first term in (B.75) is $o_p(1)$.

The second term in (B.75) has again mean zero and variance bounded by

$$\begin{aligned}
&\frac{K}{n^2||S_0^{-1}||_\infty^4}\left(\sum_i\sum_p\sum_{j<i,p}\sum_{k<i,p}|\psi_{sij}\psi_{sik}\psi_{spj}\psi_{spk}|\right)\\
\le&\frac{K}{n^2||S_0^{-1}||_\infty^4}\sum_i\sum_p\sum_j\sum_k|\psi_{sij}\psi_{sik}|\left(\psi_{spj}^2+\psi_{spk}^2\right)\\
\le&\frac{K}{n^2||S_0^{-1}||_\infty^4}\left(\sup_i\sum_k|\psi_{sik}|\right)\left(\sup_j\sum_i|\psi_{sij}|\right)\sum_p\sum_j\psi_{spj}^2\\
+&\frac{K}{n^2||S_0^{-1}||_\infty^4}\left(\sup_i\sum_j|\psi_{sij}|\right)\left(\sup_k\sum_i|\psi_{sik}|\right)\sum_p\sum_k\psi_{spk}^2=O\left(\frac{1}{n||S_0^{-1}||_\infty}\right),
\end{aligned} \tag{B.80}$$

again from Lemma 4(f) and since, for $s = 1, ..., k+1$, $||\Psi_s||_\infty = O(||S_0^{-1}||_\infty)$. For each $0 \le \gamma \le 1$, we conclude that the second term in (B.75) is $o_p(1)$.

The proof that (B.76) is $o_p(1)$ is virtually identical and it is omitted to avoid repetitions. We prove (B.73) by observing that the the typical element at the lhs of (B.73) is

$$\frac{1}{n||S_0^{-1}||_\infty^2}\mu_0^{(3)}\sum_i\tilde{\psi}_{sii}\sum_{j<i}\psi_{tij}\epsilon_j, \quad s,t = 1, ..., k+1, \tag{B.81}$$

where

$$\tilde{\psi}_{sii} = \psi_{sii} - \frac{1}{n}\sum_t \psi_{stt}. \tag{B.82}$$

The term in (B.81) has mean zero and variance bounded by

$$\begin{aligned}
&\frac{K}{n^2||S_0^{-1}||_\infty^4}\sum_i\sum_k\sum_{j<i,k}|\tilde{\psi}_{sii}\tilde{\psi}_{skk}\psi_{tij}\psi_{tkj}| \le \frac{K}{n^2||S_0^{-1}||_\infty^4}\sum_i\sum_k\sum_j|\psi_{tij}||\psi_{tkj}|(\tilde{\psi}_{sii}^2+\tilde{\psi}_{skk}^2)\\
&\le \frac{K}{n^2||S_0^{-1}||_\infty^4}\left(\left(\sup_j\sum_k|\psi_{tkj}|\right)\left(\sup_i\sum_j|\psi_{tij}|\right)\sum_i\tilde{\psi}_{sii}^2+\left(\sup_k\sum_j|\psi_{tkj}|\right)\left(\sup_j\sum_i|\psi_{tij}|\right)\sum_k\tilde{\psi}_{skk}^2\right)\\
=&O\left(\frac{1}{n||S_0^{-1}||_\infty}\right),
\end{aligned} \tag{B.83}$$

where the last equality follows since

$$\sum_i \tilde{\psi}_{sii}^2 \leq \sum_i \sum_j \tilde{\psi}_{sij}^2 = tr(\tilde{\Psi}_s^2) = tr\left(\left(\Psi_s - tr(\Psi_s)\frac{I}{n}\right)^2\right) = O\left(\max(n, n||S_0^{-1}||_\infty)\right) = O(n||S_0^{-1}||_\infty), \tag{B.84}$$

from Lemma 4(e) and Lemma 4(f). Thus, for all $0 \leq \gamma \leq 1$ the term in (B.73) is thus $o_p(1)$.

We finally need to show (B.74) for a typical element, i.e. we consider

$$\frac{\sigma_0^2}{n||S_0^{-1}||_\infty^2} \sum_i \sum_j \sum_{t<i} \beta_0' z_j \phi_{sij} \psi_{vit} \epsilon_t \quad \text{for} \quad s, t = 1, ..., k+1. \tag{B.85}$$

