

# The Spatial Random Effects and the Spatial Fixed Effects Model: The Hausman Test in a Cliff and Ord Panel Model

Jan Mutl\* and Michael Pfaffermayr†

December 4, 2008

## Abstract

This paper studies the spatial random effects and spatial fixed effects model. The model includes a Cliff and Ord type spatial lag of the dependent variable as well as a spatially lagged one-way error component structure, accounting for both heterogeneity and spatial correlation across units. We discuss instrumental variable estimation under both the fixed and the random effects specification and propose a spatial Hausman test which compares these two models accounting for spatial autocorrelation in the disturbances. We derive the large sample properties of our estimation procedures and show that the test statistic is asymptotically chi-square distributed. A small Monte Carlo study demonstrates that this test works well even in small panels.

**Keywords:** Spatial Econometrics; Panel Data; Random Effects Estimator; Within Estimator; Hausman test

**JEL Classification:** C21; C23

---

\*Jan Mutl: Institute for Advanced Studies, Stumpergasse 56, 1060 Vienna, Austria. Email: mutl@ihs.ac.at

†Michael Pfaffermayr: Department of Economics, University of Innsbruck, Universitätsstrasse 15, 6020 Innsbruck and Austrian Institute of Economic Research, P.O.-Box 91, A-1103 Vienna, Austria; CESIFO, Germany. Financial support from the Austrian science fund is gratefully acknowledged. Email: Michael.Pfaffermayr@uibk.ac.at.

# 1 Introduction\*

The panel literature offers the random effects and the fixed effects model to account for heterogeneity across units. While the random effects estimator is more efficient than the fixed effects estimator, in many non-spatial empirical applications the random effects model is rejected in favour of the fixed effects model. Often there are plausible arguments that the explanatory variables are correlated with unit specific effects. For example, in earnings equations unobserved ability of individuals may be reflected in both the unit specific effects and the explanatory variables such as the years of schooling. The estimation of gravity equations to model bilateral trade flows is another important example where the assumptions of the random effects model are often found to be violated.<sup>1</sup> It is perfectly sensible that this issue also comes up in spatial panel models.

This paper contributes to the literature by introducing a spatial generalized methods of moments estimator for panel data models with Cliff and Ord type spatial autocorrelation and one-way error components. Our work complements the seminal paper of Kapoor, Kelejian and Prucha (2007) who provide a spatial generalized least squares (spatial GLS) estimator for the spatial random effects model. In addition to their work, our model allows for an endogenous spatial lag of the dependent variable. We discuss the proper instrumentation of the endogenous spatial lag and suggest an instrumental variable (IV) procedure for both the spatial within estimator and the spatial GLS estimator. In order to discriminate between the two spatial panel models, we also propose a Hausman test that accounts for spatially autocorrelated disturbances. Specifically, we derive the joint asymptotic distribution of the spatial GLS and the spatial within estimators, as well as the asymptotic distribution of the spatial Hausman test for random versus fixed effects. This test should enable applied researchers to choose between these two models, when spatial correlation of the endogenous variable and/or the disturbances is present.

Our paper is not the first that considers spatial within or fixed effects estimators. Case (1991) seems to be among the first in estimating spatial random and fixed effects models. Korniotis (2008) introduces a bias-corrected estimator for a spatial dynamic panel model with fixed effects. Lee and

---

\*We would like to thank Robert Kunst and Ingmar Prucha for helpful comments and suggestions.

<sup>1</sup>See also the papers cited in Baltagi (2008) for more examples.

Yu (2008) establish the asymptotic properties of quasi-maximum likelihood estimators for fixed effects spatial autoregressive (SAR) panel data models with SAR disturbances, where the time periods and/or the number of spatial units can be finite or large in all combinations except that both are finite (see also Yu, de Jong and Lee, 2006 and 2007).

In the next section we specify our model and spell out the maintained assumptions. Section 3 defines the two estimators under consideration and shows that the random effects and the within estimators are jointly asymptotically normally distributed under the random effects assumption. Section 4 introduces the feasible counterparts of the considered estimators based on an initial instrumental variable estimator. We show that the initial estimator is consistent and asymptotically normal and derive its asymptotic distribution. We also demonstrate that the true and feasible estimators have the same asymptotic distribution. Section 5 defines the spatial Hausman test that allows to discriminate between the two spatial panel models. It provides its asymptotic distribution under the null and also shows that the test statistics diverges in probability under the alternative hypothesis. In Section 6 we report the results of Monte Carlo experiments that assess both the size and the power of the proposed spatial Hausman test in finite samples. Finally, the last section concludes.

## 2 The Spatial Panel Model

Consider the following spatial panel model:

$$y_{it,N} = \lambda \sum_{j=1}^N w_{ij,N} y_{jt,N} + \mathbf{x}_{it,N} \boldsymbol{\beta} + \alpha + u_{it,N}. \quad (2.1)$$

Index  $i = 1, \dots, N$  denotes the cross-sectional dimension of the panel while the index  $t = 1, \dots, T$  refers to the time series dimension of the panel.  $y_{it,N}$  is the (scalar) dependent variable and  $\sum_{j=1}^N w_{ij,N} y_{jt,N}$  denotes the spatial lag of the dependent variable with  $w_{ij,N}$  being observable non-stochastic spatial weights.  $\mathbf{x}_{it,N}$  is a  $1 \times K - 1$  vector of exogenous variables,  $\boldsymbol{\beta}$  is a  $K - 1 \times 1$  parameter vector,  $\lambda$  a scalar parameters and  $\alpha$  refers to the constant.  $u_{it,N}$  is the overall disturbance term.

We allow for cross-sectional sectional correlation of the disturbances and, in particular, we assume that the disturbances follow a Cliff and Ord type

spatial autocorrelation (SAR(1) in terminology of Anselin, 1988) as proposed by Kapoor, Kelejian and Prucha (2007):

$$u_{it,N} = \rho \sum_{j=1}^N m_{ij,N} u_{jt,N} + \varepsilon_{it,N}, \quad (2.2)$$

where  $\rho$  is a scalar parameter and  $m_{ij,N}$  are observable spatial weights (possibly the same as the weights  $w_{ij,N}$ ). The innovations  $\varepsilon_{it,N}$  have the following one-way error component structure:

$$\varepsilon_{it,N} = \mu_{i,N} + \nu_{it,N}. \quad (2.3)$$

$\nu_{it,N}$  are independent innovations and  $\mu_{i,N}$  are individual effects, which can be either fixed or random.

We index all variables by the sample size  $N$ , since they form triangular arrays. This is necessary because the model involves inverses of matrices whose size depends on  $N$ , and hence their elements must change with  $N$ . Thus at the minimum  $y_{it,N}$  and  $u_{it,N}$  are triangular arrays in the present specification.

We sort the data so that the fast index is  $i$  and the slow index is  $t$ . Stacking the model over the  $N$  cross-sections for a single period  $t$  yields

$$\begin{aligned} \mathbf{y}_{t,N} &= \lambda \mathbf{W}_N \mathbf{y}_{t,N} + \mathbf{X}_{t,N} \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_N + \mathbf{u}_{t,N}, \\ \mathbf{u}_{t,N} &= \rho \mathbf{M}_N \mathbf{u}_{t,N} + \boldsymbol{\varepsilon}_{t,N}, \\ \boldsymbol{\varepsilon}_{t,N} &= \boldsymbol{\mu}_N + \boldsymbol{\nu}_{t,N} \end{aligned} \quad (2.4)$$

where  $\boldsymbol{\iota}_N$  is a  $(N \times 1)$  vector of ones and

$$\begin{aligned} \mathbf{y}_{t,N} &= \begin{pmatrix} y_{t1,N} \\ \vdots \\ y_{tN,N} \end{pmatrix}, \quad \mathbf{X}_{t,N} = \begin{pmatrix} \mathbf{x}_{t1,N} \\ \vdots \\ \mathbf{x}_{tN,N} \end{pmatrix}, \quad \mathbf{u}_{t,N} = \begin{pmatrix} u_{t1,N} \\ \vdots \\ u_{tN,N} \end{pmatrix} \\ \boldsymbol{\varepsilon}_{t,N} &= \begin{pmatrix} \varepsilon_{t1,N} \\ \vdots \\ \varepsilon_{tN,N} \end{pmatrix}, \quad \boldsymbol{\nu}_{t,N} = \begin{pmatrix} \nu_{t1,N} \\ \vdots \\ \nu_{tN,N} \end{pmatrix}, \quad \boldsymbol{\mu}_N = \begin{pmatrix} \mu_{1,N} \\ \vdots \\ \mu_{N,N} \end{pmatrix}, \\ \mathbf{W}_N &= \begin{pmatrix} w_{11,N} & \cdots & w_{1N,N} \\ \vdots & \ddots & \vdots \\ w_{N1,N} & \cdots & w_{NN,N} \end{pmatrix}, \quad \mathbf{M}_N = \begin{pmatrix} m_{11,N} & \cdots & m_{1N,N} \\ \vdots & \ddots & \vdots \\ m_{N1,N} & \cdots & m_{NN,N} \end{pmatrix}. \end{aligned} \quad (2.5)$$

Stacking over time periods, we write our model compactly as

$$\begin{aligned}
\mathbf{y}_N &= \lambda \mathbf{W}_N \mathbf{y}_N + \mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT} + \mathbf{u}_N & (2.6) \\
&= \mathbf{Z}_N \boldsymbol{\delta} + \mathbf{u}_N, \\
\mathbf{u}_N &= \rho \mathbf{M}_N \mathbf{u}_N + \boldsymbol{\varepsilon}_N, \\
\boldsymbol{\varepsilon}_N &= (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \boldsymbol{\mu}_N + \boldsymbol{\nu}_N.
\end{aligned}$$

where  $\mathbf{W}_N = (\mathbf{I}_T \otimes \mathbf{W}_N)$ ,  $\mathbf{M}_N = (\mathbf{I}_T \otimes \mathbf{M}_N)$ ,  $\mathbf{Z}_N = (\boldsymbol{\iota}_{NT}, \mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)$ ,  $\boldsymbol{\delta} = (\alpha, \lambda, \boldsymbol{\beta}')$ ,  $\boldsymbol{\iota}_T$  is a  $T \times 1$  vector of ones,  $\boldsymbol{\iota}_{NT}$  is an  $NT \times 1$  vector of ones and

$$\begin{aligned}
\mathbf{y}_N &= \begin{pmatrix} \mathbf{y}_{1,N} \\ \vdots \\ \mathbf{y}_{T,N} \end{pmatrix}, \quad \mathbf{X}_N = \begin{pmatrix} \mathbf{x}_{1,N} \\ \vdots \\ \mathbf{x}_{T,N} \end{pmatrix}, \quad \mathbf{u}_N = \begin{pmatrix} \mathbf{u}_{1,N} \\ \vdots \\ \mathbf{u}_{T,N} \end{pmatrix}, & (2.7) \\
\boldsymbol{\varepsilon}_N &= \begin{pmatrix} \boldsymbol{\varepsilon}_{1,N} \\ \vdots \\ \boldsymbol{\varepsilon}_{T,N} \end{pmatrix}, \quad \boldsymbol{\nu}_N = \begin{pmatrix} \boldsymbol{\nu}_{1,N} \\ \vdots \\ \boldsymbol{\nu}_{T,N} \end{pmatrix}.
\end{aligned}$$

Throughout we maintain the following basic assumptions, which follow closely those postulated in Kapoor, Kelejian and Prucha (2007).

**Assumption 1**

The elements of  $\boldsymbol{\nu}_N$  are independently and identically distributed with finite absolute  $4 + \delta_\nu$  moments for some  $\delta_\nu > 0$ . Furthermore,  $E(\nu_{i,N}^2) = \sigma_\nu^2 > 0$ .

**Assumption 2**

The spatial weights collected in  $\mathbf{M}_N$  and  $\mathbf{W}_N$  are non-stochastic and

- (a)  $m_{ii,N} = 0$  and  $w_{ii,N} = 0$ .
- (b) The absolute row and column sums of the matrices  $\mathbf{M}_N$ ,  $\mathbf{W}_N$ ,  $(\mathbf{I}_N - \rho \mathbf{M}_N)^{-1}$ ,  $(\mathbf{I}_N - \lambda \mathbf{W}_N)^{-1}$  are uniformly bounded in absolute value, i.e.  $\sum_{i=1}^N |a_{ij,N}| \leq k < \infty$ , where  $k$  does not depend on  $N$  (but may depend on parameters of the model, i.e. on  $\rho$  or  $\lambda$ , respectively) and  $a_{ij,N}$  denotes elements of the above matrices.
- (c)  $|\rho| \leq k_\rho < 1/\lambda_{\max}(\mathbf{M}_N)$ ,  $|\lambda| \leq k_\lambda < 1/\lambda_{\max}(\mathbf{W}_N)$ , where  $\lambda_{\max}(\cdot)$  denotes the largest absolute eigenvalue of a matrix.
- (d) The matrices  $(\mathbf{I}_N - \rho \mathbf{M}_N)$  and  $(\mathbf{I}_N - \lambda \mathbf{W}_N)$  are non-singular.

### Assumption 3

*The exogenous variables collected in  $\mathbf{X}_N$  are non-stochastic and vary with both individuals and time. Their elements are uniformly bounded in absolute value.*

Assumption 1 is a restriction on the higher moments of the disturbances required for asymptotic results. Assumption 2(a) is a typical normalization (but is not necessary for our asymptotic results). Assumptions 2(b) and (c) are satisfied in many empirical applications and hold, for example, if the spatial weights matrices  $\mathbf{W}_N$  and  $\mathbf{M}_N$  are (maximum)-row normalized and  $|\lambda| \leq k_\lambda < 1$  and  $|\rho| \leq k_\rho < 1$ , respectively. Observe that Assumption 2(d) follows from 2(c).<sup>2</sup> Assumptions like 2(b) - 2(d) are typically maintained in spatial models (see Kelejian and Prucha, 1999) and restrict the extent of spatial dependence among cross-section units. It will be satisfied if the spatial weighting matrix is sparse so that each unit possess a limited number of neighbors, or if the spatial weights decline sufficiently fast in distance.

## 3 The Estimation of Spatial Panel models

In their seminal paper Kapoor, Kelejian and Prucha (2007) concentrate on the random effects model, assuming that the explanatory variables and the unit specific error terms are independent. Yet, in applied work exactly this assumption often does not hold and a fixed effects specification is employed instead. Examples in a non-spatial setting, where the unit specific effects and the explanatory variables may be correlated include earnings equations. In this setting the unobserved individual ability of an individual is typically correlated with the years of schooling, which enters the earnings equation as explanatory variables (see e.g. Baltagi, 2008, p. 79). Also in models explaining bilateral trade flows, the random effects model is typically rejected in favour of the fixed effects model to mention another example (see Egger, 2000). In addition, the considered spatial panel model allows for an endogenous spatial lag. First, we analyze the spatial random effects estimator for this general spatial panel model.

