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†

January 26, 2010

Abstract

This paper studies the random effects model and the fixed effects model for spatial panel data. 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, Universitaetsstrasse 15, 6020 Innsbruck; Austrian Institute of Economic Research, P.O.-Box 91, A-1103 Vienna, Austria; and 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.[1] 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

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

[1]See also the papers cited in Baltagi (2008) for more examples.

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 Yu (2010) 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, 2007 and 2008).

In the next section we specify our model and spell out the maintained assumptions. Section 3 defines the two estimators under consideration and derives their joint asymptotic distribution. 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:

$$y_{it,N} = \lambda \sum_{j=1}^{N} w_{ij,N} y_{jt,N} + \mathbf{x}_{it,N}\boldsymbol{\beta} + \mathbf{d}_{i,N}\boldsymbol{\gamma} + u_{it,N}. \tag{2.1}$$

Index $i = 1, .., N$ denotes the cross-sectional dimension of the panel while the index $t = 1, ..., T$ refers to the time series dimension of the panel. Throughout we assume that T is fixed, i.e. our asymptotic analysis refers to large N. $y_{it,N}$ 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. λ is the associated scalar parameter. $\mathbf{x}_{it,N}$ denotes a $1 \times (K-1)$ vector of time varying exogenous variables and $\boldsymbol{\beta}$ is the corresponding $(K-1) \times 1$ parameter vector. $\mathbf{d}_{i,N}$ is $1 \times L$ vector of time invariant variables, including the constant, with $L \times 1$ parameter vector $\boldsymbol{\gamma}$. Lastly, $u_{it,N}$ is the overall disturbance term.

We allow for cross-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):

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

where ρ 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:

$$\varepsilon_{it,N} = \mu_{i,N} + \nu_{it,N}. \tag{2.3}$$

$\nu_{it,N}$ 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

$$\begin{aligned} \mathbf{y}_{t,N} &= \lambda \mathbf{W}_N \mathbf{y}_{t,N} + \mathbf{X}_{t,N}\boldsymbol{\beta} + \mathbf{D}_N \boldsymbol{\gamma} + \mathbf{u}_{t,N}, \\ \mathbf{u}_{t,N} &= \rho \mathbf{M}_N \mathbf{u}_{t,N} + \boldsymbol{\varepsilon}_{t,N}, \\ \boldsymbol{\varepsilon}_{t,N} &= \boldsymbol{\mu}_N + \boldsymbol{\nu}_{t,N} \end{aligned} \tag{2.4}$$

where

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

Stacking over time periods, we write our model compactly as

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

where $\mathtt{W}_N = (\mathbf{I}_T \otimes \mathbf{W}_N)$, $\mathtt{M}_N = (\mathbf{I}_T \otimes \mathbf{M}_N)$, $\mathtt{D}_N = (\boldsymbol{\iota}_T \otimes \mathbf{D}_N)$, $\mathbf{Z}_N = (\mathtt{D}_N, \mathtt{W}_N \mathbf{y}_N, \mathbf{X}_N)$,

$\boldsymbol{\delta} = (\boldsymbol{\gamma}', \lambda, \boldsymbol{\beta}')'$, and

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

Throughout, we maintain the following basic assumptions, which follow closely those postulated in Kapoor, Kelejian and Prucha (2007). Note that we reserve the symbols $\sigma_\nu^2, \sigma_\mu^2$, λ and ρ for the true parameter values.

Assumption 1

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

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 matrices $(\mathbf{I}_N - \rho\mathbf{M}_N)$ and $(\mathbf{I}_N - \lambda\mathbf{W}_N)$ are non-singular.*

(c) *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., $sup_j \sum_{i=1}^N |a_{ij,N}| \leq k < \infty$, where k does not depend on N (but may depend on parameters of the* *model, i.e. on ρ or λ, respectively) and $a_{ij,N}$ denotes elements of the above matrices.*

Assumption 3

The exogenous variables collected in $\mathbf{X}_N$ and $\mathtt{D}_N$ are non-stochastic. 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 regularity conditions. Assumptions 2(b) and (c) 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. From Assumption 2(c) follows that $|\rho| \leq k_\rho < 1/\lambda_{\max}(\mathbf{M}_N)$, $|\lambda| \leq k_\lambda < 1/\lambda_{\max}(\mathbf{W}_N)$, where $\lambda_{\max}(.)$ denotes the largest absolute eigenvalue of a matrix and k_ρ and k_λ are constants.[2] Assumptions like 2(b) - 2(c) are typically maintained in spatial models (see Kelejian and Prucha, 1999) and restrict the extent of spatial dependence among cross-section units. They 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

[2] This follows from Corrolary 5.6.16 in Horn and Johnsonn (1985) using their Lemma 5.6.10.

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).

First, we analyze the spatial random effects estimator for this general spatial panel model.

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 (i) $E\left(\mu_{i,N}^2\right) = \sigma_\mu^2 > 0$. (ii) Furthermore, the elements of $\boldsymbol{\mu}_N$ are independent of the process for $\nu_{it,N}$ and $E\left(\mu_{i,N}\right) = 0$ are for all i.

The random effects assumptions maintains that the individual effects exhibit homoskedastic variances and that they are orthogonal to each of the explanatory variables. The latter will always hold if the explanatory variables are non-stochastic and $E[\boldsymbol{\mu}_N] = 0$ (see Wooldridge 2002, p. 257 and 259). The disturbances are generated as

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

Under Assumptions 1 and 4 (RE), the variance covariance matrix of the disturbances is given by

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

where $\boldsymbol{\Omega}_{u,N} = (\mathbf{I}_{NT} - \rho \mathtt{M}_N)^{-1} \left[\frac{\sigma_\mu^2}{\sigma_\nu^2} (\boldsymbol{\iota}_T \boldsymbol{\iota}_T' \otimes \mathbf{I}_N) + \mathbf{I}_{NT}\right] (\mathbf{I}_{NT} - \rho \mathtt{M}_N')^{-1}$. 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 $\mathbf{Q}_{i,N}$ $(i = 0, 1)$:

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

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)

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

where

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

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

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

where

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

If the parameter values ρ, σ_ν^2 and σ_μ^2 (and, therefore, σ_1^2) are known, the efficient GLS estimation procedure is to transform the model by the square root of $\boldsymbol{\Omega}_{u,N}$ given by

$$\boldsymbol{\Omega}_{u,N}^{-1/2} = \boldsymbol{\Omega}_{\varepsilon,N}^{-1/2} \left(\mathbf{I}_{NT} - \rho \mathtt{M}_N\right). \tag{3.8}$$

This is equivalent to first applying the spatial counterpart of the Cochrane-Orcutt transformation $(\mathbf{I}_{NT} - \rho \mathtt{M}_N)$ that eliminates the spatial correlation from the disturbances and then the familiar panel GLS transformation $\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2}$ 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 $\boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta})$ 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 $\mathtt{W}_N \mathbf{y}_N$ is endogenous in the (transformed) model, we adapt the instrumental variable procedure described in Kelejian and Prucha (1998).[3] 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 $[\mathbf{Q}_{0,N}\mathbf{X}_N, \mathbf{Q}_{1,N}\mathbf{X}_N, \mathbf{Q}_{1,N}\mathtt{D}_N]$. Observe that

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

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

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

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

$$\mathbf{H}_{R,N} = [\mathbf{H}_{Q,N}, \mathbf{H}_{P,N}] = [\mathbf{Q}_{0,N}\mathbf{G}_{0,N}, \mathbf{Q}_{1,N}\mathbf{G}_{1,N}],$$

[3] It is possible to use other sets of instruments that are similar to those proposed in Lee (2003) or Kelejian, Prucha and Yuzefovich (2004).