The latter has mean zero and variance bounded by

$$\frac{K}{n^2||S_0^{-1}||_\infty^4} \left| \sum_i \sum_j \sum_u \sum_h \sum_{t<i,u} \phi_{sij} \phi_{suh} \psi_{vit} \psi_{vut} \right| \leq \frac{K}{n^2||S_0^{-1}||_\infty^4} \sum_i \sum_j \sum_u \sum_h \sum_t |\phi_{sij} \phi_{suh} \psi_{vit} \psi_{vut}|$$
$$\frac{K}{n||S_0^{-1}||_\infty^4} \sup_i \sum_j |\phi_{sij}| \sup_u \sum_h |\phi_{shu}| \sup_t \sum_i |\psi_{vit}| \sup_u \sum_t |\psi_{vut}| = O\left(\frac{1}{n}\right) \tag{B.86}$$

where the last equality follows since from Lemma 1 and basic norm inequalities, we have

$$||\Psi_v||_\infty = O(||S_0^{-1}||_\infty) \qquad ||\Phi_s||_\infty = O(||S_0^{-1}||_\infty) \quad \text{for} \quad s, v = 1, ..., k+1. \tag{B.87}$$

By Markov's inequality (B.74) holds, concluding the proof of (A.34). ■

Proof of (A.35)

We prove (A.35) by showing the sufficient Lyapunov condition

$$\sum_i \mathbb{E}|v_i|^{2+\delta} \to 0, \quad \text{for some} \quad \delta > 0 \tag{B.88}$$

and showing, for a typical standardized element of u_i, $s = 1, ..., k+1$,

$$\left(\frac{1}{(n||S_0^{-1}||_\infty^2)^{1/2}}\right)^{2+\delta} \sum_i \mathbb{E}|u_{si}|^{2+\delta} = \left(\frac{1}{(n||S_0^{-1}||_\infty^2)^{1/2}}\right)^{2+\delta} \sum_i \mathbb{E}\left(\mathbb{E}|u_{si}|^{2+\delta}|\epsilon_j, j < i\right) \to 0. \tag{B.89}$$

We have, by the c_r inequality,

$$\left(\frac{1}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\sum_i \mathbb{E}\left(\mathbb{E}|u_{si}|^{2+\delta}|\epsilon_j, j<i\right)$$
$$\leq\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\sum_i|\tilde{\psi}_{sii}|^{2+\delta}+\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\sum_i\mathbb{E}|\sum_{j<i}\psi_{sij}\epsilon_j|^{2+\delta}$$
$$+\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\sum_i\left|\sum_j\beta_0'z_j\phi_{sij}\right|^{2+\delta}. \tag{B.90}$$

The first term of (B.90) is

$$\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\left(\sup_i|\tilde{\psi}_{sii}|^\delta\right)\sum_i\tilde{\psi}_{sii}^2=O\left(\frac{1}{n^{\delta/2}||S_0^{-1}||_\infty^{1+\delta}}\right)=o(1), \tag{B.91}$$

since the second factor is $O(1)$, given (B.8) and Lemma 4(e), and the third factor is $O(n||S_0^{-1}||_\infty)$ from (B.84).

The second term in (B.90), by the Burkholder and Von Bahr/Esseen inequality, is bounded by

$$\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\sum_i\mathbb{E}|\sum_{j<i}\psi_{sij}^2\epsilon_j^2|^{1+\delta/2}\leq\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\sum_i\sum_{j<i}\psi_{sij}^{2+\delta}$$
$$\leq\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\sum_i\left(\sum_{j<i}\psi_{sij}^2\right)^{1+\delta/2}\leq\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2}\left(\sup_i\sum_j\psi_{sij}^2\right)^{\delta/2}\sum_i\sum_j\psi_{sij}^2$$
$$=O\left(\frac{1}{n^{\delta/2}||S_0^{-1}||_\infty}\right)=o(1) \tag{B.92}$$

using (B.78), Lemma 4(f) and (B.79).

We show that the third term in (B.90) is $o(1)$ by observing that, under Assumption 7,

$$|\sum_j\beta_0'z_j\phi_{sij}|^{2+\delta}\leq K\sup_j|\beta_0'z_j|^{2+\delta}\sum_j|\phi_{sij}|^{2+\delta} \tag{B.93}$$

and thus the third term in (B.90) is bounded by

$$\left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2} K\sum_i\sum_j|\phi_{sij}|^{2+\delta} \leq \left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2} K\sum_i\left(\sum_j\phi_{sij}^2\right)^{1+\delta/2}$$

$$\leq \left(\frac{K}{n||S_0^{-1}||_\infty^2}\right)^{1+\delta/2} K\left(\sup_i\sum_j\phi_{sij}^2\right)^{\delta/2}\sum_i\sum_j\phi_{sij}^2 = O\left(\frac{1}{n^{\delta/2}||S_0^{-1}||_\infty}\right) \tag{B.94}$$

where the last equality follows from an argument identical to that used to derive (B.78) and (B.79), using again Lemma 4(f). ■

References

[1] Andrews, D.W.K (1994), Asymptotics for semiparametric econometric models via stochastic equicontinuity, Econometrica, 62, 43-72.