---

<sup>2</sup>This follows from Corrolary 5.6.16 in Horn and Johnsonn (1985) using their Lemma 5.6.10.

### 3.1 The Spatial Random Effects Estimator

Under the random effects specification, the unit specific effects  $\mu_{i,N}$  are assumed to be random and the following standard assumption is maintained.

#### Assumption 4 (RE)

The elements of  $\boldsymbol{\mu}_N$  are independently and identically distributed with finite absolute  $4 + \delta_\mu$  moments for some  $\delta_\mu > 0$  and  $E(\mu_{i,N}^2) = \sigma_\mu^2 > 0$ . Furthermore, the elements of  $\boldsymbol{\mu}_N$  are independent of the process for  $\nu_{it,N}$  and  $E(\mu_{i,N} | \mathbf{X}_N) = \mathbf{0}$  for all  $i$  and  $t$ .

Under Assumptions 1, 2 and 4 (RE), it follows that the disturbances are generated as

$$\mathbf{u}_N = (\mathbf{I}_{NT} - \rho \mathbf{M}_N)^{-1} \boldsymbol{\varepsilon}_N. \quad (3.1)$$

Hence, under the random effects specification, the variance covariance matrix of the disturbances is given by

$$\begin{aligned} \boldsymbol{\Omega}_{u,N} &= E(\mathbf{u}_N \mathbf{u}_N') \\ &= (\mathbf{I}_{NT} - \rho \mathbf{M}_N)^{-1} [\sigma_\mu^2 (\boldsymbol{\nu}_T \boldsymbol{\nu}_T' \otimes \mathbf{I}_N) + \sigma_\nu^2 \mathbf{I}_{NT}] (\mathbf{I}_{NT} - \rho \mathbf{M}_N')^{-1}. \end{aligned} \quad (3.2)$$

It proves to be useful to use the notation  $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$  and define the following standard within and between transformation matrices  $\mathbf{Q}_{i,N}$  ( $i = 0, 1$ ):

$$\begin{aligned} \mathbf{Q}_{0,N} &= (\mathbf{I}_T - \frac{1}{T} \mathbf{J}_T) \otimes \mathbf{I}_N \\ \mathbf{Q}_{1,N} &= \frac{1}{T} \mathbf{J}_T \otimes \mathbf{I}_N, \end{aligned} \quad (3.3)$$

where  $\mathbf{J}_T$  is a  $T \times T$  matrix of unit elements. The matrices  $\mathbf{Q}_{i,N}$  are the standard transformation matrices utilized in the error component literature but adjusted for the different stacking of the data (compare Kapoor, Kelejian and Prucha, 2007 and Baltagi, 2008). The matrices  $\mathbf{Q}_{i,N}$  are symmetric and idempotent and mutually orthogonal. The variance covariance matrix of the disturbances can then be written as (see Baltagi, 2008, p. 18)

$$\boldsymbol{\Omega}_{u,N} = (\mathbf{I}_{NT} - \rho \mathbf{M}_N)^{-1} \boldsymbol{\Omega}_{\varepsilon,N} (\mathbf{I}_{NT} - \rho \mathbf{M}_N')^{-1}, \quad (3.4)$$

where

$$\boldsymbol{\Omega}_{\varepsilon,N} = E(\boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}_N') = \sigma_\nu^2 \mathbf{Q}_{0,N} + \sigma_1^2 \mathbf{Q}_{1,N}. \quad (3.5)$$

Furthermore, the inverse of  $\mathbf{\Omega}_{u,N}$  can then be expressed as

$$\mathbf{\Omega}_{u,N}^{-1} = (\mathbf{I}_{NT} - \rho\mathbf{M}_N) \mathbf{\Omega}_{\varepsilon,N}^{-1} (\mathbf{I}_{NT} - \rho\mathbf{M}'_N), \quad (3.6)$$

where

$$\mathbf{\Omega}_{\varepsilon,N}^{-1} = \sigma_\nu^{-2} \mathbf{Q}_{0,N} + \sigma_1^{-2} \mathbf{Q}_{1,N}. \quad (3.7)$$

If the parameter values  $\rho$ ,  $\sigma_\nu^2$  and  $\sigma_\mu^2$  (and, therefore,  $\sigma_1^2$ ) are known, the efficient GLS estimation procedure is to transform the model by the square root of the inverse of the variance-covariance matrix of the disturbances

$$\sigma_\nu \mathbf{\Omega}_{u,N}^{-1/2} = \sigma_\nu \mathbf{\Omega}_{\varepsilon,N}^{-1/2} (\mathbf{I}_{NT} - \rho\mathbf{M}_N). \quad (3.8)$$

This is equivalent to first applying the spatial counterpart of the Cochrane-Orcutt transformation  $(\mathbf{I}_{NT} - \rho\mathbf{M}_N)$  that eliminates the spatial correlation from the disturbances and then the familiar panel GLS transformation  $\sigma_\nu \mathbf{\Omega}_{\varepsilon,N}^{-1/2}$  that accounts of the variance-covariance structure of the innovations induced by the random effects. To simplify the exposition, we collect the parameters of the variance covariance matrix in a vector  $\boldsymbol{\vartheta} = (\rho, \sigma_\nu^2, \sigma_1^2)$  and use the notation  $\mathbf{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta})$  to explicitly note the dependence of the GLS transformation on these parameters. Observe that in a balanced panel the order with which the transformations are applied is irrelevant (see also Remark A1 in Kapoor, Kelejian and Prucha, 2007).

Since the spatial lag  $\mathbf{W}_N \mathbf{y}_N$  is endogenous in the (transformed) model, we adapt the instrumental variable procedure described in Kelejian and Prucha (1998).<sup>3</sup> Specifically, we first eliminate the spatial correlation eliminated from the error term using the Cochrane-Orcutt transformation. Then we apply the instrumental variable procedure for random effects models suggested by Baltagi and Li (1992), Cornwell, Schmidt and Wyhowsky (1992) and surveyed by Baltagi (2008). These authors show in a non-spatial setting that the optimal set of instruments for a random effects model with endogenous variables is comprised of  $[\mathbf{Q}_{0,N} \mathbf{X}_N, \mathbf{Q}_{1,N} \mathbf{X}_N, \boldsymbol{\iota}_{NT}]$ . Observe that

$$\begin{aligned} E[\mathbf{y}_N] &= (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} (\mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT}) \\ &= \left[ \sum_{k=0} \lambda^k \mathbf{W}_N^k \right] (\mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT}), \end{aligned} \quad (3.9)$$

---

<sup>3</sup>It is possible to use other sets of instruments, such as those proposed in Lee (2003) or Kelejian, Prucha and Yuzefovich (2004).

where  $\mathbf{W}_N^0 = \mathbf{I}_{NT}$ . Hence under the present assumptions, the ideal set of instruments is based on

$$\begin{aligned} \sigma_\nu \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta}) \mathbf{W}_N E[\mathbf{y}_N] &= \sigma_\nu \sum_{k=0} \lambda^k \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta}) \mathbf{W}_N^{k+1} (\mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT}) \quad (3.10) \\ &= \sum_{k=0} \lambda^k (\mathbf{Q}_{0,N} + \frac{\sigma_\nu}{\sigma_1} \mathbf{Q}_{1,N}) \mathbf{W}_N^{k+1} (\mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT}) \\ &\quad - \rho \sum_{k=0} \lambda^k (\mathbf{Q}_{0,N} + \frac{\sigma_\nu}{\sigma_1} \mathbf{Q}_{1,N}) \mathbf{M}_N \mathbf{W}_N^{k+1} (\mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT}) \end{aligned}$$

Therefore, the transformed endogenous variable  $\sigma_\nu \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta}) \mathbf{W}_N \mathbf{y}_N$  is best instrumented by

$$\begin{aligned} \mathbf{H}_{R,N} &= [\mathbf{H}_{Q,N}, \mathbf{H}_{P,N}] \quad (3.11) \\ &= [\mathbf{Q}_{0,N} \mathbf{G}_{0,N}, \mathbf{Q}_{1,N} \mathbf{G}_{1,N}], \end{aligned}$$

where  $\mathbf{G}_{0,N}$  contains a subset of the the linearly independent columns of

$$[\mathbf{X}_N, \mathbf{W}_N \mathbf{X}_N, \mathbf{W}_N^2 \mathbf{X}_N, \dots, \mathbf{M}_N \mathbf{X}_N, \mathbf{M}_N \mathbf{W}_N \mathbf{X}_N, \mathbf{M}_N \mathbf{W}_N^2 \mathbf{X}_N, \dots]$$

and  $\mathbf{G}_{1,N}$  contains a subset of the the linearly independent columns of

$$[\mathbf{G}_{0,N}, \boldsymbol{\iota}_{NT}, \mathbf{M}_N \boldsymbol{\iota}_{NT}, \mathbf{M}_N \mathbf{W}_N \boldsymbol{\iota}_{NT}, \mathbf{M}_N \mathbf{W}_N^2 \boldsymbol{\iota}_{NT}, \dots].$$

The columns in  $\mathbf{G}_{0,N}$  and  $\mathbf{G}_{1,N}$  must be chosen so that the columns of  $\mathbf{H}_{R,N}$  are linearly independent. In the special case where  $\mathbf{W}_N = \mathbf{M}_N$  the set of instruments is based on  $\mathbf{G}_{0,N} = [\mathbf{X}_N, \mathbf{W}_N \mathbf{X}_N, \mathbf{W}_N^2 \mathbf{X}_N, \dots]$  and  $\mathbf{G}_{1,N} = [\mathbf{G}_{0,N}, \boldsymbol{\iota}_{NT}, \mathbf{W}_N \boldsymbol{\iota}_{NT}, \dots]$ . If the spatial weighting matrices are row normalized, the set of instruments in  $\mathbf{G}_{1,N}$  includes only  $\boldsymbol{\iota}_{NT}$  besides  $\mathbf{G}_{0,N}$ , since in this case  $\mathbf{M}_N \boldsymbol{\iota}_{NT} = \mathbf{W}_N \boldsymbol{\iota}_{NT} = \boldsymbol{\iota}_{NT}$ .

In the following we assume that the  $NT \times p$  matrix of instruments denoted by  $\mathbf{H}_{R,N}$  is of the form described above. In order to derive asymptotic properties of the considered estimators, we maintain the following additional assumption for the matrix of instruments and the explanatory variables of the model collected in  $\mathbf{Z}_N = [\boldsymbol{\iota}_{NT}, \mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N]$ :

### Assumption 5

Let  $\tilde{\mathbf{Z}}_N(\boldsymbol{\vartheta}) = \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta}) \mathbf{Z}_N$ . The matrix of instruments  $\mathbf{H}_{R,N}$  has full column rank and consists of a subset of linearly independent columns of  $[\mathbf{Q}_{0,N} \mathbf{G}_{0,N}, \mathbf{Q}_{1,N} \mathbf{G}_{1,N}]$ . Furthermore, it satisfies the following conditions:

- (a)  $\mathbf{M}_{H_R H_R} = \lim_{N \rightarrow \infty} (NT)^{-1} \mathbf{H}'_{R,N} \mathbf{H}_{R,N}$  exists and is finite and non-singular,
- (b)  $\mathbf{M}_{H_R \tilde{\mathbf{Z}}} = p \lim_{N \rightarrow \infty} (NT)^{-1} \mathbf{H}'_{R,N} \tilde{\mathbf{Z}}_N$  exists and is finite with full column rank.

The spatial random effects estimator of  $\boldsymbol{\delta} = (\alpha, \lambda, \boldsymbol{\beta}')'$  is then defined as

$$\hat{\boldsymbol{\delta}}_{GLS,N} = \left[ \hat{\tilde{\mathbf{Z}}}_N(\boldsymbol{\vartheta})' \tilde{\mathbf{Z}}_N(\boldsymbol{\vartheta}) \right]^{-1} \hat{\tilde{\mathbf{Z}}}'_N(\boldsymbol{\vartheta}) \tilde{\mathbf{y}}_N(\boldsymbol{\vartheta}), \quad (3.12)$$

with  $\hat{\tilde{\mathbf{Z}}}_N(\boldsymbol{\vartheta}) = \mathbf{P}_{H_R,N} \tilde{\mathbf{Z}}_N(\boldsymbol{\vartheta})$ ,  $\tilde{\mathbf{Z}}_N(\boldsymbol{\vartheta}) = \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta}) \mathbf{Z}_N$  and  $\tilde{\mathbf{y}}_N(\boldsymbol{\vartheta}) = \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta}) \mathbf{y}_N$ .  $\mathbf{P}_{H_R,N} = \mathbf{H}_{R,N} (\mathbf{H}'_{R,N} \mathbf{H}_{R,N})^{-1} \mathbf{H}'_{R,N}$  is the projection matrix based on the instruments  $\mathbf{H}_{R,N}$ .

The joint asymptotic distribution of the spatial random effects estimator and the spatial within estimator under known nuisance parameter vector  $\boldsymbol{\vartheta}$  is given in Theorem 1 below. This theorem is based on the random effects Assumption 4 (RE) and forms the basis of the Hausman test. The asymptotic properties of this test and its feasible counterparts are given in Theorem 2 in Section 5.

### 3.2 The Spatial Within Estimator

As an alternative to the random effects assumption above, the spatial within estimator allows for possible correlation of the unit specific effects and the explanatory variables. We follow Mundlak (1978) and maintain the following fixed effects assumption:

#### Assumption 6 (FE)

The vector individual effects is given by

$$\boldsymbol{\mu}_N = (\mathbf{I}_N - \rho \mathbf{M}_N) \bar{\mathbf{X}}_N \boldsymbol{\pi} + \boldsymbol{\xi}_N,$$

where  $\boldsymbol{\pi} \neq \mathbf{0}_{(K-1) \times 1}$ , the  $N \times (K-1)$  matrix  $\bar{\mathbf{X}}_N$  contains the time averages of the explanatory variables and is given by

$$\bar{\mathbf{X}}_N = \left( \frac{1}{T} \right) (\boldsymbol{\iota}'_T \otimes \mathbf{I}_N) \mathbf{X}_N.$$

Finally, the elements of the  $N \times 1$  random vector  $\boldsymbol{\xi}_N$  satisfy Assumption 4 (RE).