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

$$\left[\mathbf{X}_N, \mathtt{W}_N\mathbf{X}_N, \mathtt{W}_N^2\mathbf{X}_N..., \mathtt{M}_N\mathbf{X}_N, \mathtt{M}_N\mathtt{W}_N\mathbf{X}_N, \mathtt{M}_N\mathtt{W}_N^2\mathbf{X}_N, ...\right]$$

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

$$\left[\mathbf{G}_{0,N}, \mathtt{D}_N, \mathtt{W}_N\mathtt{D}_N, \mathtt{W}_N^2\mathtt{D}_N, ..., \mathtt{M}_N\mathtt{D}_N, \mathtt{M}_N\mathtt{W}_N\mathtt{D}_N, \mathtt{M}_N\mathtt{W}_N^2\mathtt{D}_N...\right].$$

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. We do no include the constant and the other time invariant variables in $\mathbf{G}_{0,N}$, since $\mathbf{Q}_{0,N}\mathtt{W}_N\mathtt{D}_N = \left[\left(\mathbf{I}_T - \frac{1}{T}\mathbf{J}_T\right)\iota_T \otimes \mathbf{W}_N\mathbf{D}_N\right] = \mathbf{0}$. In the special case where $\mathtt{W}_N = \mathtt{M}_N$ the set of instruments is based on $\mathbf{G}_{0,N} = [\mathbf{X}_N, \mathtt{W}_N\mathbf{X}_N, \mathtt{W}_N^2\mathbf{X}_N, ...]$ and $\mathbf{G}_{1,N} = [\mathbf{G}_{0,N}, \mathtt{D}_N, \mathtt{W}_N\mathtt{D}_N, ...]$. If the spatial weighting matrices are row normalized, the set of instruments in $\mathbf{G}_{1,N}$ excludes spatial lags of the time invariant variables, since in this case $\mathtt{M}_N\boldsymbol{\iota}_{NT} = \mathtt{W}_N\boldsymbol{\iota}_{NT} = \boldsymbol{\iota}_{NT}$, where $\boldsymbol{\iota}_{NT}$ is an $NT \times 1$ vector of ones. With regards to the choice of the instruments in practical applications, note that it is usually not advisable to use powers of the spatial weight matrices higher than two; see also the discussion of the choice of the instruments in Kelejian and Prucha (1998).

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 that are collected in $\mathbf{Z}_N = (\mathtt{D}_N, \mathtt{W}_N\mathbf{y}_N, \mathbf{X}_N)$:

Assumption 5

Let $\widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta}) = \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta})\,\mathbf{Z}_N$. *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) $\mathbf{M}_{H_R H_R} = \lim_{N\to\infty}(NT)^{-1}\,\mathbf{H}'_{R,N}\mathbf{H}_{R,N}$ *exists and is finite and non-singular,*

(b) $\mathbf{M}_{H_R \widetilde{Z}} = p\lim_{N\to\infty}(NT)^{-1}\,\mathbf{H}'_{R,N}\widetilde{\mathbf{Z}}_N$ *exists and is finite with full column rank.*[4]

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

$$\widehat{\boldsymbol{\delta}}_{GLS,N} = \left[\widehat{\widetilde{\mathbf{Z}}}_N(\boldsymbol{\vartheta})'\,\widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta})\right]^{-1}\widehat{\widetilde{\mathbf{Z}}}'_N(\boldsymbol{\vartheta})\,\widetilde{\mathbf{y}}_N(\boldsymbol{\vartheta}), \tag{3.11}$$

with $\widehat{\widetilde{\mathbf{Z}}}_N(\boldsymbol{\vartheta}) = \mathbf{P}_{H_{R,N}}\widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta})$, $\widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta}) = \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta})\,\mathbf{Z}_N$ and $\widetilde{\mathbf{y}}_N(\boldsymbol{\vartheta}) = \boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta})\,\mathbf{y}_N$. $\mathbf{P}_{H_{R,N}} = \mathbf{H}_{R,N}\left(\mathbf{H}'_{R,N}\mathbf{H}_{R,N}\right)^{-1}\mathbf{H}'_{R,N}$ 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

In situations where the random effects assumption might be violated one can use the spatial within estimator that remains unaffected by possible correlation of the unit specific effects and the explanatory variables and, hence a violation

[4] Observe that the $\widetilde{Z}_N$ depends on the parameter vector $\boldsymbol{\vartheta}$ and, hence, this assumption is meant to apply at the true parameter values.

of assumption 4 (RE). One can apply the within transformation $\mathbf{Q}_{0,N}$ to wipe out the individual effects (see e.g. Baltagi, 2008 and Mundlak, 1978). Using $\mathbf{Q}_{0,N}\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N\right)=\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N\right)\mathbf{Q}_{0,N}$, one obtains

$$
\begin{aligned}
\mathbf{Q}_{0,N}\mathbf{u}_N &= \left(\mathbf{E}_T\otimes\mathbf{I}_N\right)\left[\rho\left(\mathbf{I}_T\otimes\mathbf{M}_N\right)\mathbf{u}_N+\left(\boldsymbol{\iota}_T\otimes\mathbf{I}_N\right)\boldsymbol{\mu}_N+\boldsymbol{\nu}_N\right], \\
&= \rho\left(\mathbf{I}_T\otimes\mathbf{M}_N\right)\left(\mathbf{E}_T\otimes\mathbf{I}_N\right)\mathbf{u}_N+\left(\mathbf{E}_T\otimes\mathbf{I}_N\right)\boldsymbol{\nu}_N \\
&= \rho\left(\mathbf{I}_T\otimes\mathbf{M}_N\right)\mathbf{Q}_{0,N}\mathbf{u}_N+\mathbf{Q}_{0,N}\boldsymbol{\nu}_N. \qquad (3.12)
\end{aligned}
$$

or

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

where $\mathbf{E}_T=\left(\mathbf{I}_T-\frac{1}{T}\mathbf{J}_T\right)$. Hence, one can apply the Cochrane-Orcutt type transformation on the within transformed model to obtain the fixed effects generalized least squares (FEGLS) estimator. Note that the parameters of the time invariant variables, collected in the vector $\boldsymbol{\gamma}$, remain unidentified with this approach.

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 even if the unit specific effects depend on $\mathbf{X}_N$. 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 and the time invariant variables are wiped out in the spatial within estimator, we define the $(NT\times K)$ matrix $\mathbf{Z}_{Q,N}=\mathbf{Q}_{0,N}\left[\mathtt{W}_N\mathbf{y}_N,\mathbf{X}_N\right]$ 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 for $\boldsymbol{\theta}$, we maintain the following additional assumptions for the matrix of instruments used in the spatial within model:

Assumption 6

Let $\mathbf{Z}_{Q,N} = \mathbf{Q}_{0,N}[\mathtt{W}_N\mathbf{y}_N, \mathbf{X}_N]$. *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) $\mathbf{M}_{H_Q H_Q} = \lim_{N\to\infty}(NT)^{-1}\mathbf{H}'_{Q,N}\mathbf{H}_{Q,N}$ *is finite and non-singular with full column rank.*