[2] Bailey, N., S. Holly and M.H. Pesaran (2016), A two stage approach to spatiotemporal analysis with strong and weak cross-sectional dependence, Journal of Applied Econometrics, 3, 249-280.

[3] Baltagi, B.H., C. Kao and L. Liu (2013), The Estimation and Testing of a Linear Regression with Near Unit Root in the Spatial Autoregressive Error Term, Spatial Economic Analysis, 8, 241-270.

[4] Chan, N. H. and C. Z. Wei (1987), Asymptotic inference for nearly nonstationary AR(1) processes, Annals of Statistics, 15, 1050–1063.

[5] Chudik A and M.H. Pesaran (2015), Large panel data models with cross-sectional dependence: a survey. In The Oxford Handbook of Panel Data, Baltagi B. (ed), 3- 45. Oxford University Press, New York.

[6] Delgado,M. A. and P. M. Robinson (2015), Non-nested testing of spatial correlation, Journal of Econometrics, 187, 385-401.

[7] Gayer G., I. Gilboa and O. Lieberman (2007), Empirical similarity, B.E. Journal in Theoretical Economics, 7(1)(Advances), Article 10.

[8] Gayer G., O. Lieberman and O. Yaffe (2019), Similarity-based model for ordered categorical data, Econometric Reviews, 38, 263-278.

[9] Gilboa, I, O. Lieberman and D. Schmeidler (2006), Empirical similarity, Review of Economics and Statistics, 88, 433-444.

[10] Gilboa, I., O. Lieberman and D. Schmeidler (2010), On the definition of objective probabilities by empirical similarity, Synthese, 172, 79-95.

[11] Gilboa, I., O. Lieberman and D. Schmeidler (2011), A similarity-based approach to prediction, Journal of Econometrics, 162, 124-131.

[12] Gupta, A. and P.M. Robinson (2015), Inference on higher-order spatial autoregressive models with increasingly many parameters, Journal of Econometrics, 186 (1), 19-31.

[13] Gupta, A. and P.M. Robinson (2018), Pseudo maximum likelihood estimation of spatial autoregressive models with increasing dimension, Journal of Econometrics, 202 (1), 92-107.

[14] Hamilton, J.D. (1994), Time Series Analysis. Princeton University Press, New Jersey.

[15] Kapetanios, G., J. Mitchell and Y. Shin, (2013), A nonlinear panel data model of cross-sectional dependence, Journal of Econometrics, 179, 134-157.

[16] Kelejian, H.H. and I.R. Prucha (1998), A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, Journal of Real Estate Finance and Economics, 17, 99-121.

[17] Kelejian, H.H. and I.R. Prucha (1999), A generalized moments estimator for the autoregressive parameter in a spatial model, International Economic Review, 40, 509-533.

[18] Kelejian H. H. and I.R. Prucha (2010), Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, Journal of Econometrics, 157, 53-67.

[19] Lee, L.F. (2004), Asymptotic distribution of quasi-maximum likelihood estimates for spatial autoregressive models, Econometrica, 72, 1899-1925.

[20] Lee, L.F. and J. Yu (2013), Near Unit Root in the Spatial Autoregressive Model, Spatial Economic Analysis, 8, 314-351.

[21] Lieberman, O. (2010), Asymptotic theory for empirical similarity models, Econometric Theory, 26, 1032-1059.

[22] Lieberman, O. (2012), A similarity-based approach to time-varying coefficient nonstationary autoregression, Journal of Time Series Analysis, 33, 484-502.

[23] Lieberman, O. and P.C.B. Phillips, (2014), Norming rates and limit theory for some time-varying coefficient autoregressions, Journal of Time Series Analysis, 35, 592-623.

[24] Phillips, P.C.B. (1987), Towards a unified asymptotic theory for autoregression, Biometrika, 74, 535-547.

[25] Phillips, P.C.B. and T. Magdalinos (2007), Limit theory for moderate deviations from a unit root, Journal of Econometrics, 136, 115-130.

[26] Robinson P.M (2011), Asymptotic theory for nonparametric regression with spatial data, Journal of Econometrics, 165, 5-19.

[27] Rossi, F. and O. Lieberman (2021), Inference in a spatial autoregressive model with an extended coefficient range and a similarity-based weight matrix, submitted to publication.

[28] Scott, D.J. (1973), Central limit theorems for martingales using a Skorokhod representation approach, Advances in Applied Probability, 5, 119-137.

[29] Teitelbaum, J. C. (2013), Asymmetric empirical similarity, Mathematical Social Sciences, 66(3), 346-351.