We can redefine the innovations of the model under the fixed effect assumption to be

$$\boldsymbol{\varepsilon}_N = (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \boldsymbol{\xi}_N + \boldsymbol{\nu}_N, \quad (3.13)$$

i.e. replace the individual effects  $\boldsymbol{\mu}_N$  with only their independent component  $\boldsymbol{\xi}_N$ . Observe that the fixed effects assumption then implies that the disturbances of the model are given by

$$\mathbf{u}_N = (\mathbf{I}_{NT} - \rho \mathbf{M}_N)^{-1} \boldsymbol{\varepsilon}_N + (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \bar{\mathbf{X}}_N \boldsymbol{\pi}. \quad (3.14)$$

Clearly, the explanatory variables are uncorrelated with the individual effects if and only if  $\boldsymbol{\pi} = \mathbf{0}$  and the random effects model defined under Assumption 4 (RE) arises as a special case of the fixed effects model defined under Assumption 6 (FE). Observe that under Assumption 6 (FE) with  $\boldsymbol{\pi} \neq \mathbf{0}$  the spatial GLS estimator is inconsistent. However, one can apply the within transformation  $\mathbf{Q}_{0,N}$  to wipe out the individual effects (see e.g. Baltagi, 2008 and Mundlak 1978). Observe that  $\mathbf{Q}_{0,N} (\mathbf{I}_{NT} - \rho \mathbf{M}_N) = (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{Q}_{0,N}$ , and hence one obtains

$$\begin{aligned} \mathbf{Q}_{0,N} \mathbf{u}_N &= (\mathbf{E}_T \otimes \mathbf{I}_N) [\rho (\mathbf{I}_T \otimes \mathbf{M}_N) \mathbf{u}_N + (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) [(\mathbf{I}_N - \rho \mathbf{M}_N) \bar{\mathbf{X}}_N \boldsymbol{\pi} + \boldsymbol{\xi}_N] + \boldsymbol{\nu}_N], \\ &= \rho (\mathbf{I}_T \otimes \mathbf{M}_N) (\mathbf{E}_T \otimes \mathbf{I}_N) \mathbf{u}_N + (\mathbf{E}_T \otimes \mathbf{I}_N) \boldsymbol{\nu}_N \\ &= \rho (\mathbf{I}_T \otimes \mathbf{M}_N) \mathbf{Q}_{0,N} \mathbf{u}_N + \mathbf{Q}_{0,N} \boldsymbol{\nu}_N. \end{aligned} \quad (3.15)$$

or

$$\mathbf{Q}_{0,N} \mathbf{u}_N = (\mathbf{I}_{NT} - \rho \mathbf{M}_N)^{-1} \mathbf{Q}_{0,N} \boldsymbol{\nu}_N$$

where  $\mathbf{E}_T = (\mathbf{I}_T - \frac{1}{T} \mathbf{J}_T)$ . Hence, one can apply the Cochrane-Orcutt type transformation on the within transformed model to obtain the fixed effects generalized least squares (FEGLS) estimator.

More importantly, one can base the method of moment estimator of  $(\rho, \sigma_\nu^2)$  on the initial within transformed residuals of the initial within estimator as given by  $\mathbf{Q}_{0,N} \mathbf{u}_N$ , which are consistently estimated under both the spatial random effects model and the spatial fixed effects model. Obviously, the set of instruments denoted by  $\mathbf{H}_{Q,N}$  now comprises the linear independent columns of  $\mathbf{Q}_{0,N} \mathbf{G}_{0,N}$ . Since the constant is wiped out in the spatial within estimator, we define the  $(NT \times K)$  matrix  $\mathbf{Z}_{Q,N} = \mathbf{Q}_{0,N} [\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N]$  with the corresponding  $(K \times 1)$  parameter vector  $\boldsymbol{\theta} = (\lambda, \boldsymbol{\beta}')'$ .

In order to derive the asymptotic properties of the spatial within estimator, we maintain the following additional assumptions for the matrix of instruments used in the spatial within model:

**Assumption 7**

Let  $\mathbf{Z}_{Q,N} = \mathbf{Q}_{0,N} [\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N]$ . The matrix of instruments  $\mathbf{H}_{Q,N}$  has full column rank and consists of a subset of linearly independent columns of  $\mathbf{Q}_{0,N} \mathbf{G}_{0,N}$ . Furthermore, it satisfies the following conditions:

- (a)  $\mathbf{M}_{H_Q H_Q} = \lim_{N \rightarrow \infty} (NT)^{-1} \mathbf{H}'_{Q,N} \mathbf{H}_{Q,N}$  is finite and non-singular with full column rank.
- (b)  $\mathbf{M}_{H_Q Z^*} = p \lim_{N \rightarrow \infty} (NT)^{-1} \mathbf{H}'_{Q,N} (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{Z}_{Q,N}$  exists and is finite with full column rank.

Again, treating  $\rho$  as known, we apply the Cochrane-Orcutt type transformation to the within transformed model yielding:

$$\begin{aligned} \mathbf{Q}_{0,N} (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{y}_N &= (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{Q}_{0,N} \mathbf{y}_N & (3.16) \\ &= (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{Q}_{0,N} \mathbf{Z}_{Q,N} + \mathbf{Q}_{0,N} \boldsymbol{\nu}_N \\ \mathbf{y}_N^* (\rho) &= \mathbf{Z}_N^* (\rho) \boldsymbol{\theta} + \boldsymbol{\nu}_N^*, \end{aligned}$$

where

$$\begin{aligned} \mathbf{y}_N^* (\rho) &= (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{Q}_{0,N} \mathbf{y}_N, & (3.17) \\ \mathbf{Z}_N^* (\rho) &= (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{Z}_{Q,N} \\ \boldsymbol{\nu}_N^* &= \mathbf{Q}_{0,N} \boldsymbol{\nu}_N. \end{aligned}$$

The spatial within estimator is then obtained by applying IV to the transformed model to obtain

$$\hat{\boldsymbol{\theta}}_{W,N} = \left[ \hat{\mathbf{Z}}_N^* (\rho)' \hat{\mathbf{Z}}_N^* (\rho) \right]^{-1} \hat{\mathbf{Z}}_N^* (\rho)' \mathbf{y}_N^* (\rho), \quad (3.18)$$

with

$$\begin{aligned} \hat{\mathbf{Z}}_N^* (\rho) &= \mathbf{P}_{H_{Q,N}} \mathbf{Z}_N^* (\rho) & (3.19) \\ &= \mathbf{P}_{H_{Q,N}} (\mathbf{I}_{NT} - \rho \mathbf{M}_N) \mathbf{Z}_N, \end{aligned}$$

where  $\mathbf{P}_{H_{Q,N}} = \mathbf{H}_{Q,N} (\mathbf{H}'_{Q,N} \mathbf{H}_{Q,N})^{-1} \mathbf{H}'_{Q,N}$  is the projection matrix based on the instruments  $\mathbf{H}_{Q,N}$ .

The following theorem establishes our main asymptotic result concerning the common asymptotic distribution of the spatial random effects and the spatial within estimators under random effects Assumption 4 (RE). The Hausman test for spatial panels derived below will be based on this result. Since the random effects estimator includes the constant, we define  $\hat{\boldsymbol{\delta}}_{GLS,N} = \left( \hat{\alpha}_{GLS,N}, \hat{\boldsymbol{\theta}}'_{GLS,N} \right)'$ .

**Theorem 1** *Let Assumptions 1-5 and 7 hold. Then*

$$\sqrt{NT} \begin{pmatrix} \widehat{\boldsymbol{\theta}}_{GLS,N} - \boldsymbol{\theta} \\ \widehat{\boldsymbol{\theta}}_{W,N} - \boldsymbol{\theta} \end{pmatrix} \xrightarrow{d} N \left( \mathbf{0}, \begin{bmatrix} \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_{GLS} \\ \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_W \end{bmatrix} \right),$$

where  $\boldsymbol{\Sigma}_W = \sigma_\nu^2 \left( \mathbf{M}'_{H_Q Z^*} \mathbf{M}_{H_Q H_Q}^{-1} \mathbf{M}_{H_Q Z^*} \right)^{-1}$  and  $\boldsymbol{\Sigma}_{GLS}$  is the lower-right  $K \times K$  block of the matrix  $\sigma_\nu^2 \left( \mathbf{M}'_{H_R \tilde{Z}} \mathbf{M}_{H_R H_R}^{-1} \mathbf{M}_{H_R \tilde{Z}} \right)^{-1}$ .

Proof: See the Appendix.

## 4 Feasible Estimation

The spatial GLS and spatial within estimators defined above are based on the unknown parameters  $\rho$ ,  $\sigma_\nu^2$  and  $\sigma_\mu^2$  which have to be estimated. The feasible estimation procedure starts by estimating the within transformed model using the instruments  $\mathbf{H}_{Q,N} = \mathbf{Q}_{0,N} \mathbf{G}_{0,N}$  as described above to obtain initial within IV estimates. This initial estimator is consistent under both the random effects and the fixed effects specification (see Baltagi, 2008) and it can be written as

$$\widehat{\boldsymbol{\theta}}_{I,N} = \left[ \widehat{\mathbf{Z}}'_{Q,N} \mathbf{Q}_{0,N} \mathbf{Z}_{Q,N} \right]^{-1} \widehat{\mathbf{Z}}'_{Q,N} \mathbf{Q}_{0,N} \mathbf{y}_N, \quad (4.1)$$

where  $\widehat{\mathbf{Z}}_{Q,N} = \mathbf{P}_{H_{Q,N}} \mathbf{Q}_{0,N} \mathbf{Z}_{Q,N}$ . The following proposition gives the asymptotic distribution of the initial estimator.

**Proposition 1** *Let the limit  $\mathbf{M}_{H_Q Z} = p \lim_{N \rightarrow \infty} (NT)^{-1} \mathbf{H}'_{Q,N} \mathbf{Z}_{Q,N}$  exist and be finite with full column rank. Let Assumptions 1-3, 7 and either 4 (RE), or 6 (FE) hold. Then*

$$\sqrt{NT} \left( \widehat{\boldsymbol{\theta}}_{I,N} - \boldsymbol{\theta} \right) \xrightarrow{d} N \left( \mathbf{0}, \boldsymbol{\Sigma}_I \right),$$

where  $\boldsymbol{\Sigma}_I = \sigma_\nu^2 \left( \mathbf{M}'_{H_Q Z} \mathbf{M}_{H_Q H_Q}^{-1} \mathbf{M}_{H_Q Z} \right)^{-1}$ .

Proof: See the Appendix.

The projected residuals then give consistent initial estimates of  $\widehat{\mathbf{Q}}_{0,N}\mathbf{u}_N$  which can be used in the spatial generalized moments (GM) estimator as suggested by Kapoor, Kelejian and Prucha (2007). These authors use OLS residuals, which are consistent under the random effects assumption 4 (RE) but would be biased under the fixed effects assumption 6 (FE). The spatial GM estimator for  $\rho, \sigma_\nu^2$  can then be based on the first three moment conditions given in Kapoor, Kelejian and Prucha (2007).

Using  $\mathbf{Q}_{i,N}\mathbf{M}_N = \mathbf{M}_N\mathbf{Q}_{i,N}$  and the notation  $\bar{\boldsymbol{\varepsilon}}_N = \mathbf{M}_N\boldsymbol{\varepsilon}_N$ , we can formulate the first three moment conditions in terms of  $\mathbf{Q}_{0,N}\mathbf{u}_N$  as

$$\begin{aligned}
& E \begin{bmatrix} \frac{1}{N(T-1)}\boldsymbol{\varepsilon}'_N\mathbf{Q}_{0,N}\boldsymbol{\varepsilon}_N \\ \frac{1}{N(T-1)}\bar{\boldsymbol{\varepsilon}}'_N\mathbf{Q}_{0,N}\bar{\boldsymbol{\varepsilon}}_N \\ \frac{1}{N(T-1)}\bar{\boldsymbol{\varepsilon}}'_N\mathbf{Q}_{0,N}\boldsymbol{\varepsilon}_N \end{bmatrix} \tag{4.2} \\
&= \begin{bmatrix} \frac{1}{N(T-1)}\mathbf{u}'_N\mathbf{Q}_{0,N}(\mathbf{I}_{NT} - \rho\mathbf{M}'_N)(\mathbf{I}_{NT} - \rho\mathbf{M}_N)\mathbf{Q}_{0,N}\mathbf{u}_N \\ \frac{1}{N(T-1)}\mathbf{u}'_N\mathbf{Q}_{0,N}(\mathbf{I}_{NT} - \rho\mathbf{M}'_N)\mathbf{M}'_N\mathbf{M}_N(\mathbf{I}_{NT} - \rho\mathbf{M}_N)\mathbf{Q}_{0,N}\mathbf{u}_N \\ \frac{1}{N(T-1)}\mathbf{u}'_N\mathbf{Q}_{0,N}(\mathbf{I}_{NT} - \rho\mathbf{M}'_N)\mathbf{M}'_N(\mathbf{I}_{NT} - \rho\mathbf{M}_N)\mathbf{Q}_{0,N}\mathbf{u}_N \end{bmatrix} \\
&= \begin{bmatrix} \sigma_\nu^2 \\ \sigma_\nu^2\frac{1}{N}\text{tr}(\mathbf{M}'_N\mathbf{M}_N) \\ 0 \end{bmatrix}
\end{aligned}$$

Under the random effects model, Assumption 4 (RE), we add a fourth moment condition.

$$\begin{aligned}
& E \left[ \frac{1}{N}\boldsymbol{\varepsilon}'_N\mathbf{Q}_{1,N}\boldsymbol{\varepsilon}_N \right] \tag{4.3} \\
&= \left[ \frac{1}{N}\mathbf{u}'_N\mathbf{Q}_{1,N}(\mathbf{I}_{NT} - \rho\mathbf{M}'_N)(\mathbf{I}_{NT} - \rho\mathbf{M}_N)\mathbf{Q}_{1,N}\mathbf{u}_N \right] \\
&= \sigma_1^2
\end{aligned}$$

With the solution of the first three moment conditions at hand, one can solve the fourth moment condition to obtain an estimate of  $\sigma_1^2$ . Theorem 1 in Kapoor, Kelejian and Prucha (2007, p. 108 ) shows that the estimators for  $\rho, \sigma_\nu^2$  and  $\sigma_1^2$  based on these moment conditions and some additional assumptions (see their Assumption 5) are consistent as long as the initial estimator  $\widehat{\boldsymbol{\theta}}_{I,N}$  is consistent. Note, that this is the case for the spatial within estimator under both the Assumption 4 (RE) and Assumption 6 (FE).