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

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

$$\begin{aligned}\mathbf{Q}_{0,N}(\mathbf{I}_{NT}-\rho\mathtt{M}_N)\mathbf{y}_N &= (\mathbf{I}_{NT}-\rho\mathtt{M}_N)\mathbf{Q}_{0,N}\mathbf{y}_N \\ &= (\mathbf{I}_{NT}-\rho\mathtt{M}_N)\mathbf{Q}_{0,N}\mathbf{Z}_{Q,N}+\mathbf{Q}_{0,N}\boldsymbol{\nu}_N \\ \mathbf{y}^*_N(\rho) &= \mathbf{Z}^*_N(\rho)\boldsymbol{\theta}+\boldsymbol{\nu}^*_N,\end{aligned} \tag{3.13}$$

where $\mathbf{y}^*_N(\rho) = (\mathbf{I}_{NT}-\rho\mathtt{M}_N)\mathbf{Q}_{0,N}\mathbf{y}_N$, $\mathbf{Z}^*_N(\rho) = (\mathbf{I}_{NT}-\rho\mathtt{M}_N)\mathbf{Z}_{Q,N}$ and $\boldsymbol{\nu}^*_N = \mathbf{Q}_{0,N}\boldsymbol{\nu}_N$. The spatial within estimator for $\boldsymbol{\theta}$ is then obtained by applying IV to the transformed model to obtain

$$\widehat{\boldsymbol{\theta}}_{W,N} = \left[\widehat{\mathbf{Z}}^*_N(\rho)'\mathbf{Z}^*_N(\rho)\right]^{-1}\widehat{\mathbf{Z}}^*_N(\rho)'\mathbf{y}^*(\rho), \tag{3.14}$$

with $\widehat{\mathbf{Z}}^*_N(\rho) = \mathbf{P}_{H_{Q,N}}\mathbf{Z}^*_N(\rho) = \mathbf{P}_{H_{Q,N}}(\mathbf{I}_{NT}-\rho\mathtt{M}_N)\mathbf{Z}_N$ and
$\mathbf{P}_{H_{Q,N}} = \mathbf{H}_{Q,N}\left(\mathbf{H}'_{Q,N}\mathbf{H}_{Q,N}\right)^{-1}\mathbf{H}'_{Q,N}$ being 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 time invariant variables including the constant, we define $\widehat{\boldsymbol{\delta}}_{GLS,N} = \left(\widehat{\gamma}'_{GLS,N}, \widehat{\boldsymbol{\theta}}'_{GLS,N}\right)' = \left(\widehat{\gamma}'_{GLS,N}, \widehat{\lambda}'_{GLS,N}, \widehat{\boldsymbol{\beta}}'_{GLS,N}\right)'$.

Theorem 1 *Let Assumptions 1-6 hold. Then*

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

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

Proof: See the Appendix.

This Theorem forms the basis for the spatial Hausman test derived below under the null hypothesis that Assumption 4 (RE) is true.

4 Feasible Estimation

The spatial GLS and spatial within estimators defined above are based on the unknown parameters ρ, σ_ν^2 and σ_μ^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 (see Baltagi, 2008) and it can be written as

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

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* $\mathbf{M}_{H_QZ} = p\lim_{N\to\infty}(NT)^{-1}\mathbf{H}'_{Q,N}\mathbf{Z}_{Q,N}$ *exist and be finite with full column rank. Let Assumptions 1-3 and 6 hold. Then*

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

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

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 only consistent under the random effects Assumption 4 (RE). The spatial GM estimator for ρ, σ_ν^2 can then be based on the first three moment conditions given in Kapoor, Kelejian and Prucha (2007).

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

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

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

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

With the solution of the first three moment conditions at hand, one can solve the fourth moment condition to obtain an estimate of σ_1^2. Theorem 1 in Kapoor, Kelejian and Prucha (2007, p. 108) shows that the estimators for ρ, σ_ν^2 and σ_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.

Proposition 2 demonstrates that the parameters ρ, σ_ν^2 and σ_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 ρ, σ_ν^2 and σ_1^2.

Proposition 2 *Let the feasible estimators $\widehat{\boldsymbol{\theta}}_{FGLS,N}$ and $\widehat{\boldsymbol{\theta}}_{FW,N}$ be based on consistent estimators of ρ, σ_ν^2, and σ_1^2. Then under Assumptions 1-6 we have*

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

Proof: See the Appendix.

5 Hausman Specification Test

The spatial within estimator is consistent 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 $E[\boldsymbol{\mu}_N]=0$, implying that the spatial random effects model is inconsistent if the random effects Assumption 4 (RE) does not hold. The Hausman test (Hausman, 1978) suggests comparing these two estimators and to test whether the random effects assumption 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 the null hypothesis both considered estimators are consistent, while under the fixed effects assumption, the spatial random effects estimator is inconsistent, but the spatial within estimator is consistent. We summarize these properties in the following Lemma.[5]

Lemma 1 *Under Assumptions 1-6 we have:*

(a)

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

where $\boldsymbol{\Sigma}_{W}-\boldsymbol{\Sigma}_{GLS}$ *is positive definite at the true parameter values* $\boldsymbol{\vartheta}$*.*

(b) Furthermore

$$\widehat{\boldsymbol{\Sigma}}_{W,N}-\widehat{\boldsymbol{\Sigma}}_{GLS,N}\xrightarrow{p.}\boldsymbol{\Sigma}_{W}-\boldsymbol{\Sigma}_{GLS}$$

where

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

and

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

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

Proof: See the Appendix.

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

[5] To operationalize this lemma, we need to provide a consistent estimator $\widehat{\boldsymbol{\vartheta}}_{N}$ of $\boldsymbol{\vartheta}$. Our suggestion (see the summary of our estimation procedure below) is to construct both $\widehat{\boldsymbol{\Sigma}}_{W,N}$ and $\widehat{\boldsymbol{\Sigma}}_{GLS,N}$ using $\widehat{\boldsymbol{\vartheta}}_{N}=\left(\widehat{\rho}_{N},\widehat{\sigma}^{2}_{\nu,N},\widehat{\sigma}^{2}_{1,N}\right)'$, where $\widehat{\rho}_{N}$ and $\widehat{\sigma}^{2}_{\nu,N}$ are estimated from three moment conditions using the within estimates (see 4.2), while $\widehat{\sigma}^{2}_{1,N}$ is derived from a fourth moment condition that exploits the between variation.

Theorem 2 *Assume that Assumptions 1- 6 hold and that*

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

and

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

Then $\widehat{H}_N - H_N \xrightarrow{p.} 0$, *where* H_N *is asymptotically* χ^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:

1. Calculate a consistent initial instrumental variables within estimator $\widehat{\boldsymbol{\theta}}_{I,N}$ which wipes out the individual effects using the within transformation. This estimator ignores the spatial correlation in the disturbances.

2. Use the resulting estimated disturbances of the within transformed model in a spatial GM procedure as described in Kapoor, Kelejian and Prucha (2007) and obtain a (consistent) estimator $\widehat{\boldsymbol{\vartheta}}_N = \left(\widehat{\rho}_N, \widehat{\sigma}^2_{\nu,N}, \widehat{\sigma}^2_{1,N}\right)'$.

3. Transform the model by the spatial Cochrane-Orcutt transformation and then use either the GLS, or the within transformation to obtain the feasible spatial GLS and spatial within estimators $\widehat{\boldsymbol{\theta}}_{FGLS,N}$ and $\widehat{\boldsymbol{\theta}}_{FW,N}$ respectively.

4. Calculate the Hausman test statistics $\widehat{H}_N$ to make a decision whether the random or fixed effects specification is more appropriate. Note, once the variables are spatially Cochrane-Orcutt transformed standard econometric software can calculate the spatial Hausman test.

We now investigate the power properties of the test statistics, i.e. its behavior when the null hypothesis is violated. In order to do so, we need to specify a particular alternative. Note that all results obtained up to this point were independent of a particular choice of the alternative hypothesis. We now follow Mundlak (1978) and assume an alternative (fixed effects) model in which the explanatory variables and the unit specific effects are related in the following way:[6]

Assumption 4 (FE)

The vector of individual effects is given by

$$\boldsymbol{\mu}_N = \overline{\mathbf{X}}_N \boldsymbol{\pi} + \boldsymbol{\xi}_N,$$

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

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

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

Note that it would be trivial to generalize our assumption to an alternative model that has been defined by Chamberlain (1982), where $\mu_{i,N} = \sum_{t=1}^{T} \mathbf{x}_{it,N} \boldsymbol{\pi}_t + \xi_{i,N}$

[6] There are other specifications of the unit effects possible. In the present case, unit effects are decomposed in a spatially correlated errer term simialar to that defined in Assumption 4 (RE) and a systematic component without spatial effects.

or $\boldsymbol{\mu}_N = \sum_{t=1}^{T} \mathbf{X}_{t,N} \boldsymbol{\pi}_t + \boldsymbol{\xi}_N$. Clearly, setting $\boldsymbol{\pi}_t = \frac{1}{T} \boldsymbol{\pi}$ yields Mundlak's formulation.

Under the FE assumption, we have

$$\begin{aligned} \boldsymbol{\varepsilon}_N &= (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \left[\left(\tfrac{1}{T} \boldsymbol{\iota}_T' \otimes \mathbf{I}_N \right) \mathbf{X}_N + \boldsymbol{\xi}_N \right] + \boldsymbol{\nu}_N \qquad (5.1) \\ &= \mathbf{Q}_{1,N} \mathbf{X}_N \boldsymbol{\pi} + (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \boldsymbol{\xi}_N + \boldsymbol{\nu}_N, \\ \mathbf{u}_N &= (\mathbf{I}_{NT} - \rho \mathtt{M}_N)^{-1} \left[\mathbf{Q}_{1,N} \mathbf{X}_N \boldsymbol{\pi} + (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \boldsymbol{\xi}_N + \boldsymbol{\nu}_N \right]. \end{aligned}$$

The individual effects do not depend on the explanatory variables if and only if $\boldsymbol{\pi} = \mathbf{0}$ and the random effects model defined under Assumption 4 (RE) arises as a special case of Assumption 4 (FE). Observe that under Assumption 4 (FE) with $\boldsymbol{\pi} \neq \mathbf{0}$ the spatial GLS estimator is biased as one obtains

$$\begin{aligned} \widehat{\boldsymbol{\delta}}_{GLS,N} &= \left[\widehat{\widetilde{\mathbf{Z}}}_N (\boldsymbol{\vartheta})' \widetilde{\mathbf{Z}}_N (\boldsymbol{\vartheta}) \right]^{-1} \widehat{\widetilde{\mathbf{Z}}}_N' (\boldsymbol{\vartheta}) \widetilde{\mathbf{y}}_N (\boldsymbol{\vartheta}) \qquad (5.2) \\ &= \boldsymbol{\delta} + \left[\widehat{\widetilde{\mathbf{Z}}}_N (\boldsymbol{\vartheta})' \widetilde{\mathbf{Z}}_N (\boldsymbol{\vartheta}) \right]^{-1} \widehat{\widetilde{\mathbf{Z}}}_N' (\boldsymbol{\vartheta}) \boldsymbol{\Omega}_{u,N}^{-1/2} \\ &\quad \cdot \left[\mathbf{Q}_{1,N} \mathbf{X}_N \boldsymbol{\pi} + (\boldsymbol{\iota}_T \otimes \mathbf{I}_N) \boldsymbol{\xi}_N + \boldsymbol{\nu}_N \right]. \end{aligned}$$

On the other hand, the within transformation wipes out the individual effects and, hence, the within estimator 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 $N \to \infty$ for an arbitrary significance level of the test.

Proposition 3 *Let Assumptions 1-3 and 4(FE), 5 and 6 hold and assume that*

$$\mathbf{M}_{H_R \widetilde{X}} = lim_{N \to \infty} (NT)^{-1} \mathbf{H}_{R,N} \boldsymbol{\Omega}_{u,N}^{-1/2} \mathbf{Q}_{1,N} \mathbf{X}_N$$

exists and is finite with full column rank. Let $h > 0$ be some positive constant. Then $\lim_{N\to\infty} P(H_N > h) = 1$.

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:

$$y_{it,N} = \lambda \sum_{j=1}^{N} w_{ij,N} y_{jt,N} + \beta x_{it,N} + \alpha + u_{it,N}, \quad i = 1, ..., N \text{ and } t = 1, ..., T. \quad (6.1)$$

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 $[a, b]$. In accordance with Assumption 3, 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 $\overline{x}_i$, setting $\mu_i = \mu_{i0} + \pi\overline{x}_i$, where μ_{i0} is drawn from a normal distribution, i.e. $\mu_{i0} \sim i.i.d.\ N(0, 10\phi)$ and π 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 μ_i so that its mean is 0 and its variance 10ϕ, 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 λ and ρ vary over the set $\{-0.8, -0.4, 0, 0.4, 0.8\}$. The parameter $\boldsymbol{\pi}$ takes its values in $\{-0.2, -0.1, 0, 0.1, 0.2\}$. Based on the discussion above we use the instruments $\mathbf{H}_{Q,N} = [\mathbf{Q}_{0,N}\mathbf{x}_N, \mathbf{Q}_{0,N}\mathtt{W}_N\mathbf{x}_N,$ $\mathbf{Q}_{0,N}\mathtt{W}_N^2\mathbf{x}_N]$, while $\mathbf{H}_{R,N}$ is composed of $[\mathbf{H}_{Q,N}, \mathbf{Q}_{1,N}\mathbf{x}_N, \mathbf{Q}_{1,N}\mathtt{W}_N\mathbf{x}_N, \mathbf{Q}_{1,N}\mathtt{W}_N^2\mathbf{x}_N,$ $\iota_{NT}, \mathbf{Q}_{1,N}\mathtt{W}_N\iota_{NT}]$.

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 the Table 1, 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 λ, where the test is slightly oversized. For example, at $\lambda = 0.8$ and $\rho = 0.8$ the size of the spatial Hausman test is 0.092. At negative values of ρ, this phenomenon is not observed. The power of the test by and large remains unaffected by variations of ρ and λ, although it seems somewhat lower at high absolute values of ρ or high absolute values of λ.

A larger cross-section ($N = 324$, $NT = 1620$) improves both the size and the power of the test as expected (see Table 2), especially the size distortion at high positive values of λ is now reduced. In Table 3 we extend the time series dimension and set $T = 11$ ($NT = 1584$). The size distortion at high values of λ 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 $N = 144$, $T = 5$ and $\phi = 0.8$, so that $\sigma_\mu^2 = 8$ and σ_ν^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 and the size distortion observed in the baseline scenario now vanishes. Also, the power of the test is significantly higher.

We performed a series of robustness checks and the full set of tables with these results is available upon request. In particular we considered the case where serial correlation in the disturbances is present but is ignored. Our results show that the size of the Hausman test remains nearly unaffected. However, the spatial Hausman test has lower power especially if the serial correlation is high.

In addition, we investigated the performance of the proposed Hausman test when the exogenous variable has weak effects on the dependent variable. In our simulation experiment we reduced the variance of the generated explanatory variable by one half, which results in weaker instruments as compared to the base line experiment. The results show that a high positive values of both λ and ρ the spatial Hausman test tends to over-reject, if the explanatory variables are weak. Also, the power of the test is lower as compared to the base line experiment. The tables corresponding to these additional experiments are available upon request from the authors.