Proposition 2 demonstrates that the parameters  $\rho$ ,  $\sigma_\nu^2$  and  $\sigma_1^2$  are nuisance parameters and that the feasible spatial random effects and the feasible spatial within estimates have the same asymptotic distribution as their counterparts based on the true values of  $\rho$ ,  $\sigma_\nu^2$  and  $\sigma_1^2$ .

**Proposition 2** *Let the feasible estimators  $\hat{\boldsymbol{\theta}}_{FGLS,N}$  and  $\hat{\boldsymbol{\theta}}_{FW,N}$  be based on consistent estimators of  $\rho$ ,  $\sigma_\nu^2$ , and  $\sigma_1^2$ . Then under Assumptions 1-5 and 7 we have*

$$\begin{aligned}\sqrt{NT} \left( \hat{\boldsymbol{\theta}}_{FGLS,N} - \hat{\boldsymbol{\theta}}_{GLS,N} \right) &\xrightarrow{p} \mathbf{0}, \\ \sqrt{NT} \left( \hat{\boldsymbol{\theta}}_{FW,N} - \hat{\boldsymbol{\theta}}_{W,N} \right) &\xrightarrow{p} \mathbf{0}.\end{aligned}$$

Proof: See the Appendix.

## 5 Hausman Specification Test

The spatial within estimator is consistent under both the random effects assumption 4 (RE) and the fixed effects assumption 6 (FE) since it wipes out the unit specific effects by applying the within transformation. The critical assumption for the validity of the spatial random effects model is that  $E(u_{i,N}|\mathbf{X}_N) = \mathbf{0}$ , implying that the spatial random effects model is inconsistent under the fixed effects Assumption 6 (FE). The Hausman test (Hausman, 1978) suggests comparing these two estimators and to test whether the random effects assumption maintaining  $E(u_{i,N}|\mathbf{X}_N) = \mathbf{0}$  holds true. The spatial GLS estimator of the random effects model is more efficient than the spatial within estimator under the random effects assumption 4 (RE). Moreover, under  $H_0 : \boldsymbol{\pi} = \mathbf{0}$  both considered estimators are consistent, while under  $H_1 : \boldsymbol{\pi} \neq \mathbf{0}$  the spatial random effects estimator is inconsistent, but the spatial within estimator is consistent.

The theorem below defines the Hausman test statistic for spatial panels and provides its asymptotic distribution under the null.

**Theorem 2** *Under Assumptions 1-5 and 7, we have*

$$(a) \quad \sqrt{NT} \left( \hat{\boldsymbol{\theta}}_{GLS,N} - \hat{\boldsymbol{\theta}}_{W,N} \right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}),$$

where  $\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}$  is positive definite.

(b)

$$\widehat{\Sigma}_{W,N} - \widehat{\Sigma}_{GLS,N} \xrightarrow{p} \Sigma_W - \Sigma_{GLS}$$

where

$$\widehat{\Sigma}_{W,N} = \widehat{\sigma}_{\nu,N}^2 NT \left[ \mathbf{Z}'_{Q,N} (\mathbf{I}_{NT} - \widehat{\rho}_N \mathbf{M}'_N) \mathbf{P}_{H_{Q,N}} (\mathbf{I}_{NT} - \widehat{\rho}_N \mathbf{M}_N) \mathbf{Z}_{Q,N} \right]^{-1},$$

and

$$\widehat{\Sigma}_{GLS,N} = \widehat{\sigma}_{\nu,N}^2 NT \left[ \mathbf{Z}'_N \boldsymbol{\Omega}_{u,N}^{-1/2} \left( \widehat{\boldsymbol{\vartheta}}_N \right) \mathbf{P}_{H_{R,N}} \boldsymbol{\Omega}_{u,N}^{-1/2} \left( \widehat{\boldsymbol{\vartheta}}_N \right) \mathbf{Z}_N \right]^{-1},$$

with  $\widehat{\boldsymbol{\vartheta}}_N = (\widehat{\rho}_N, \widehat{\sigma}_{\nu,N}^2, \widehat{\sigma}_{1,N}^2)'$  being some consistent estimator of  $\boldsymbol{\vartheta}$ .

(c) Let

$$\widehat{H}_N = NT \left( \widehat{\boldsymbol{\theta}}_{FGLS,N} - \widehat{\boldsymbol{\theta}}_{FW,N} \right)' \left( \widehat{\Sigma}_{W,N} - \widehat{\Sigma}_{GLS,N} \right)^{-1} \left( \widehat{\boldsymbol{\theta}}_{FGLS,N} - \widehat{\boldsymbol{\theta}}_{FW,N} \right),$$

and

$$H_N = NT \left( \widehat{\boldsymbol{\theta}}_{GLS,N} - \widehat{\boldsymbol{\theta}}_{W,N} \right)' (\Sigma_W - \Sigma_{GLS})^{-1} \left( \widehat{\boldsymbol{\theta}}_{GLS,N} - \widehat{\boldsymbol{\theta}}_{W,N} \right).$$

$\widehat{H}_N - H_N \xrightarrow{p} 0$ , where  $H_N$  is asymptotically  $\chi^2$  distributed with  $K$  degrees of freedom.

Proof: See the Appendix.

Given Theorem 2 in Kapoor, Kelejian and Prucha (2007) and the results in this paper, a feasible estimation and testing procedure can be summarized as follows. First calculate a consistent initial instrumental variables estimator  $\widehat{\boldsymbol{\theta}}_{I,N}$  which ignores the spatial correlation in the disturbances and wipes out the individual effects using the within transformation. Second, use the resulting estimated disturbances in a spatial GM procedure as described in Kapoor, Kelejian and Prucha (2007) and obtain a (consistent) estimator  $\widehat{\boldsymbol{\vartheta}}_N = (\widehat{\rho}_N, \widehat{\sigma}_{\nu,N}^2, \widehat{\sigma}_{1,N}^2)'$ . Third, transform the model by the spatial Cochrane-Orcutt transformation and then either the within, or the GLS transformation to obtain the spatial GLS and spatial within estimators  $\widehat{\boldsymbol{\theta}}_{FGLS,N}$  and  $\widehat{\boldsymbol{\theta}}_{FW,N}$ . Finally, calculate the Hausman test statistics  $\widehat{H}_N$  and use it to make a decision whether the random or fixed effects specification is more appropriate.

Under the alternative Assumption 6 (FE) one obtains

$$\begin{aligned}
\widehat{\boldsymbol{\delta}}_{GLS,N} &= \left[ \widehat{\mathbf{Z}}_N(\boldsymbol{\vartheta})' \widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta}) \right]^{-1} \widehat{\mathbf{Z}}_N'(\boldsymbol{\vartheta}) \widetilde{\mathbf{y}}_N(\boldsymbol{\vartheta}) \\
&= \boldsymbol{\delta} + \left[ \widehat{\mathbf{Z}}_N(\boldsymbol{\vartheta})' \widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta}) \right]^{-1} \widehat{\mathbf{Z}}_N'(\boldsymbol{\vartheta}) \sigma_\nu \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2} \cdot \\
&\quad \left[ (\boldsymbol{\nu}_T \otimes \mathbf{I}_N) \overline{\mathbf{X}}_N^*(\boldsymbol{\vartheta}) \boldsymbol{\pi} + (\boldsymbol{\nu}_T \otimes \mathbf{I}_N) \boldsymbol{\xi}_N + \boldsymbol{\nu}_N \right],
\end{aligned} \tag{5.1}$$

where  $\overline{\mathbf{X}}_N^*(\boldsymbol{\vartheta}) = (\mathbf{I}_N - \rho \mathbf{M}_N) \overline{\mathbf{X}}_N$ . On the other hand, the within estimator remains to be given as

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}_{W,N} &= \left[ \widehat{\mathbf{Z}}_N^*(\rho)' \mathbf{Z}_N^*(\rho) \right]^{-1} \widehat{\mathbf{Z}}_N^*(\rho)' \mathbf{y}^*(\rho) \\
&= \boldsymbol{\theta} + \left[ \widehat{\mathbf{Z}}_N^*(\rho)' \mathbf{Z}_N^*(\rho) \right]^{-1} \widehat{\mathbf{Z}}_N^*(\rho)' \mathbf{Q}_{0,N} \boldsymbol{\nu}_N,
\end{aligned} \tag{5.2}$$

and hence is the same as under  $H_0 : \boldsymbol{\pi} = \mathbf{0}$ . The following Proposition shows that the Hausman test statistic is a consistent statistic, i.e., the power of the test approaches unity as  $N \rightarrow \infty$  for an arbitrary significance level of the test.

**Proposition 3** *Let Assumptions 1-3 and 5-7 hold and let  $\gamma > 0$  be some positive constant. Then  $\lim_{N \rightarrow \infty} P(H_N > \gamma) = 1$ .*

Proof: See the Appendix

## 6 Monte Carlo Evidence

The Monte Carlo analysis investigates the small sample properties of the proposed spatial Hausman test. For this we use a simple spatial panel model that includes one explanatory variable and a constant:

$$y_{it} = \beta x_{it} + \alpha + u_{it}, \quad i = 1, \dots, N \text{ and } t = 1, \dots, T. \tag{6.1}$$

We set  $\beta = 0.5$  and  $\alpha = 5$ . The explanatory variable is generated as  $x_{it} = \zeta_i + z_{it}$  with  $\zeta_i \sim i.i.d. U[-7.5, 7.5]$  and  $z_{it} \sim i.i.d. U[-7.5, 7.5]$  with  $U[a, b]$  denoting the uniform distribution on the interval  $[a, b]$ .  $x_{it}$  is treated as non-stochastic variable and it is held fixed in repeated samples. The individual-specific effects are allowed to be correlated with  $\bar{x}_i$ , setting  $\mu_i = \mu_{i0} + \pi \bar{x}_i$ ,

where  $\mu_{i0}$  is drawn from a normal distribution, i.e.  $\mu_{i0} \sim i.i.d. N(0, 10\phi)$  and  $\pi$  is a constant parameter. This mimics the fixed effects assumption 2(FE) with  $\pi \neq 0$ . At  $\pi = 0$  the random effects assumption 4 holds and it forms the null for the spatial Hausman test. We normalize  $\mu_i$  so that its mean is 0 and its variance  $10\phi$ , where  $\phi = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\varepsilon^2}$ ,  $0 < \phi < 1$ , denotes the proportion of the total variance due to the presence of the individual-specific effects. For the remainder error we assume  $\varepsilon_{it} \sim i.i.d. N(0, 10(1 - \phi))$ . This implies that total the variance of the disturbances is  $\sigma_\mu^2 + \sigma_\varepsilon^2 = 10$ .

The row normalized spatial weighting matrix uses a regular lattice with 144 and 324 cells, respectively, containing one observation each. The spatial weighting scheme is based on a rook design, where every unit is surrounded by four neighbors. The corresponding spatial weighting matrix is maximum-row normalized following Kelejian and Prucha (2007). We will use the same spatial weighting matrix to generate both the endogenous spatial lag and the spatial lag of the error term.

The spatial parameters  $\lambda$  and  $\rho$  vary over the set  $\{-0.8, -0.4, 0, 0.4, 0.8\}$ . The parameter  $\pi$  takes its values in  $\{-0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3\}$ . Based on the discussion above we use the instruments  $\mathbf{H}_{Q,N} = [\mathbf{Q}_{0,N}\mathbf{x}_N, \mathbf{Q}_{0,N}\mathbf{W}_N\mathbf{x}_N, \mathbf{Q}_{0,N}\mathbf{W}_N^2\mathbf{x}_N]$ , while  $\mathbf{H}_{R,N}$  is composed of  $[\mathbf{H}_{Q,N}, \mathbf{Q}_{1,N}\mathbf{x}_N, \mathbf{Q}_{1,N}\mathbf{W}_N\mathbf{x}_N, \mathbf{Q}_{1,N}\mathbf{W}_N^2\mathbf{x}_N, \iota_{NT}, \mathbf{Q}_{1,N}\mathbf{W}_N\iota_{NT}]$ .

In each experiment we calculate the size of the Hausman test, which is given by the share of rejections at  $\pi = 0$ . The power of the spatial Hausman test is given by the share of rejections at  $\pi \neq 0$ .

===== Tables 1-4 =====

The baseline scenario is reported in Table 1 setting  $N = 144$ ,  $T = 5$  and  $\phi = 0.5$ . The results show that the proposed spatial IVGLS estimators work well and that the spatial Hausman test exhibits good performance for almost all considered parameter configurations. In the experiments reported in this Table, the spatial Hausman test comes close to the nominal size of 0.05 in most of the cases. Exceptions are only observed for high values of  $\lambda$ , where the test is slightly oversized. For example, at  $\lambda = 0.8$  and  $\rho = 0.8$  the size of the spatial Hausman test is 0.09. At negative values of  $\rho$ , this phenomenon is not observed. The power of the test by and large remains unaffected by variations of  $\rho$  and  $\lambda$ , although it seems somewhat lower at high absolute values of  $\rho$  or high absolute values of  $\lambda$ .

A larger cross-section ( $N = 324$ ) improves both the size and the power of the test as expected (see Table 2). The size distortion at high positive values of  $\lambda$  is now reduced and the power of the spatial Hausman test is considerably higher. In Table 3 we extend the time series dimension and set  $T = 10$ . The size distortion at high values of  $\lambda$  becomes smaller as  $T$  increases and this effect seems more pronounced than in an extended cross-section as analyzed in Table 2. However, the improvement in power is much smaller as compared to extending the cross-section dimension.

In Table 4 we set  $N = 144$  and  $\phi = 0.8$ , so that  $\sigma_\mu^2 = 8$  and  $\sigma_\nu^2$  is 2. With a larger weight of the variance of the unit specific effects, we observe a better performance of the spatial Hausman test in terms of its size. The size distortion observed in the baseline scenario now vanishes. Also, the power of the test is significantly higher.

===== Tables 5-7 =====

Tables 5-7 report the root mean square error (RMSE) and the bias of the estimators of  $\beta$ ,  $\lambda$  and  $\rho$  for the basic case with  $N = 144$ ,  $T = 5$  and  $\theta = 0.5$ . Following Kapoor, Kelejian and Prucha (2007) we define the bias as the difference between the respective median of the parameter estimate and its true counterpart, while  $RMSE = \sqrt{bias^2 + \left(\frac{IQ}{1.35}\right)^2}$ , where  $IQ$  is defined as the interquartile range, i.e. the difference between the 0.75 and 0.25 quantile of the simulated parameter distribution. Under a normal distribution the median and the mean coincide and  $\frac{IQ}{1.35}$  corresponds to the standard deviation (up to a rounding error).