We investigated the root mean square error (RMSE) and the bias of the estimators of β, λ and ρ for the basic case with $N = 144$, $T = 5$ and $\theta = 0.5$.[7] 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 interquantile 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 β 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 λ and at high absolute values of ρ. A similar pattern can be found for the RMSE of λ, although the efficiency loss of the spatial within estimator is much higher as compared to that for β. Under H_1 the random effects estimator is inconsistent, leading to large biases in both β and λ. The bias of the slope parameter β is hardly affected by different degrees of spatial dependence as represented by the parameters values of λ. However, the bias is negative at low and negative values of ρ and turns to the positive if ρ gets high. With respect to the estimates of λ, we find that the bias is negative if λ or ρ take on negative values, but that it declines in λ and/or ρ. At $\lambda = 0.8$ or $\rho = 0.8$ the bias nearly vanishes. The estimates of ρ remain unaffected by deviations from H_0 as expected. These estimates are based on the spatial within estimator which is

[7] The tables corresponding to these simulation experiments are available upon request from the authors.

consistent under both H_0 and H_1.

We also assess the performance of the spatial Hausman test for non-normal disturbances against the baseline case in Table 1.[8] 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 explanatorily 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

[8] The corresponding tables are available from the authors upon request.

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.

A Appendix

Proof of Theorem 1:

We denote the $(T+1)\,N\times 1$ vector of i.i.d.$(0,1)$ innovations as $\boldsymbol{\zeta}_N = \left(\frac{\boldsymbol{\mu}_N'}{\sigma_\mu}, \frac{\boldsymbol{\nu}_N'}{\sigma_\nu}\right)'$ and we write the stacked estimators as[9]

$$\begin{pmatrix} \widehat{\boldsymbol{\delta}}_{GLS,N} - \boldsymbol{\delta} \\ \widehat{\boldsymbol{\theta}}_{W,N} - \boldsymbol{\theta} \end{pmatrix} = \widetilde{\mathbf{P}}_N \mathbf{F}_N \boldsymbol{\zeta}_N, \tag{A.1}$$

where

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

with

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

[9] 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 $\boldsymbol{\Omega}_{\varepsilon,N}^{-1/2} = \sigma_1^{-1}\mathbf{Q}_{1,N} + \sigma_\nu^{-1}\mathbf{Q}_{0,N}$.

and

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

By Assumptions (5) and (6) 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.

$$\begin{aligned}
(NT)\,\widetilde{\mathbf{P}}_{R,N} &\xrightarrow{p.} \left[\mathbf{M}'_{H_R\widetilde{Z}}\mathbf{M}^{-1}_{H_RH_R}\mathbf{M}_{H_R\widetilde{Z}}\right]^{-1}\mathbf{M}'_{H_R\widetilde{Z}}\mathbf{M}^{-1}_{H_RH_R}, \qquad (A.4)\\
(NT)\,\widetilde{\mathbf{P}}_{Q,N} &\xrightarrow{p.} \left[\mathbf{M}'_{H_QZ^*}\mathbf{M}^{-1}_{H_QH_Q}\mathbf{M}_{H_QZ^*}\right]^{-1}\mathbf{M}'_{H_QZ^*}\mathbf{M}^{-1}_{H_QH_Q}.
\end{aligned}$$

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 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[10]

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

$$= \sigma_\nu^2 \begin{pmatrix} \mathbf{H}'_{R,N} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}'_{Q,N} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{NT} & \mathbf{I}_{NT} \\ \mathbf{I}_{NT} & \mathbf{I}_{NT} \end{pmatrix} \begin{pmatrix} \mathbf{H}_{R,N} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{Q,N} \end{pmatrix}$$

and hence

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

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 (5) and (6) 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

[10] 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

$$\mathbf{H}'_{R,N}\mathbf{Q}_{0,N}\mathbf{H}_{Q,N} = \mathbf{H}'_{R,N}\mathbf{H}_{Q,N} \tag{A.5}$$

$$= \begin{pmatrix} \mathbf{G}'_{1,N}\mathbf{Q}_{1,N} \\ \mathbf{G}'_{0,N}\mathbf{Q}_{0,N} \end{pmatrix}\mathbf{Q}_{0,N}\mathbf{H}_{Q,N} = \begin{pmatrix} \mathbf{0}_{p\times q} \\ \mathbf{H}'_{Q,N}\mathbf{H}_{Q,N} \end{pmatrix}.$$

the central limit theorem it follows that[11]

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

From Assumptions (5) and (6) we also have

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

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

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta'_{12} & \Delta_{22} \end{pmatrix} \tag{A.10}$$

with

$$\begin{aligned} \Delta_{11} &= \sigma_\nu^2\left(\mathbf{M}'_{H_R\widetilde{Z}}\mathbf{M}^{-1}_{H_RH_R}\mathbf{M}_{H_R\widetilde{Z}}\right)^{-1} = \begin{pmatrix} \boldsymbol{\Sigma}_{\gamma,GLS} & \boldsymbol{\Sigma}_{\gamma\theta,GLS} \\ \boldsymbol{\Sigma}'_{\gamma\theta,GLS} & \boldsymbol{\Sigma}_{GLS} \end{pmatrix} \qquad \text{(A.11)} \\ \Delta_{12} &= \sigma_\nu^2\left(\mathbf{M}'_{H_R\widetilde{Z}}\mathbf{M}^{-1}_{H_RH_R}\mathbf{M}_{H_R\widetilde{Z}}\right)^{-1}\begin{pmatrix} \mathbf{0}_{L\times K} \\ \mathbf{I}_K \end{pmatrix}_{(K+L)\times K} = \begin{pmatrix} \boldsymbol{\Sigma}_{\gamma\theta,GLS} \\ \boldsymbol{\Sigma}_{GLS} \end{pmatrix} \\ \Delta_{22} &= \sigma_\nu^2\left(\mathbf{M}'_{H_QZ^*}\mathbf{M}^{-1}_{H_QH_Q}\mathbf{M}_{H_QZ^*}\right)^{-1} = \boldsymbol{\Sigma}_W. \end{aligned}$$

We have ordered the elements of the vector of parameters $\boldsymbol{\delta}$ such that the first elements correspond to the time invariant variables so that the asymptotic variance-

[11] Note that it can be demonstrated that $\lim_{N\to\infty}(NT)^{-1}\,\mathbf{F}_N\mathbf{F}'_N$ exists.

covariance matrix of the stacked estimators becomes

$$\mathbf{\Delta} = \begin{bmatrix} \mathbf{\Sigma}_{\gamma,GLS} & \mathbf{\Sigma}_{\gamma\theta,GLS} & \mathbf{\Sigma}_{\gamma\theta,GLS} \\ \mathbf{\Sigma}'_{\gamma\theta,GLS} & \mathbf{\Sigma}_{GLS} & \mathbf{\Sigma}_{GLS} \\ \mathbf{\Sigma}'_{\gamma\theta,GLS} & \mathbf{\Sigma}_{GLS} & \mathbf{\Sigma}_{W} \end{bmatrix}, \tag{A.12}$$

and hence

$$\sqrt{NT}\begin{pmatrix} \widehat{\boldsymbol{\gamma}}_{GLS,N} - \boldsymbol{\gamma} \\ \widehat{\boldsymbol{\theta}}_{GLS,N} - \boldsymbol{\theta} \\ \widehat{\boldsymbol{\theta}}_{W,N} - \boldsymbol{\theta} \end{pmatrix} \xrightarrow{d.} N\left(\mathbf{0}, \mathbf{\Delta}\right). \tag{A.13}$$

Proof of Proposition 1:

Note that we have

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

Given Assumption 6 and the condition in the proposition, it follows that $(NT)\,\widetilde{\mathbf{P}}_{Q,N} \xrightarrow{p.} \left[\mathbf{M}'_{H_QZ}\mathbf{M}^{-1}_{H_QH_Q}\mathbf{M}_{H_QZ}\right]^{-1}\mathbf{M}'_{H_QZ}\mathbf{M}^{-1}_{H_QH_Q}$ and

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

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 6(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)

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

and

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

where $\widehat{\boldsymbol{\vartheta}}$ and $\widehat{\rho}$ are (any) consistent estimators of $\boldsymbol{\vartheta}$ and ρ. Note that $\widehat{\widetilde{\mathbf{Z}}}_N\left(\widehat{\boldsymbol{\vartheta}}\right)=\mathbf{P}_{H_{R,N}}\widetilde{\mathbf{Z}}_N\left(\widehat{\boldsymbol{\vartheta}}\right)$ and $\widehat{\mathbf{Z}}_N^*\left(\widehat{\boldsymbol{\vartheta}}\right)=\mathbf{P}_{H_{Q,N}}\mathbf{Z}_N^*\left(\widehat{\boldsymbol{\vartheta}}\right)$, where

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

and

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

Hence

$$\begin{aligned}
&\Delta_{G1,N} &\text{(A.20)}\\
=\ & \tfrac{\rho-\widehat{\rho}}{NT}\mathbf{Z}_N'\mathtt{M}_N'\left(\mathbf{Q}_{1,N}+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Q}_{0,N}\right)\mathbf{P}_{H_{R,N}}\left(\mathbf{Q}_{1,N}+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Q}_{0,N}\right)\mathbf{Z}_N\\
&+\tfrac{\rho-\widehat{\rho}}{NT}\mathbf{Z}_N'\left(\mathbf{Q}_{1,N}+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Q}_{0,N}\right)\mathbf{P}_{H_{R,N}}\left(\mathbf{Q}_{1,N}+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Q}_{0,N}\right)\mathtt{M}_N\mathbf{Z}_N\\
&+\tfrac{\widehat{\rho}^2-\rho^2}{NT}\mathbf{Z}_N'\mathtt{M}_N'\left(\mathbf{Q}_{1,N}+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Q}_{0,N}\right)\mathbf{P}_{H_{R,N}}\left(\mathbf{Q}_{1,N}+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Q}_{0,N}\right)\mathtt{M}_N\mathbf{Z}_N\\
&+\tfrac{\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}-\frac{\sigma_\nu}{\sigma_1}}{NT}\mathbf{Z}_N'\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N'\right)\mathbf{Q}_{0,N}\mathbf{P}_{H_{R,N}}\mathbf{Q}_{1,N}\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N\right)\mathbf{Z}_N\\
&+\tfrac{\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}-\frac{\sigma_\nu}{\sigma_1}}{NT}\mathbf{Z}_N'\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N'\right)\mathbf{Q}_{1,N}\mathbf{P}_{H_{R,N}}\mathbf{Q}_{0,N}\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N\right)\mathbf{Z}_N\\
&+\tfrac{\frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2}-\frac{\sigma_\nu^2}{\sigma_1^2}}{NT}\mathbf{Z}_N'\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N'\right)\mathbf{Q}_{0,N}\mathbf{P}_{H_{R,N}}\mathbf{Q}_{0,N}\left(\mathbf{I}_{NT}-\rho\mathtt{M}_N\right)\mathbf{Z}_N=
\end{aligned}$$

$$\begin{aligned}
&\tfrac{\rho-\widehat{\rho}}{NT}\left(\mathbf{Z}_N'\mathbf{A}_{G1,N}\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Z}_N'\mathbf{A}_{G2,N}\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Z}_N'\mathbf{A}_{G3,N}\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2}\mathbf{Z}_N'\mathbf{A}_{G4,N}\mathbf{Z}_N\right)\\
&+\tfrac{\rho-\widehat{\rho}}{NT}\left(\mathbf{Z}_N'\mathbf{A}_{G1,N}'\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Z}_N'\mathbf{A}_{G2,N}'\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Z}_N'\mathbf{A}_{G3,N}'\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2}\mathbf{Z}_N'\mathbf{A}_{G4,N}'\mathbf{Z}_N\right)\\
&+\tfrac{\widehat{\rho}^2-\rho^2}{NT}\left(\mathbf{Z}_N'\mathbf{A}_{G5,N}\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Z}_N'\mathbf{A}_{G6,N}\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}\mathbf{Z}_N'\mathbf{A}_{G7,N}\mathbf{Z}_N+\frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2}\mathbf{Z}_N'\mathbf{A}_{G8,N}\mathbf{Z}_N\right)\\
&+\tfrac{\frac{\widehat{\sigma}_\nu}{\widehat{\sigma}_1}-\frac{\sigma_\nu}{\sigma_1}}{NT}\left(\mathbf{Z}_N'\mathbf{A}_{G9,N}\mathbf{Z}_N+\mathbf{Z}_N'\mathbf{A}_{G9,N}'\mathbf{Z}_N\right)\\
&+\tfrac{\frac{\widehat{\sigma}_\nu^2}{\widehat{\sigma}_1^2}-\frac{\sigma_\nu^2}{\sigma_1^2}}{NT}\mathbf{Z}_N'\mathbf{A}_{G10,N}\mathbf{Z}_N,
\end{aligned}$$

where

$$\mathbf{A}_{G1,N} = \mathtt{M}'_N \mathbf{Q}_{1,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N}, \tag{A.21}$$
$$\mathbf{A}_{G2,N} = \mathtt{M}'_N \mathbf{Q}_{1,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N},$$
$$\mathbf{A}_{G3,N} = \mathtt{M}'_N \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N},$$
$$\mathbf{A}_{G4,N} = \mathtt{M}'_N \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N},$$

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

$$\mathbf{A}_{G9,N} = \left(\mathbf{I}_{NT} - \rho \mathtt{M}'_N\right) \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{1,N} \left(\mathbf{I}_{NT} - \rho \mathtt{M}_N\right),$$
$$\mathbf{A}_{G10,N} = \left(\mathbf{I}_{NT} - \rho \mathtt{M}'_N\right) \mathbf{Q}_{0,N} \mathbf{P}_{H_{R,N}} \mathbf{Q}_{0,N} \left(\mathbf{I}_{NT} - \rho \mathtt{M}_N\right)$$

Analogically,

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

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

where

$$\mathbf{A}_{W1,N} = \mathtt{M}_N' \mathbf{P}_{H_{Q,N}}, \qquad \text{(A.23)}$$
$$\mathbf{A}_{W2,N} = \mathtt{M}_N' \mathbf{P}_{H_{Q,N}} \mathtt{M}_N.$$

In light of Assumptions 2, 5 and 6 the row and column sums of the matrices $\mathbf{A}_{Gi,N}$ and $\mathbf{A}_{Wj,N}$ are uniformly bounded in absolute value. By Lemma C3 in Mutl (2006) we then have that $(NT)^{-1} \mathbf{X}_N' \mathbf{A}_{Gi,N} \mathbf{X}_N$, $(NT)^{-1} \mathbf{X}_N' \mathbf{A}_{Wi,N} \mathbf{X}_N$, $(NT)^{-1} \mathtt{D}_N' \mathbf{A}_{Gi,N} \mathtt{D}_N$ and $(NT)^{-1} \mathtt{D}_N' \mathbf{A}_{Wi,N} \mathtt{D}_N$ 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 $(NT)^{-1} \mathbf{u}_N' \mathbf{A}_{Gi,N} \mathbf{u}_N$ is uniformly bounded in absolute value. Recall that $\mathbf{Z}_N = (\mathtt{D}_N, \mathtt{W}_N \mathbf{y}_N, \mathbf{X}_N)$, where the solution of the model yields $\mathbf{y}_N = (\mathbf{I}_{NT} - \lambda \mathtt{W}_N)^{-1} (\mathbf{X}_N \boldsymbol{\beta} + \mathtt{D}_N \boldsymbol{\gamma} + \mathbf{u}_N)$. We thus have that the elements of

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

and

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

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 $N \to \infty$.