The simulation exercises reveal a negligible bias for  $\beta$  and a somewhat higher efficiency of the random effects estimator under  $H_0$ . The gain in efficiency is especially large at high positive values of  $\lambda$  and at high absolute values of  $\rho$ . A similar pattern can be found for the RMSE of  $\lambda$ , although the efficiency loss of the spatial within estimator is much higher as compared to that for  $\beta$ . Under  $H_1$  the random effects estimator is inconsistent, leading to large biases in both  $\beta$  and  $\lambda$ . The bias of the slope parameter  $\beta$  is hardly affected by different degrees of spatial dependence as represented by the parameters values of  $\lambda$ . However, the bias is negative at low and negative values of  $\rho$  and turns to the positive if  $\rho$  gets high. With respect to the estimates of  $\lambda$ , we find that the bias is negative if  $\lambda$  or  $\rho$  take on negative values, but that it declines in  $\lambda$  and/or  $\rho$ . At  $\lambda = 0.8$  or  $\rho = 0.8$  the bias nearly vanishes. The results in Table 7 indicate that the estimates of  $\rho$

remain unaffected by deviations from  $H_0$  as expected. These estimates are based on the spatial within estimator which is consistent under both  $H_0$  and  $H_1$ .

We also assess the performance of the spatial Hausman test for non-normal disturbances.<sup>4</sup> In particular, we follow Kelejian and Prucha (1999) and assume lognormal remainder disturbances assuming  $\varepsilon_{it} = \frac{e^{\xi_{it}} - e^{0.5}}{\sqrt{e^2 - e^1}}$ , where  $\xi_{it} \sim i.i.d. N(0, 1)$ . Alternatively, we maintain that the distribution of the remainder error exhibits fatter tails than the normal and  $\varepsilon_{it} \sim i.i.d. t(5)$ . In both cases the performance of the spatial Hausman test is comparable to that under normal disturbances. However, under the  $t(5)$  error distribution the power of the test is smaller.

To summarize, the small Monte Carlo study shows that the proposed spatial Hausman test works well even in small panels. In this spatial setting, the test is able to detect deviations from the assumption that unobserved unit effects and the explanatory variables are uncorrelated, which is critical for the validity of spatial random effects models.

## 7 Conclusions

In this paper we study spatial random effects and spatial fixed effects models. We note that in many non-spatial applications the critical assumption maintained under the random effects specification, namely that unit specific effects and explanatory variables are uncorrelated, does not hold. This seems also a possibility in a spatial setting and should be tested, since the estimates of spatial random effects are inconsistent if this assumption fails to hold.

Using a spatial Cliff and Ord type model as analyzed in Kapoor, Kelejian and Prucha (2007) but augmented by an endogenous spatial lag, we introduce (feasible) instrumental variables estimators for both the spatial random effects model and a spatial fixed effects model. We derive the asymptotic distributions of these estimators as well as those of their feasible counterparts. In addition, we propose a spatial Hausman test to compare these two models, accounting for spatial autocorrelation in the disturbances. A small Monte Carlo study shows that this test works well even in small panels.

---

<sup>4</sup>The corresponding tables are available from the authors upon request.

# A Appendix

## Proof of Theorem 1:

We denote the  $(T + 1)N \times 1$  vector of i.i.d.  $(0, 1)$  innovations as  $\zeta_N = \left( \frac{\mu'_N}{\sigma_\mu}, \frac{\nu'_N}{\sigma_\nu} \right)'$  and we write the stacked estimators as<sup>5</sup>

$$\begin{pmatrix} \widehat{\delta}_{GLS,N} - \delta \\ \widehat{\theta}_{W,N} - \theta \end{pmatrix} = \widetilde{\mathbf{P}}_N \mathbf{F}_N \zeta_N, \quad (\text{A.1})$$

where

$$\widetilde{\mathbf{P}}_N = \begin{pmatrix} \widetilde{\mathbf{P}}_{R,N} & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{P}}_{Q,N} \end{pmatrix}, \quad \mathbf{F}_N = \begin{pmatrix} \mathbf{F}_{R1,N} & \mathbf{F}_{R2,N} \\ \mathbf{0} & \mathbf{F}_{Q,N} \end{pmatrix}$$

with

$$\begin{aligned} \widetilde{\mathbf{P}}_{R,N} &= \left[ \widetilde{\mathbf{Z}}'_N \mathbf{H}_{R,N} (\mathbf{H}'_{R,N} \mathbf{H}_{R,N})^{-1} \mathbf{H}'_{R,N} \widetilde{\mathbf{Z}}_N \right]^{-1}. \\ &\quad \widetilde{\mathbf{Z}}'_N \mathbf{H}_{R,N} (\mathbf{H}'_{R,N} \mathbf{H}_{R,N})^{-1}, \\ \widetilde{\mathbf{P}}_{Q,N} &= \left[ \mathbf{Z}^*{}'_{Q,N} \mathbf{H}_{Q,N} (\mathbf{H}'_{Q,N} \mathbf{H}_{Q,N})^{-1} \mathbf{H}'_{Q,N} \mathbf{Z}^*_{Q,N} \right]^{-1}. \\ &\quad \mathbf{Z}^*{}'_{Q,N} \mathbf{H}_{Q,N} (\mathbf{H}'_{Q,N} \mathbf{H}_{Q,N})^{-1}, \end{aligned} \quad (\text{A.2})$$

and

$$\begin{aligned} \mathbf{F}_{R1,N} &= \sigma_\mu \sigma_\nu \mathbf{H}'_{R,N} \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2} (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \\ &= \sigma_\mu \mathbf{H}'_{R,N} \left( \frac{\sigma_\nu}{\sigma_1} \mathbf{Q}_{1,N} + \mathbf{Q}_{0,N} \right) (\boldsymbol{\iota}_T \otimes \mathbf{I}_N), \\ \mathbf{F}_{R2,N} &= \sigma_\nu^2 \mathbf{H}'_{R,N} \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2} \\ &= \sigma_\nu \mathbf{H}'_{R,N} \left( \frac{\sigma_\nu}{\sigma_1} \mathbf{Q}_{1,N} + \mathbf{Q}_{0,N} \right), \\ \mathbf{F}_{Q,N} &= \sigma_\nu \mathbf{H}'_{Q,N} \mathbf{Q}_{0,N}. \end{aligned} \quad (\text{A.3})$$

By Assumptions (5) and (7) it follows that the sequence of the stochastic matrices  $(NT) \widetilde{\mathbf{P}}_{R,N}$  and  $(NT) \widetilde{\mathbf{P}}_{Q,N}$  converge in probability, i.e.

$$\begin{aligned} (NT) \widetilde{\mathbf{P}}_{R,N} &\xrightarrow{p} \left[ \mathbf{M}'_{HR\widetilde{Z}} \mathbf{M}^{-1}_{HR\widetilde{Z}} \mathbf{M}_{HR\widetilde{Z}} \right]^{-1} \mathbf{M}'_{HR\widetilde{Z}} \mathbf{M}^{-1}_{HR\widetilde{Z}}, \\ (NT) \widetilde{\mathbf{P}}_{Q,N} &\xrightarrow{p} \left[ \mathbf{M}'_{HQZ^*} \mathbf{M}^{-1}_{HQZ^*} \mathbf{M}_{HQZ^*} \right]^{-1} \mathbf{M}'_{HQZ^*} \mathbf{M}^{-1}_{HQZ^*}. \end{aligned} \quad (\text{A.4})$$

<sup>5</sup>We have used the properties of the  $\mathbf{Q}_{0,N}$  and  $\mathbf{Q}_{1,N}$  transformation matrices (see, e.g. Baltagi, 2008 and Kapoor, Kelejian and Prucha, 2007, Remark A1). In particular we have  $\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2} = \sigma_\nu^{-1} \mathbf{Q}_{1,N} + \sigma_1^{-1} \mathbf{Q}_{0,N}$ .

Next we apply the central limit theorem for vectors of triangular arrays given in Theorem A1 in Mutl (2006) to  $(NT)^{-\frac{1}{2}} \mathbf{F}_N \boldsymbol{\zeta}_N$ . By Assumptions 1 and 4, the vector of random variables  $\boldsymbol{\zeta}_N$  satisfies the assumptions of the central limit theorem. Observe that the matrix  $\mathbf{F}_N$  is non-stochastic and that Assumptions 2, 3 and 5 imply that the row and column sums of  $\mathbf{F}_N$  are uniformly bounded in absolute value. Hence, it remains to be demonstrated that the matrix  $(NT)^{-1} \mathbf{F}_N \mathbf{F}'_N$  has eigenvalues uniformly bounded away from zero.

One can show that<sup>6</sup>

$$\begin{aligned} (\mathbf{F}_N \mathbf{F}'_N)_{(p+2q) \times (p+2q)} &= \sigma_\nu^2 \begin{pmatrix} \mathbf{H}'_{R,N} \mathbf{H}_{R,N} & \mathbf{H}'_{R,N} \mathbf{H}_{Q,N} \\ \mathbf{H}'_{Q,N} \mathbf{H}_{R,N} & \mathbf{H}'_{Q,N} \mathbf{H}_{Q,N} \end{pmatrix} \\ &= \sigma_\nu^2 \begin{pmatrix} \mathbf{H}'_{R,N} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}'_{Q,N} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{NT} & \mathbf{I}_{NT} \\ \mathbf{I}_{NT} & \mathbf{I}_{NT} \end{pmatrix} \begin{pmatrix} \mathbf{H}_{R,N} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{Q,N} \end{pmatrix} \end{aligned} \quad (\text{A.6})$$

and hence

$$\begin{aligned} (NT)^{-1} \lambda_{\min}(\mathbf{F}_N \mathbf{F}'_N) &\geq \min \left[ (NT)^{-1} \lambda_{\min}(\mathbf{H}'_{R,N} \mathbf{H}_{R,N}), (NT)^{-1} \lambda_{\min}(\mathbf{H}'_{Q,N} \mathbf{H}_{Q,N}) \right] \\ &\quad \cdot \lambda_{\min} \left[ \begin{pmatrix} \mathbf{I}_{NT} & \mathbf{I}_{NT} \\ \mathbf{I}_{NT} & \mathbf{I}_{NT} \end{pmatrix} \right]. \end{aligned} \quad (\text{A.7})$$

Observe that  $(NT)^{-1} \mathbf{H}'_{R,N} \mathbf{H}_{R,N}$  and  $(NT)^{-1} \mathbf{H}'_{Q,N} \mathbf{H}_{Q,N}$  and  $\begin{pmatrix} \mathbf{I}_{NT} & \mathbf{I}_{NT} \\ \mathbf{I}_{NT} & \mathbf{I}_{NT} \end{pmatrix}$  are symmetric. By Assumptions (3.11) and (7) the first two matrices have full rank  $p + q$  and  $q$ , respectively. Note that the third matrix has trivially full rank as well. Hence,  $(NT)^{-1} \lambda_{\min}(\mathbf{F}_N \mathbf{F}'_N)$  is uniformly bounded away from zero. Therefore, by the central limit theorem it follows that<sup>7</sup>

$$NT^{-1/2} \mathbf{F}_N \boldsymbol{\zeta}_N \xrightarrow{d} N(\mathbf{0}, \lim_{N \rightarrow \infty} \frac{1}{NT} \mathbf{F}_N \mathbf{F}'_N). \quad (\text{A.8})$$

---

<sup>6</sup>Recall that the  $(NT \times q)$  matrix of within transformed instruments  $\mathbf{H}_{Q,N} = \mathbf{Q}_{0,N} \mathbf{G}_{0,N}$  has full column rank  $q$  and that  $\mathbf{Q}_{0,N} \mathbf{H}_{Q,N} = \mathbf{H}_{Q,N}$ . Furthermore,  $\mathbf{H}_{R,N} = [\mathbf{H}_{P,N}, \mathbf{H}_{Q,N}]$ , where  $\mathbf{H}_{P,N} = \mathbf{Q}_{1,N} \mathbf{G}_{1,N}$  with dimension  $(NT \times p)$  and full column rank  $p$ . Since  $\mathbf{Q}_{0,N} \mathbf{Q}_{1,N} = \mathbf{0}$ , it follows that

$$\begin{aligned} \mathbf{H}'_{R,N} \mathbf{Q}_{0,N} \mathbf{H}_{Q,N} &= \mathbf{H}'_{R,N} \mathbf{H}_{Q,N} \\ &= \begin{pmatrix} \mathbf{G}'_{1,N} \mathbf{Q}_{1,N} \\ \mathbf{G}'_{0,N} \mathbf{Q}_{0,N} \end{pmatrix} \mathbf{Q}_{0,N} \mathbf{H}_{Q,N} = \begin{pmatrix} \mathbf{0}_{p \times q} \\ \mathbf{H}'_{Q,N} \mathbf{H}_{Q,N} \end{pmatrix}. \end{aligned} \quad (\text{A.5})$$

<sup>7</sup>Note that it can be demonstrated that  $\lim_{N \rightarrow \infty} (NT)^{-1} \mathbf{F}_N \mathbf{F}'_N$  exists.