Next we use similar derivations to obtain

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

and

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

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

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

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

and

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

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

which by Assumptions 2, 5 and 6 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:

$$(NT)^{-1}\boldsymbol{\beta}'\mathbf{X}_N'\mathbf{A}_{Gi,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Gi,N}'\mathbf{X}_N\boldsymbol{\beta}, \tag{A.29}$$

$$(NT)^{-1}\boldsymbol{\gamma}'\mathtt{D}_N'\mathbf{A}_{Gi,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Gi,N}'\mathtt{D}_N\boldsymbol{\gamma},$$

$$(NT)^{-1}\boldsymbol{\beta}'\mathbf{X}_N'\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N'\right)^{-1}\mathtt{W}_N'\mathbf{A}_{Gi,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Gi,N}'\mathtt{W}_N\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N\right)^{-1}\mathbf{X}_N\boldsymbol{\beta},$$

$$(NT)^{-1}\boldsymbol{\gamma}'\mathtt{D}_N'\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N'\right)^{-1}\mathtt{W}_N'\mathbf{A}_{Gi,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Gi,N}'\mathtt{W}_N\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N\right)^{-1}\mathtt{D}_N\boldsymbol{\gamma},$$

and

$$(NT)^{-1}\boldsymbol{\beta}'\mathbf{X}_N'\mathbf{A}_{Wj,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Wj,N}'\mathbf{X}_N\boldsymbol{\beta},$$

$$(NT)^{-1}\boldsymbol{\gamma}'\mathtt{D}_N'\mathbf{A}_{Wj,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Wj,N}'\mathtt{D}_N\boldsymbol{\gamma},$$

$$(NT)^{-1}\boldsymbol{\beta}'\mathbf{X}_N'\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N'\right)^{-1}\mathtt{W}_N'\mathbf{A}_{Wj,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Wj,N}'\mathtt{W}_N\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N\right)^{-1}\mathbf{X}_N\boldsymbol{\beta},$$

$$(NT)^{-1}\boldsymbol{\gamma}'\mathtt{D}_N'\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N'\right)^{-1}\mathtt{W}_N'\mathbf{A}_{Wj,N}\boldsymbol{\Omega}_{u,N}\mathbf{A}_{Wj,N}'\mathtt{W}_N\left(\mathbf{I}_{NT}-\lambda\mathtt{W}_N\right)^{-1}\mathtt{D}_N\boldsymbol{\gamma}.$$

Finally, the two remaining scalar variances are given by

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

and

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

where the $\mathbf{B}$ matrices are products of $\mathbf{A}$, $\mathtt{W}_N$, $\mathtt{W}_N'$, $\left(\mathbf{I}_{NT} - \lambda \mathtt{W}_N\right)^{-1}$, and $\left(\mathbf{I}_{NT} - \lambda \mathtt{W}_N'\right)^{-1}$ matrices and hence by Assumptions 2, 5 and 6 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, $(NT)^{-1/2} \mathbf{Z}_N' \mathbf{A}_{Gj,N} \mathbf{u}_N = O_P(1)$, and $(NT)^{-1/2} \mathbf{Z}_N' \mathbf{A}_{Wj,N} \mathbf{u}_N = O_P(1)$. Thus $\Delta_{G2,N} \xrightarrow{p.} 0$ and $\Delta_{W2,N} \xrightarrow{p.} 0$ as $N \to \infty$, since $\widehat{\rho}$, $\widehat{\sigma}_\nu$, and $\widehat{\sigma}_1$ are consistent estimators.

Proof of Lemma 1

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

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

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

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

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 $\mathbf{S}_{R,N} = \widetilde{\mathbf{Z}}'_N \mathbf{P}_{\mathbf{H}_{R,N}} \widetilde{\mathbf{Z}}_N$, 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

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

with $\mathbf{A}_N = \mathbf{I}_{NT} - \rho \mathtt{M}_N$ 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 $\mathbf{S}^{22}_{R,N}$. Using the formula for partitioned inverses (see e.g. section 0.7.3 in Horn and Johnson, 1985) yields after some manipulation

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

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

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

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

so that

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

where

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

$$\mathbf{D}_N = \phi^2 (\mathtt{W}_N\mathbf{y}_N, \mathbf{X}_N)' \mathbf{A}_N' \mathbf{P}_{\mathbf{H}_{P,N}} \left[\mathbf{I}_{NT} - \mathbf{A}_N\mathtt{D}_N \left(\mathtt{D}_N'\mathbf{A}_N'\mathbf{A}_N\mathtt{D}_N\right)^{-1} \mathtt{D}_N'\mathbf{A}_N'\right] \cdot \mathbf{P}_{\mathbf{H}_{P,N}} \mathbf{A}_N (\mathtt{W}_N\mathbf{y}_N, \mathbf{X}_N) \tag{A.38}$$

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

$$(\mathbf{C}_N + \mathbf{D}_N)^{-1} = \mathbf{C}_N^{-1} - \mathbf{C}_N^{-1}\left(\mathbf{D}_N^{-1} + \mathbf{C}_N^{-1}\right)^{-1}\mathbf{C}_N^{-1} \tag{A.39}$$

and hence

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

which is clearly positive definite.

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

$$\mathbf{Z}_{Q,N}'\left(\mathbf{I}_{NT} - \widehat{\rho}_N\mathtt{M}_N'\right)\mathbf{H}_{Q,N} - \mathbf{Z}_{Q,N}'\left(\mathbf{I}_{NT} - \rho\mathtt{M}_N'\right)\mathbf{H}_{Q,N} \xrightarrow{p.} \mathbf{0}, \tag{A.41}$$

$$\mathbf{Z}_N'\boldsymbol{\Omega}_{u,N}^{-1/2}\left(\widehat{\boldsymbol{\vartheta}}_N\right)\mathbf{H}_{R,N} - \mathbf{Z}_N'\boldsymbol{\Omega}_{u,N}^{-1/2}(\boldsymbol{\vartheta})\mathbf{H}_{R,N} \xrightarrow{p.} \mathbf{0}.$$

The claim in the Lemma then follows from Assumptions 5 and 6.

Proof of Theorem 2

$\widehat{H}_N - H_N \xrightarrow{p.} 0$ follows directly from Lemma 1 and Proposition 2. 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 χ^2 distribution.