From Assumptions (3.11) and (7) we also have

$$(NT)^{\frac{1}{2}} \begin{pmatrix} \widehat{\boldsymbol{\delta}}_{GLS,N} - \boldsymbol{\delta} \\ \widehat{\boldsymbol{\theta}}_{W,N} - \boldsymbol{\theta} \end{pmatrix} \xrightarrow{d} N(\mathbf{0}, \Delta), \quad (\text{A.9})$$

where  $\Delta = p \lim_{N \rightarrow \infty} (NT) \widetilde{\mathbf{P}}_N \mathbf{F}_N \mathbf{F}'_N \widetilde{\mathbf{P}}'_N$ . Fairly straightforward calculation shows that

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta'_{12} & \Delta_{22} \end{pmatrix} \quad (\text{A.10})$$

with

$$\begin{aligned} \Delta_{11} &= \sigma_\nu^2 \left( \mathbf{M}'_{H_R \tilde{Z}} \mathbf{M}^{-1}_{H_R H_R} \mathbf{M}_{H_R \tilde{Z}} \right)^{-1} = \begin{pmatrix} \sigma_{\alpha, GLS}^2 & \boldsymbol{\Sigma}_{\alpha, GLS} \\ \boldsymbol{\Sigma}'_{\alpha, GLS} & \boldsymbol{\Sigma}_{GLS} \end{pmatrix} \quad (\text{A.11}) \\ \Delta_{12} &= \sigma_\nu^2 \left( \mathbf{M}'_{H_R \tilde{Z}} \mathbf{M}^{-1}_{H_R H_R} \mathbf{M}_{H_R \tilde{Z}} \right)^{-1} \begin{pmatrix} \mathbf{0}_{1 \times K} \\ \mathbf{I}_K \end{pmatrix}_{((K+1) \times K)} = \begin{pmatrix} \boldsymbol{\Sigma}_{\alpha, GLS} \\ \boldsymbol{\Sigma}_{GLS} \end{pmatrix} \\ \Delta_{22} &= \sigma_\nu^2 \left( \mathbf{M}'_{H_Q Z^*} \mathbf{M}^{-1}_{H_Q H_Q} \mathbf{M}_{H_Q Z^*} \right)^{-1} = \boldsymbol{\Sigma}_W. \end{aligned}$$

We have ordered the elements of the vector of parameters  $\boldsymbol{\delta}$  such that the first element is the constant so that the asymptotic variance-covariance matrix of the stacked estimators becomes

$$\Delta = \begin{bmatrix} \sigma_{\alpha, GLS}^2 & \boldsymbol{\Sigma}_{\alpha, GLS} & \boldsymbol{\Sigma}_{\alpha, GLS} \\ \boldsymbol{\Sigma}'_{\alpha, GLS} & \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_{GLS} \\ \boldsymbol{\Sigma}'_{\alpha, GLS} & \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_W \end{bmatrix}, \quad (\text{A.12})$$

and hence

$$\sqrt{NT} \begin{pmatrix} \alpha_{GLS,N} - \alpha \\ \widehat{\boldsymbol{\theta}}_{GLS,N} - \boldsymbol{\theta} \\ \widehat{\boldsymbol{\theta}}_{W,N} - \boldsymbol{\theta} \end{pmatrix} \xrightarrow{d} N \left( \mathbf{0}, \begin{bmatrix} \sigma_{\alpha, GLS}^2 & \boldsymbol{\Sigma}_{\alpha, GLS} & \boldsymbol{\Sigma}_{\alpha, GLS} \\ \boldsymbol{\Sigma}'_{\alpha, GLS} & \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_{GLS} \\ \boldsymbol{\Sigma}'_{\alpha, GLS} & \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_W \end{bmatrix} \right). \quad (\text{A.13})$$

### Proof of Proposition 1:

Inspection of the proof of Theorem 1 above reveals that it remains valid after replacing  $\widehat{\boldsymbol{\theta}}_{W,N}$  with  $\widehat{\boldsymbol{\theta}}_{I,N}$  and Assumption 6(b) with the one made in the proposition. It then trivially implies the claim under Assumption 4 (RE). It

thus remains to be shown that the initial estimator converges in distribution under the Assumption 6 (FE). Note that we have

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{I,N} - \boldsymbol{\theta} &= \left[ \mathbf{Z}'_{Q,N} \mathbf{H}_{Q,N} (\mathbf{H}'_{Q,N} \mathbf{H}_{Q,N})^{-1} \mathbf{H}'_{Q,N} \mathbf{Z}_{Q,N} \right]^{-1} \quad (\text{A.14}) \\ &\quad \cdot \mathbf{Z}'_{Q,N} \mathbf{H}_{Q,N} (\mathbf{H}'_{Q,N} \mathbf{H}_{Q,N})^{-1} \mathbf{H}'_{Q,N} \mathbf{Q}_{0,N} \boldsymbol{\nu}_N \\ &= \widetilde{\mathbf{P}}_{Q,N} \mathbf{H}'_{Q,N} \mathbf{Q}_{0,N} \boldsymbol{\nu}_N.\end{aligned}$$

Given Assumption 7 and the condition in the proposition, it follows that  $(NT) \widetilde{\mathbf{P}}_{Q,N} \xrightarrow{p} \left[ \mathbf{M}'_{HQZ} \mathbf{M}_{HQ}^{-1} \mathbf{M}_{HQZ} \right]^{-1} \mathbf{M}'_{HQZ} \mathbf{M}_{HQ}^{-1}$  and

$$(NT) \lambda_{\min} (\mathbf{H}'_{Q,N} \mathbf{Q}_{0,N} \mathbf{Q}'_{0,N} \mathbf{H}_{Q,N}) = (NT) \lambda_{\min} (\mathbf{H}_{Q,N} \mathbf{H}'_{Q,N}), \quad (\text{A.15})$$

which is uniformly bounded away from zero. Given Assumption 1, the conditions of Theorem A1 in Mutl (2006) are satisfied and in light of Assumption 7(a), we have the desired result.

### Proof of Proposition 2:

The proof follows closely the proof of Theorem 4, part 2 in Kapoor, Kelejian and Prucha (2007) and Theorem 3 in Mutl (2006). In particular, it will be sufficient to show that (see e.g. Schmidt, 1976)

$$\begin{aligned}\Delta_{G1,N} &= (NT)^{-1} \left[ \widehat{\mathbf{Z}}_N (\widehat{\boldsymbol{\vartheta}})' \widetilde{\mathbf{Z}}_N (\widehat{\boldsymbol{\vartheta}}) - \widehat{\mathbf{Z}}_N (\boldsymbol{\vartheta})' \widetilde{\mathbf{Z}}_N (\boldsymbol{\vartheta}) \right] \xrightarrow{p} 0 \quad (\text{A.16}) \\ \Delta_{W1,N} &= (NT)^{-1} \left[ \widehat{\mathbf{Z}}_N^* (\widehat{\boldsymbol{\rho}})' \mathbf{Z}_N^* (\widehat{\boldsymbol{\rho}}) - \widehat{\mathbf{Z}}_N^* (\boldsymbol{\rho})' \mathbf{Z}_N^* (\boldsymbol{\rho}) \right] \xrightarrow{p} 0,\end{aligned}$$

and

$$\begin{aligned}\Delta_{G2,N} &= (NT)^{-1/2} \left( \widehat{\mathbf{Z}}_N' (\widehat{\boldsymbol{\vartheta}}) \widetilde{\mathbf{u}}_N (\widehat{\boldsymbol{\vartheta}}) - \widehat{\mathbf{Z}}_N' (\boldsymbol{\vartheta}) \widetilde{\mathbf{u}}_N (\boldsymbol{\vartheta}) \right) \xrightarrow{p} 0 \quad (\text{A.17}) \\ \Delta_{W2,N} &= (NT)^{-1/2} \left( \widehat{\mathbf{Z}}_N^* (\widehat{\boldsymbol{\rho}})' \mathbf{u}^* (\widehat{\boldsymbol{\rho}}) - \widehat{\mathbf{Z}}_N^* (\boldsymbol{\rho})' \mathbf{u}^* (\boldsymbol{\rho}) \right) \xrightarrow{p} 0,\end{aligned}$$

where  $\widehat{\boldsymbol{\vartheta}}$  and  $\widehat{\boldsymbol{\rho}}$  are (any) consistent estimators of  $\boldsymbol{\vartheta}$  and  $\boldsymbol{\rho}$ . Note that  $\widehat{\mathbf{Z}}_N (\widehat{\boldsymbol{\vartheta}}) = \mathbf{P}_{HR,N} \widetilde{\mathbf{Z}}_N (\widehat{\boldsymbol{\vartheta}})$  and  $\widehat{\mathbf{Z}}_N^* (\widehat{\boldsymbol{\rho}}) = \mathbf{P}_{HQ,N} \mathbf{Z}_N^* (\widehat{\boldsymbol{\rho}})$ , where

$$\begin{aligned}\widetilde{\mathbf{Z}}_N (\widehat{\boldsymbol{\vartheta}}) &= \widehat{\sigma}_\nu \widehat{\boldsymbol{\Omega}}_{\varepsilon,N}^{-1/2} (\mathbf{I}_{NT} - \widehat{\boldsymbol{\rho}} \mathbf{M}_N) \mathbf{Z}_N \quad (\text{A.18}) \\ &= \left( \mathbf{Q}_{1,N} + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Q}_{0,N} \right) (\mathbf{I}_{NT} - \widehat{\boldsymbol{\rho}} \mathbf{M}_N) \mathbf{Z}_N,\end{aligned}$$

and

$$\mathbf{Z}_N^* \left( \widehat{\boldsymbol{\vartheta}} \right) = \mathbf{Q}_{0,N} \left( \mathbf{I}_{NT} - \widehat{\rho} \mathbf{M}_N \right) \mathbf{Z}_N. \quad (\text{A.19})$$

Hence

$$\begin{aligned} & \Delta_{G1,N} \quad (\text{A.20}) \\ = & \frac{\rho - \widehat{\rho}}{NT} \mathbf{Z}'_N \mathbf{M}'_N \left( \mathbf{Q}_{1,N} + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Q}_{0,N} \right) \mathbf{P}_{H_{R,N}} \left( \mathbf{Q}_{1,N} + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Q}_{0,N} \right) \mathbf{Z}_N \\ & + \frac{\rho - \widehat{\rho}}{NT} \mathbf{Z}'_N \left( \mathbf{Q}_{1,N} + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Q}_{0,N} \right) \mathbf{P}_{H_{R,N}} \left( \mathbf{Q}_{1,N} + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Q}_{0,N} \right) \mathbf{M}_N \mathbf{Z}_N \\ & + \frac{\widehat{\rho}^2 - \rho^2}{NT} \mathbf{Z}'_N \mathbf{M}'_N \left( \mathbf{Q}_{1,N} + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Q}_{0,N} \right) \mathbf{P}_{H_{R,N}} \left( \mathbf{Q}_{1,N} + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Q}_{0,N} \right) \mathbf{M}_N \mathbf{Z}_N \\ & + \frac{\widehat{\sigma}_\nu - \sigma_\nu}{\widehat{\sigma}_1 - \sigma_1} \mathbf{Z}'_N \left( \mathbf{I}_{NT} - \rho \mathbf{M}'_N \right) \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N} \left( \mathbf{I}_{NT} - \rho \mathbf{M}_N \right) \mathbf{Z}_N \\ & + \frac{\widehat{\sigma}_\nu - \sigma_\nu}{\widehat{\sigma}_1 - \sigma_1} \mathbf{Z}'_N \left( \mathbf{I}_{NT} - \rho \mathbf{M}'_N \right) \mathbf{Q}_{1,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N} \left( \mathbf{I}_{NT} - \rho \mathbf{M}_N \right) \mathbf{Z}_N \\ & + \frac{\widehat{\sigma}_\nu^2 - \sigma_\nu^2}{\widehat{\sigma}_1^2 - \sigma_1^2} \mathbf{Z}'_N \left( \mathbf{I}_{NT} - \rho \mathbf{M}'_N \right) \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N} \left( \mathbf{I}_{NT} - \rho \mathbf{M}_N \right) \mathbf{Z}_N = \\ & \frac{\rho - \widehat{\rho}}{NT} \left( \mathbf{Z}'_N \mathbf{A}_{G1,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G2,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G3,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2} \mathbf{Z}'_N \mathbf{A}_{G4,N} \mathbf{Z}_N \right) \\ & + \frac{\rho - \widehat{\rho}}{NT} \left( \mathbf{Z}'_N \mathbf{A}'_{G1,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}'_{G2,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}'_{G3,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2} \mathbf{Z}'_N \mathbf{A}'_{G4,N} \mathbf{Z}_N \right) \\ & + \frac{\widehat{\rho}^2 - \rho^2}{NT} \left( \mathbf{Z}'_N \mathbf{A}_{G5,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G6,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G7,N} \mathbf{Z}_N + \frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2} \mathbf{Z}'_N \mathbf{A}_{G8,N} \mathbf{Z}_N \right) \\ & + \frac{\widehat{\sigma}_\nu - \sigma_\nu}{\widehat{\sigma}_1 - \sigma_1} \left( \mathbf{Z}'_N \mathbf{A}_{G9,N} \mathbf{Z}_N + \mathbf{Z}'_N \mathbf{A}'_{G9,N} \mathbf{Z}_N \right) \\ & + \frac{\widehat{\sigma}_\nu^2 - \sigma_\nu^2}{\widehat{\sigma}_1^2 - \sigma_1^2} \mathbf{Z}'_N \mathbf{A}_{G10,N} \mathbf{Z}_N, \end{aligned}$$

where

$$\begin{aligned} \mathbf{A}_{G1,N} &= \mathbf{M}'_N \mathbf{Q}_{1,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N}, \\ \mathbf{A}_{G2,N} &= \mathbf{M}'_N \mathbf{Q}_{1,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N}, \\ \mathbf{A}_{G3,N} &= \mathbf{M}'_N \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N}, \\ \mathbf{A}_{G4,N} &= \mathbf{M}'_N \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N}, \end{aligned} \quad (\text{A.21})$$

$$\begin{aligned}
\mathbf{A}_{G5,N} &= \mathbf{M}'_N \mathbf{Q}_{1,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N} \mathbf{M}_N, \\
\mathbf{A}_{G6,N} &= \mathbf{M}'_N \mathbf{Q}_{1,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N} \mathbf{M}_N, \\
\mathbf{A}_{G7,N} &= \mathbf{M}'_N \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N} \mathbf{M}_N, \\
\mathbf{A}_{G8,N} &= \mathbf{M}'_N \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N} \mathbf{M}_N,
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{G9,N} &= (\mathbf{I}_{NT} - \rho \mathbf{M}'_N) \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N} (\mathbf{I}_{NT} - \rho \mathbf{M}_N), \\
\mathbf{A}_{G10,N} &= (\mathbf{I}_{NT} - \rho \mathbf{M}'_N) \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N} (\mathbf{I}_{NT} - \rho \mathbf{M}_N)
\end{aligned}$$

Analogically,

$$\begin{aligned}
\Delta_{W1,N} &= (\rho - \widehat{\rho}) (NT)^{-1} \mathbf{Z}'_N \mathbf{M}'_N \mathbf{P}_{H_{Q,N}} \mathbf{Z}_N & (\text{A.22}) \\
&\quad + (\rho - \widehat{\rho}) (NT)^{-1} \mathbf{Z}'_N \mathbf{P}_{H_{Q,N}} \mathbf{M}_N \mathbf{Z}_N \\
&\quad + (\widehat{\rho}^2 - \rho^2) (NT)^{-1} \mathbf{Z}'_N \mathbf{M}'_N \mathbf{P}_{H_{Q,N}} \mathbf{M}_N \mathbf{Z}_N
\end{aligned}$$

$$= (\rho - \widehat{\rho}) (NT)^{-1} (\mathbf{Z}'_N \mathbf{A}_{W1,N} \mathbf{Z}_N + \mathbf{Z}'_N \mathbf{A}'_{W1,N} \mathbf{Z}_N) + (\widehat{\rho}^2 - \rho^2) (NT)^{-1} \mathbf{Z}'_N \mathbf{A}_{W2,N} \mathbf{Z}_N,$$

where

$$\begin{aligned}
\mathbf{A}_{W1,N} &= \mathbf{M}'_N \mathbf{P}_{H_{Q,N}}, & (\text{A.23}) \\
\mathbf{A}_{W2,N} &= \mathbf{M}'_N \mathbf{P}_{H_{Q,N}} \mathbf{M}_N.
\end{aligned}$$