Proof of Proposition 3:

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

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}^1_{GLS,N} - \widehat{\boldsymbol{\theta}}_{W,N} &= \left(\mathbf{0}_{K\times L}, \mathbf{I}_K\right)\left[\widehat{\widetilde{\mathbf{Z}}}_N(\boldsymbol{\vartheta})' \widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta})\right]^{-1} \\
&\quad \widehat{\widetilde{\mathbf{Z}}}'_N(\boldsymbol{\vartheta})\, \boldsymbol{\Omega}^{-1/2}_{u,N}\left[\mathbf{Q}_{1,N}\mathbf{X}_N\boldsymbol{\pi} + \left(\boldsymbol{\iota}_T \otimes \mathbf{I}_N\right)\boldsymbol{\xi}_N + \boldsymbol{\nu}_N\right] - \widehat{\boldsymbol{\theta}}_{W,N} \\
&= \widehat{\boldsymbol{\theta}}^0_{GLS,N} - \widehat{\boldsymbol{\theta}}_{W,N} + \mathbf{C}_N\boldsymbol{\pi},
\end{aligned} \tag{A.42}$$

where the nonstochastic matrix $\mathbf{C}_N$ is given by

$$\mathbf{C}_N = \left(\mathbf{0}_{K\times L}, \mathbf{I}_K\right)\left[\widehat{\widetilde{\mathbf{Z}}}_N(\boldsymbol{\vartheta})' \widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta})\right]^{-1} \widehat{\widetilde{\mathbf{Z}}}'_N(\boldsymbol{\vartheta})\, \boldsymbol{\Omega}^{-1/2}_{u,N}\mathbf{Q}_{1,N}\mathbf{X}_N. \tag{A.43}$$

Hence the test statistics under the alternative can be written as

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

Observe that Assumption 4 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 $(NT)^{-1/2}\left(\widehat{\boldsymbol{\theta}}_{GLS,N}^0-\widehat{\boldsymbol{\theta}}_{W,N}\right) \xrightarrow{d.} N(\mathbf{0},\boldsymbol{\Sigma}_W-\boldsymbol{\Sigma}_{GLS})$ where by Lemma 1, $(\boldsymbol{\Sigma}_W-\boldsymbol{\Sigma}_{GLS})$ is a positive definite matrix. Thus also $p\lim_{N\to\infty}\left(\widehat{\boldsymbol{\theta}}_{GLS,N}^0-\widehat{\boldsymbol{\theta}}_{W,N}\right)=0$. Furthermore, it follows from Assumption 5 and the condition stipulated in the Proposition that limit of $\mathbf{C}_N$ exists and is given as

$$\begin{aligned} \mathbf{C} &= \lim_{N\to\infty}\mathbf{C}_N=\lim_{N\to\infty}(\mathbf{0}_{K\times L},\mathbf{I}_K)\left[\widehat{\widetilde{\mathbf{Z}}}_N(\boldsymbol{\vartheta})'\widetilde{\mathbf{Z}}_N(\boldsymbol{\vartheta})\right]^{-1}\widehat{\widetilde{\mathbf{Z}}}_N'(\boldsymbol{\vartheta})\boldsymbol{\Omega}_{u,N}^{-1/2}\mathbf{Q}_{1,N}\mathbf{X}_N \\ &= (\mathbf{0}_{K\times L},\mathbf{I}_K)\left(\mathbf{M}_{H_R\widetilde{Z}}'\mathbf{M}_{H_RH_R}^{-1}\mathbf{M}_{H_R\widetilde{Z}}\right)^{-1}\mathbf{M}_{H_R\widetilde{Z}}'\mathbf{M}_{H_RH_R}^{-1}\mathbf{M}_{H_R\widetilde{X}}, \end{aligned} \tag{A.45}$$

where the matrices $\mathbf{M}_{H_R\widetilde{Z}}$ and $\mathbf{M}_{H_R\widetilde{X}}$ have full column rank. As a result, the quadratic form defined by the matrix $\mathbf{C}'(\boldsymbol{\Sigma}_W-\boldsymbol{\Sigma}_{GLS})^{-1}\mathbf{C}$ is positive definite.

Observe now that

$$
\begin{aligned}
H_N^1 &= H_N^0 + (NT)\,\boldsymbol{\pi}'\mathbf{C}_N'\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1}\mathbf{C}_N\boldsymbol{\pi} \\
&\quad -2NT\boldsymbol{\pi}\mathbf{C}_N'\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1}\left(\widehat{\boldsymbol{\theta}}^0_{GLS,N} - \widehat{\boldsymbol{\theta}}_{W,N}\right) \\
&= H_N^0 + (NT)\,\boldsymbol{\pi}'\mathbf{C}_N'\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1}\mathbf{C}_N\boldsymbol{\pi} \qquad \text{(A.46)} \\
&\quad -2\sqrt{NT}\left[\boldsymbol{\pi}'\mathbf{C}' + o(1)\right]\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1} o_p(1) \\
&= o_p(1) + (NT)\,\boldsymbol{\pi}\mathbf{C}_N'\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1}\mathbf{C}_N\boldsymbol{\pi} - \sqrt{NT}o(1)\,o_p(1) \\
&= (NT)\,\boldsymbol{\pi}'\mathbf{C}_N'\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1}\mathbf{C}_N\boldsymbol{\pi} + o_p\left(\sqrt{NT}\right),
\end{aligned}
$$

where

$$
\lim_{N\to\infty}\mathbf{C}_N'\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1}\mathbf{C}_N = \mathbf{C}'\left(\boldsymbol{\Sigma}_W - \boldsymbol{\Sigma}_{GLS}\right)^{-1}\mathbf{C}, \qquad \text{(A.47)}
$$

is a positive definite matrix. Therefore, for any $\gamma > 0$, we have

$$
\lim_{N\to\infty} P\left(H_N^1 > h\right) = 1.
$$

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] Chamberlain, G., (1982), Multivariate Regression Models for Panel, *Journal of Econometrics* 18, 5-46.

[6] Cornwell, Ch., Schmidt, P. and D. Wyhowski, 1992, Simultaneous Equations and Panel Data, *Journal of Econometrics* 51, 151-181.

[7] Egger, P., 2000, A Note on the Proper Econometric Specification of the Gravity Equation , *Economics Letters* 66, 25-31.

[8] Greene, W., 2003, *Econometric Analysis*, 5th ed., Prentice Hall, New Jersey.

[9] Hausman, J.A., 1978, Specification Tests in Econometrics, *Econometrica* 46, 1251-1271.

[10] Horn, R.A. and Ch. R. Johnson, 1985, *Matrix Analysis*, Cambridge University Press, Cambridge UK.

[11] Kapoor, M., H.H. Kelejian and I.R. Prucha, 2007, Panel Data Models with Spatially Correlated Error Components, *Journal of Econometrics* 140, 97-130.

[12] 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.

[13] 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.

[14] 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.*

[15] 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.

[16] Korniotis, G. M., 2008, Estimating Panel Models with Internal and External Habit Formation, forthcoming in the *Journal of Business and Economic Statistics.*

[17] Lee, L., 2003, Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances, *Econometric Reviews* 22, 307-335.

[18] Lee L. and Yu J.R., 2010, Estimation of spatial autoregressive panel data models with fixed effects, *Journal of Econometrics* 154, 165-185.

[19] Mundlak, Y. 1978, On Pooling Time Series and Cross Section Data, *Econometrica* 46, 69-85.

[20] Mutl, Jan, 2006, *Dynamic Panel Data Models with Spatially Correlated Disturbances*, Dissertation, University of Maryland.

[21] Schmidt, P., 1976, *Econometrics*, Marcel Dekker, New York.

[22] Wooldridge, J., 2002, Econometric Analysis if Cross Section And Panel Data, MIT-Press, Cambridge MA.

[23] 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.

[24] Yu, J. R. de Jong and Lee, L., 2008, Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both n and T Are Large, *Journal of Econometrics* 146, 118-134.