In light of Assumptions 2, 5 and 7 the row and column sums of the matrices  $\mathbf{A}_{G_i,N}$  and  $\mathbf{A}_{W_j,N}$  are uniformly bounded in absolute value. By Lemma C3 in Mutl (2006) we then have that  $(NT)^{-1} \mathbf{X}'_N \mathbf{A}_{G_i,N} \mathbf{X}_N$ ,  $(NT)^{-1} \mathbf{X}'_N \mathbf{A}_{W_i,N} \mathbf{X}_N$ ,  $(NT)^{-1} \boldsymbol{\iota}'_N \mathbf{A}_{G_i,N} \boldsymbol{\iota}_N$  and  $(NT)^{-1} \boldsymbol{\iota}'_N \mathbf{A}_{W_i,N} \boldsymbol{\iota}_N$  have elements uniformly bounded in absolute value. Furthermore, by Lemma B2 in Mutl (2006), using Assumptions 1, 2 and 4, the elements of  $\mathbf{u}_N$  have uniformly bounded 4-th moments. Therefore, the variance of  $(NT)^{-1} \mathbf{u}'_N \mathbf{A}_{G_i,N} \mathbf{u}_N$  is uniformly bounded in absolute value. Recall that  $\mathbf{Z}_N = (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N, \boldsymbol{\iota}_N)$ , where the solution of the model yields  $\mathbf{y}_N = (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} (\mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT} + \mathbf{u}_N)$ . We thus have that the elements of

$$(NT)^{-2} E (\mathbf{Z}'_N \mathbf{A}_{G_i,N} \mathbf{Z}_N \mathbf{Z}'_N \mathbf{A}_{G_i,N} \mathbf{Z}_N),$$

and

$$(NT)^{-2} E (\mathbf{Z}'_N \mathbf{A}_{W_i,N} \mathbf{Z}_N \mathbf{Z}'_N \mathbf{A}_{W_i,N} \mathbf{Z}_N),$$

are uniformly bounded in absolute value. Since  $\widehat{\rho}$ ,  $\widehat{\sigma}_v$ , and  $\widehat{\sigma}_1$  are consistent estimators, it follows that  $\Delta_{G1,N} \xrightarrow{p} 0$  and  $\Delta_{W1,N} \xrightarrow{p} 0$  as  $N \rightarrow \infty$ .

Next we use similar derivations to obtain

$$\begin{aligned}
\Delta_{G2,N} = & \tag{A.24} \\
& \frac{\rho - \hat{\rho}}{(NT)^{1/2}} \left( \mathbf{Z}'_N \mathbf{A}_{G1,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu}{\hat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G2,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu}{\hat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G3,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu^2}{\hat{\sigma}_1^2} \mathbf{Z}'_N \mathbf{A}_{G4,N} \mathbf{u}_N \right) \\
& + \frac{\rho - \hat{\rho}}{(NT)^{1/2}} \left( \mathbf{Z}'_N \mathbf{A}'_{G1,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu}{\hat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}'_{G2,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu}{\hat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}'_{G3,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu^2}{\hat{\sigma}_1^2} \mathbf{Z}'_N \mathbf{A}'_{G4,N} \mathbf{u}_N \right) \\
& + \frac{\hat{\rho}^2 - \rho^2}{(NT)^{1/2}} \left( \mathbf{Z}'_N \mathbf{A}_{G5,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu}{\hat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G6,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu}{\hat{\sigma}_1} \mathbf{Z}'_N \mathbf{A}_{G7,N} \mathbf{u}_N + \frac{\hat{\sigma}_\nu^2}{\hat{\sigma}_1^2} \mathbf{Z}'_N \mathbf{A}_{G8,N} \mathbf{u}_N \right) \\
& + \frac{\frac{\hat{\sigma}_\nu}{\hat{\sigma}_1} - \frac{\sigma_\nu}{\sigma_1}}{(NT)^{1/2}} \left( \mathbf{Z}'_N \mathbf{A}_{G9,N} \mathbf{u}_N + \mathbf{Z}'_N \mathbf{A}'_{G9,N} \mathbf{u}_N \right) \\
& + \frac{\frac{\hat{\sigma}_\nu^2}{\hat{\sigma}_1^2} - \frac{\sigma_\nu^2}{\sigma_1^2}}{(NT)^{1/2}} \mathbf{Z}'_N \mathbf{A}_{G10,N} \mathbf{u}_N,
\end{aligned}$$

and

$$\begin{aligned}
\Delta_{W2,N} = & \frac{\rho - \hat{\rho}}{(NT)^{1/2}} \left( \mathbf{Z}'_N \mathbf{A}_{W1,N} \mathbf{u}_N + \mathbf{Z}'_N \mathbf{A}'_{W1,N} \mathbf{u}_N \right) \tag{A.25} \\
& + \frac{\hat{\rho}^2 - \rho^2}{(NT)^{1/2}} \mathbf{Z}'_N \mathbf{A}_{W2,N} \mathbf{u}_N,
\end{aligned}$$

where  $\mathbf{Z}_N = (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N, \boldsymbol{\iota}_N)$ , with  $\mathbf{y}_N = (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} (\mathbf{X}_N \boldsymbol{\beta} + \alpha \boldsymbol{\iota}_{NT} + \mathbf{u}_N)$ . Observe that since the matrices  $\mathbf{A}_{G_i,N}$  and  $\mathbf{A}_{W_j,N}$ , as well as the vectors  $\mathbf{X}_N$  and  $\boldsymbol{\iota}_{NT}$  are non-stochastic. We have that

$$\begin{aligned}
E \left[ (NT)^{-1/2} \boldsymbol{\beta}' \mathbf{X}'_N \mathbf{A}_{G_i,N} \mathbf{u}_N \right] &= 0, \tag{A.26} \\
E \left[ (NT)^{-1/2} \alpha \boldsymbol{\iota}'_{NT} \mathbf{A}_{G_i,N} \mathbf{u}_N \right] &= 0, \\
E \left[ (NT)^{-1/2} \boldsymbol{\beta}' \mathbf{X}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{G_i,N} \mathbf{u}_N \right] &= 0, \\
E \left[ (NT)^{-1/2} \alpha \boldsymbol{\iota}'_{NT} (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{G_i,N} \mathbf{u}_N \right] &= 0,
\end{aligned}$$

$$\begin{aligned}
E \left[ (NT)^{-1/2} \boldsymbol{\beta}' \mathbf{X}'_N \mathbf{A}_{Wj,N} \mathbf{u}_N \right] &= 0, \\
E \left[ (NT)^{-1/2} \alpha \boldsymbol{\iota}'_{NT} \mathbf{A}_{Wj,N} \mathbf{u}_N \right] &= 0, \\
E \left[ (NT)^{-1/2} \boldsymbol{\beta}' \mathbf{X}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} \mathbf{u}_N \right] &= 0, \\
E \left[ (NT)^{-1/2} \alpha \boldsymbol{\iota}'_{NT} (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{A}_{Wj,N} \mathbf{W}'_N \mathbf{u}_N \right] &= 0.
\end{aligned}$$

and

$$\begin{aligned}
&E \left[ (NT)^{-1/2} \mathbf{u}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} \mathbf{u}_N \right] \tag{A.27} \\
&= \frac{\sigma_u^2}{(NT)^{1/2}} tr \left[ (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}_N)^{-1} \mathbf{Q}_{0,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}'_N)^{-1} \right] \\
&\quad + \frac{\sigma_1^2}{(NT)^{1/2}} tr \left[ (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}_N)^{-1} \mathbf{Q}_{1,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}'_N)^{-1} \right],
\end{aligned}$$

$$\begin{aligned}
&E \left[ (NT)^{-1/2} \mathbf{u}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} \mathbf{u}_N \right] \tag{A.28} \\
&= \frac{\sigma_u^2}{(NT)^{1/2}} (NT)^{-1/2} tr \left[ (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}_N)^{-1} \mathbf{Q}_{0,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}'_N)^{-1} \right] \\
&\quad + \frac{\sigma_1^2}{(NT)^{1/2}} (NT)^{-1/2} tr \left[ (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}_N)^{-1} \mathbf{Q}_{1,N} (\mathbf{I}_{NT} - \lambda \mathbf{M}'_N)^{-1} \right],
\end{aligned}$$

which by Assumptions 2, 5 and 7 are uniformly bounded in absolute value.

Therefore, the elements of  $E \left[ (NT)^{-1/2} \mathbf{Z}'_N \mathbf{A}_{Gj,N} \mathbf{u}_N \right]$ , and  $E \left[ (NT)^{-1/2} \mathbf{Z}'_N \mathbf{A}_{Wj,N} \mathbf{u}_N \right]$  are uniformly bounded in absolute value.

Next consider the corresponding variance covariance matrices:

$$\begin{aligned}
(NT)^{-1} \boldsymbol{\beta}' \mathbf{X}'_N \mathbf{A}_{Gi,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Gi,N} \mathbf{X}_N \boldsymbol{\beta}, \tag{A.29} \\
(NT)^{-1} \alpha^2 \boldsymbol{\iota}'_{NT} \mathbf{A}_{Gi,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Gi,N} \boldsymbol{\iota}_{NT}, \\
(NT)^{-1} \boldsymbol{\beta}' \mathbf{X}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Gi,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \mathbf{X}_N \boldsymbol{\beta}, \\
(NT)^{-1} \alpha^2 \boldsymbol{\iota}'_{NT} (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Gi,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \boldsymbol{\iota}_{NT},
\end{aligned}$$

and

$$\begin{aligned}
(NT)^{-1} \boldsymbol{\beta}' \mathbf{X}'_N \mathbf{A}_{Wj,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Wj,N} \mathbf{X}_N \boldsymbol{\beta}, \\
(NT)^{-1} \alpha^2 \boldsymbol{\iota}'_{NT} \mathbf{A}_{Wj,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Wj,N} \boldsymbol{\iota}_{NT}, \\
(NT)^{-1} \boldsymbol{\beta}' \mathbf{X}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Wj,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \mathbf{X}_N \boldsymbol{\beta}, \\
(NT)^{-1} \alpha^2 \boldsymbol{\iota}'_{NT} (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} \boldsymbol{\Omega}_{u,N} \mathbf{A}'_{Wj,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \boldsymbol{\iota}_{NT}.
\end{aligned}$$

Finally, the two remaining scalar variances are given by

$$\begin{aligned}
& (NT)^{-1} E \left[ \mathbf{u}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \mathbf{u}_N \mathbf{u}'_N \right. \\
& \quad \left. (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \mathbf{u}_N \right] \\
= & (NT)^{-1} E \left( \text{tr} \left[ (\mathbf{I}_{NT} - \lambda \mathbf{M}'_N)^{-1} (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Gi,N} \right. \right. \\
& \quad \cdot \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} (\mathbf{I}_{NT} - \lambda \mathbf{M}_N)^{-1} \boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}'_N (\mathbf{I}_{NT} - \lambda \mathbf{M}'_N)^{-1} (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \\
& \quad \left. \left. \cdot \mathbf{A}_{Gi,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} (\mathbf{I}_{NT} - \lambda \mathbf{M}_N)^{-1} \boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}'_N \right] \right) \\
= & (NT)^{-1} E \left( \text{tr} [\mathbf{B}_{1,Gi,N} \boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}'_N \mathbf{B}_{2,Gi,N} \boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}'_N] \right), \tag{A.30}
\end{aligned}$$

and

$$\begin{aligned}
& (NT)^{-1} E \left[ \mathbf{u}'_N (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \mathbf{u}_N \mathbf{u}'_N \right. \\
& \quad \left. \cdot (\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1} \mathbf{W}'_N \mathbf{A}_{Wj,N} \mathbf{W}_N (\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1} \mathbf{u}_N \right] \tag{A.31} \\
= & (NT)^{-1} E \left( \text{tr} [\mathbf{B}_{1,Wj,N} \boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}'_N \mathbf{B}_{2,Wj,N} \boldsymbol{\varepsilon}_N \boldsymbol{\varepsilon}'_N] \right),
\end{aligned}$$

where the  $\mathbf{B}$  matrices are products of  $\mathbf{A}$ ,  $\mathbf{W}_N$ ,  $\mathbf{W}'_N$ ,  $(\mathbf{I}_{NT} - \lambda \mathbf{W}_N)^{-1}$ , and  $(\mathbf{I}_{NT} - \lambda \mathbf{W}'_N)^{-1}$  matrices and hence by Assumptions 2, 5 and 7 have row and column sums uniformly bounded in absolute value. Hence the row and column sums of the variance covariance matrices are uniformly bounded in absolute value and, therefore,  $E \left[ (NT)^{-1/2} \mathbf{Z}'_N \mathbf{A}_{Gj,N} \mathbf{u}_N \right] = O_P(1)$ , and  $E \left[ (NT)^{-1/2} \mathbf{Z}'_N \mathbf{A}_{Wj,N} \mathbf{u}_N \right] = O_P(1)$ . Thus  $\Delta_{G2,N} \xrightarrow{p} 0$  and  $\Delta_{W2,N} \xrightarrow{p} 0$  as  $N \rightarrow \infty$ , since  $\hat{\rho}$ ,  $\hat{\sigma}_\nu$ , and  $\hat{\sigma}_1$  are consistent estimators.

### Proof of Theorem 2

Part (a): From Theorem 1 it follows that  $\sqrt{NT} \left( \hat{\boldsymbol{\theta}}_{GLS,N} - \hat{\boldsymbol{\theta}}_{W,N} \right) \xrightarrow{d} N(0, \boldsymbol{\Psi})$ , where

$$\boldsymbol{\Psi} = (\mathbf{I}_K, -\mathbf{I}_K) \begin{pmatrix} \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_{GLS} \\ \boldsymbol{\Sigma}_{GLS} & \boldsymbol{\Sigma}_W \end{pmatrix} \begin{pmatrix} \mathbf{I}_K \\ -\mathbf{I}_K \end{pmatrix} = \boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}. \tag{A.32}$$

To show that  $\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}$  is positive definite, recall that

$$\begin{aligned}
\boldsymbol{\Sigma}_W &= \sigma_\nu^2 \lim_{N \rightarrow \infty} \left[ NT (\mathbf{Z}_N^{*'} \mathbf{P}_{\mathbf{H}_{Q,N}} \mathbf{Z}_N^*)^{-1} \right], \tag{A.33} \\
\boldsymbol{\Sigma}_{GLS} &= \sigma_\nu^2 (\mathbf{0}_{K \times 1}, \mathbf{I}_K) \lim_{N \rightarrow \infty} \left[ NT \left( \tilde{\mathbf{Z}}_N' \mathbf{P}_{\mathbf{H}_{R,N}} \tilde{\mathbf{Z}}_N \right)^{-1} \right] \cdot (\mathbf{0}_{K \times 1}, \mathbf{I}_K)'.
\end{aligned}$$

Since the instrument sets are given by  $\mathbf{H}_{Q,N} = \mathbf{Q}_{0,N}\mathbf{G}_{0,N}$ , and  $\mathbf{H}_{R,N} = [\mathbf{H}_{Q,N}, \mathbf{H}_{P,N}]$  one can show that  $\mathbf{P}_{\mathbf{H}_{R,N}} = \mathbf{P}_{\mathbf{H}_{P,N}} + \mathbf{P}_{\mathbf{H}_{Q,N}}$  and hence denoting  $\mathbf{S}_{R,N} = \tilde{\mathbf{Z}}_N' \mathbf{P}_{\mathbf{H}_{R,N}} \tilde{\mathbf{Z}}_N$ , we have

$$\mathbf{S}_{R,N} = \begin{bmatrix} \mathbf{S}_{R,N,11} & \mathbf{S}_{R,N,21} \\ \mathbf{S}_{R,N,12} & \mathbf{S}_{R,N,22} \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{S}_{R,N,11} &= \boldsymbol{\iota}'_{NT} \mathbf{A}'_N (\mathbf{P}_{\mathbf{H}_{Q,N}} + \phi^2 \mathbf{P}_{\mathbf{H}_{P,N}}) \mathbf{A}_N \boldsymbol{\iota}_{NT} & (\text{A.34}) \\ \mathbf{S}_{R,N,12} &= \mathbf{S}'_{R,N,21} = (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}'_N (\mathbf{P}_{\mathbf{H}_{Q,N}} + \phi^2 \mathbf{P}_{\mathbf{H}_{P,N}}) \mathbf{A}_N \boldsymbol{\iota}_{NT} \\ \mathbf{S}_{R,N,22} &= (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}'_N (\mathbf{P}_{\mathbf{H}_{Q,N}} + \phi^2 \mathbf{P}_{\mathbf{H}_{P,N}}) \mathbf{A}_N (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N), \end{aligned}$$

with  $\mathbf{A}_N = \mathbf{I}_{NT} - \rho \mathbf{M}_N$  and  $\phi = \frac{\sigma_\nu}{\sigma_1}$ . We are interested in the lower-right  $K \times K$  block of the inverse of  $\mathbf{S}_{R,N}$  which we denote by  $\mathbf{S}_{R,N}^{22}$ . Using the formula for partitioned inverses (see e.g. section 0.7.3 in Horn and Johnson, 1985) yields after some manipulation

$$\mathbf{S}_{R,N}^{22} = (\mathbf{S}_{R,N,22} - \mathbf{S}_{R,N,21} \mathbf{S}_{R,N,11}^{-1} \mathbf{S}_{R,N,12})^{-1} \quad (\text{A.35})$$

$$\begin{aligned} &= \{ (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}'_N \mathbf{P}_{\mathbf{H}_{Q,N}} \mathbf{A}_N (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N) \\ &\quad + \phi^2 (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}'_N \mathbf{P}_{\mathbf{H}_{P,N}} \left[ \mathbf{I}_{NT} - \mathbf{A}_N \mathbf{J}_{NT} \mathbf{A}'_N (\boldsymbol{\iota}'_{NT} \mathbf{A}'_N \mathbf{A}_N \boldsymbol{\iota}_{NT})^{-1} \right] \cdot \\ &\quad \mathbf{P}_{\mathbf{H}_{P,N}} \mathbf{A}_N (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N) \}^{-1}. \end{aligned}$$

Defining  $\mathbf{S}_{Q,N} = \mathbf{Z}_N^* \mathbf{P}_{\mathbf{H}_{Q,N}} \mathbf{Z}_N^*$  we have

$$\mathbf{S}_{Q,N} = (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}'_N \mathbf{P}_{\mathbf{H}_{Q,N}} \mathbf{A}_N (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N),$$

so that

$$\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS} = \sigma_\nu^2 \lim_{N \rightarrow \infty} NT (\mathbf{S}_{Q,N}^{-1} - \mathbf{S}_{R,N}^{22}) = \lim_{N \rightarrow \infty} NT (\mathbf{C}_N^{-1} - (\mathbf{C}_N + \mathbf{D}_N)^{-1}) \quad (\text{A.36})$$

where

$$\mathbf{C}_N = (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}'_N \mathbf{P}_{\mathbf{H}_{Q,N}} \mathbf{A}_N (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N) \quad (\text{A.37})$$

$$\begin{aligned} \mathbf{D}_N &= \phi^2 (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}'_N \mathbf{P}_{\mathbf{H}_{P,N}} \left[ \mathbf{I}_{NT} - \mathbf{A}_N \mathbf{J}_{NT} \mathbf{A}'_N (\boldsymbol{\iota}'_{NT} \mathbf{A}'_N \mathbf{A}_N \boldsymbol{\iota}_{NT})^{-1} \right] \cdot \\ &\quad \mathbf{P}_{\mathbf{H}_{P,N}} \mathbf{A}_N (\mathbf{W}_N \mathbf{y}_N, \mathbf{X}_N) \end{aligned} \quad (\text{A.38})$$

Using Greene (2003, p. 822) it follows that

$$(\mathbf{C}_N + \mathbf{D}_N)^{-1} = \mathbf{C}_N^{-1} - \mathbf{C}_N^{-1} (\mathbf{D}_N^{-1} + \mathbf{C}_N^{-1})^{-1} \mathbf{C}_N^{-1} \quad (\text{A.39})$$

and hence

$$\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS} = \sigma_\nu^2 \lim_{N \rightarrow \infty} \left( \mathbf{C}_N^{-1} (\mathbf{D}_N^{-1} + \mathbf{C}_N^{-1})^{-1} \mathbf{C}_N^{-1} \right), \quad (\text{A.40})$$

which is clearly positive definite.

Part (b): This follows directly from Proposition 2.

Part (c): Since  $\widehat{\boldsymbol{\vartheta}} = (\widehat{\rho}_N, \widehat{\sigma}_{\nu,N}^2, \widehat{\sigma}_{1,N}^2)'$  is a consistent estimator, it follows that

$$\begin{aligned} \mathbf{Z}'_{Q,N} (\mathbf{I}_{NT} - \widehat{\rho}_N \mathbf{M}'_N) \mathbf{H}_{Q,N} - \mathbf{Z}'_{Q,N} (\mathbf{I}_{NT} - \rho \mathbf{M}'_N) \mathbf{H}_{Q,N} &\xrightarrow{p} \mathbf{0}, \quad (\text{A.41}) \\ \mathbf{Z}'_N \boldsymbol{\Omega}_{u,N}^{-1/2} \left( \widehat{\boldsymbol{\vartheta}}_N \right) \mathbf{H}_{R,N} - \mathbf{Z}'_N \boldsymbol{\Omega}_{u,N}^{-1/2} (\boldsymbol{\vartheta}) \mathbf{H}_{R,N} &\xrightarrow{p} \mathbf{0}. \end{aligned}$$

The claim in the Theorem then follows from Assumptions 5 and 7.

Part (d):  $\widehat{H}_N - H_N \xrightarrow{p} 0$  is a direct consequence of (a), (b) and (c). Given Theorem 1, the (true) Hausman test statistics  $H_N$  is a asymptotically distributed as a quadratic form of normally distributed random variables and, hence, it has an asymptotic  $\chi^2$  distribution.

### Proof of Proposition 3:

We denote by  $\widehat{\boldsymbol{\theta}}_{GLS,N}^0$  and  $\widehat{\boldsymbol{\theta}}_{GLS,N}^1$  the algebraic expressions for the GLS estimators under the null and alternative hypothesis respectively (i.e. when  $\mathbf{u}_N$  is given by either Assumption 4, or 6). Analogously, we denote the algebraic expressions for the Hausman test statistics under the two hypotheses by  $H_N^0$  and  $H_N^1$ . We then have:

$$\begin{aligned} H_N^1 &= NT \left( \widehat{\boldsymbol{\theta}}_{GLS,N}^1 - \widehat{\boldsymbol{\theta}}_{W,N} \right)' (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \left( \widehat{\boldsymbol{\theta}}_{GLS,N}^1 - \widehat{\boldsymbol{\theta}}_{W,N} \right) \quad (\text{A.42}) \\ &= H_N^0 + NT \mathbf{c}'_N (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \left[ \mathbf{c}_N - 2 \left( \widehat{\boldsymbol{\theta}}_{GLS,N}^0 - \widehat{\boldsymbol{\theta}}_{W,N} \right) \right], \end{aligned}$$

where

$$\mathbf{c}_N = (\mathbf{0}_{K \times 1}, \mathbf{I}_K) \left[ \widetilde{\mathbf{Z}}'_N \mathbf{P}_{\mathbf{H}_{R,N}} \widetilde{\mathbf{Z}}_N \right]^{-1} \widetilde{\mathbf{Z}}'_N \mathbf{P}_{\mathbf{H}_{P,N}} (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \overline{\mathbf{X}}_N^* (\boldsymbol{\vartheta}) \boldsymbol{\pi}. \quad (\text{A.43})$$

Observe that Assumption 6 contains the random effects assumption for the independent component of the individual effects. Therefore, by Theorem 2,

we have  $H_N^0 \xrightarrow{d} \chi^2(K)$  and  $(NT)^{-1/2} \left( \widehat{\boldsymbol{\theta}}_{GLS,N}^0 - \widehat{\boldsymbol{\theta}}_{W,N} \right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})$ .

Thus also  $p \lim_{N \rightarrow \infty} \left( \widehat{\boldsymbol{\theta}}_{GLS,N}^0 - \widehat{\boldsymbol{\theta}}_{W,N} \right) = \mathbf{0}$ . Furthermore, since  $\mathbf{H}_{R,N} = [\mathbf{H}_{Q,N}, \mathbf{H}_{P,N}]$ , it follows from Assumption 5 that probability limit of  $\mathbf{c}_N$  exists and is finite, say  $p \lim_{N \rightarrow \infty} \mathbf{c}_N = \mathbf{c}$ . Therefore

$$\begin{aligned}
H_N^1 &= H_N^0 + NT \mathbf{c}'_N (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \mathbf{c}_N - 2NT \mathbf{c}'_N (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \left( \widehat{\boldsymbol{\theta}}_{GLS,N}^0 - \widehat{\boldsymbol{\theta}}_{W,N} \right) \\
&= H_N^0 + NT \mathbf{c}' (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \mathbf{c} - 2\sqrt{NT} \mathbf{c}'_N (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \sqrt{NT} \left( \widehat{\boldsymbol{\theta}}_{GLS,N}^0 - \widehat{\boldsymbol{\theta}}_{W,N} \right) \\
&= o_p \left( \sqrt{NT} \right) + NT \mathbf{c}' (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \mathbf{c} - 2\sqrt{NT} [\mathbf{c}' + o_p(1)] (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} o_p(1) \\
&= o_p \left( \sqrt{NT} \right) + NT \mathbf{c}' (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \mathbf{c} + \sqrt{NT} o_p(1) \\
&= NT \mathbf{c}' (\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS})^{-1} \mathbf{c} + o_p \left( \sqrt{NT} \right). \tag{A.44}
\end{aligned}$$

Therefore, for any  $\gamma > 0$ , we have

$$\lim_{N \rightarrow \infty} P(H_N^1 > \gamma) = 1.$$

## References

- [1] Anselin, L., 1988, *Spatial Econometrics: Methods and Models* (Kluwer Academic Publishers, Boston).
- [2] Baltagi, B.H., 2008, *Econometric Analysis of Panel Data*, Wiley and Sons, Chichester, 4th edition.
- [3] Baltagi, B.H. and Q. Li, 1992, A Note on the Estimation of Simultaneous Equations with Error Components, *Econometric Theory* 8, 113,119.
- [4] Case, A. C., 1991, Spatial Patterns in Household Demand, *Econometrica* 59, 953-965.
- [5] Cornwell, Ch., Schmidt, P. and D. Wyhowski, 1992, Simultaneous Equations and Panel Data, *Journal of Econometrics* 51, 151-181.
- [6] Egger, P., 2000, A Note on the Proper Econometric Specification of the Gravity Equation, *Economics Letters* 66, 25-31.
- [7] Greene, W., 2003, *Econometric Analysis*, 5th ed., Prentice Hall, New Jersey.
- [8] Hausman, J.A., 1978, Specification Tests in Econometrics, *Econometrica* 46, 1251-1271.
- [9] Horn, R.A. and Ch. R. Johnson, 1985, *Matrix Analysis*, Cambridge University Press, Cambridge UK.
- [10] Kapoor, M., H.H. Kelejian and I.R. Prucha, 2007, Panel Data Models with Spatially Correlated Error Components, *Journal of Econometrics* 140, 97-130.
- [11] 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.
- [12] 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.

- [13] Kelejian, H.H. and I.R. Prucha, 2007, Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances, forthcoming in the *Journal of Econometrics*.
- [14] Kelejian H. H., I.R. Prucha and Y. Yuzefovich, 2004, Instrumental Variable Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large and Small Sample Results, in J. LeSage and R.K. Pace, *Advances in Econometrics*, New York: Elsevier, 163-198.
- [15] Korniotis, G. M., 2008, Estimating Panel Models with Internal and External Habit Formation, forthcoming in the *Journal of Business and Economic Statistics*.
- [16] Lee, L., 2003, Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances, *Econometric Reviews* 22, 307-335.
- [17] Lee L. and Yu J.R., 2008, Estimation of spatial autoregressive panel data models with fixed effects, Working paper, Ohio State University.
- [18] Mundlak, Y., On Pooling Time Series and Cross Section Data, *Econometrica* 46, 69-85.
- [19] Mutl, Jan, 2006, *Dynamic Panel Data Models with Spatially Correlated Disturbances*, Dissertation, University of Maryland.
- [20] Schmidt, P., 1976, *Econometrics*, Marcel Dekker, New York.
- [21] Yu, J. R. de Jong and Lee, L., 2006, Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both  $n$  and  $T$  Are Large, Working paper, Ohio State University.
- [22] Yu, J. R. de Jong and Lee, L., 2007, Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both  $n$  and  $T$  Are Large: A Nonstationary Case, Working paper, Ohio State University.