Improving the J test in the SARAR model by likelihood-based estimation

Peter Burridge
Department of Economics and Related Studies, University of York

....

Preliminary and incomplete, please do not cite this version.

May 5, 2010

Abstract

It has been demonstrated recently that the empirical significance levels of the J-type tests introduced by Kelejian (2008) can be controlled in many cases by the use of a bootstrap to construct a reference distribution. When the spatial parameter estimates lie outside the invertibility region constructing bootstrap samples is problematic, however, and the present paper explores how far this practical obstacle may be removed by the use of likelihood-based moment conditions to construct the parameter estimates. The effects of different spatial weight patterns and sample size on the empirical significance levels and power of the tests are also investigated, with sample size found to be important but weight pattern less so.

1 Introduction

Kelejian (2008), introduces spatial extensions of the J-test of Davidson and MacKinnon (1981) for testing a null model, Model$_0$ against g alternatives in which it is not nested. The tests are constructed for SARAR models using the GMM-type estimators of Kelejian and Prucha (1998, 1999, 2008), and as shown by Burridge and Fingleton (2010) (BF) they can be somewhat liberal in various parts of the parameter space. To correct this size-inflation problem, BF propose use of a parametric bootstrap; this is quite effective, but is found to be subject to difficulties when the GMM estimators applied to the original data deliver spatial parameter estimates that lie outside the invertibility region. If the device of bootstrapping the tests is to be widely applied, this problem must be resolved, and the present paper explores an approach based on the use of moment conditions obtained from the Gaussian likelihood that ensures that the estimated model is invertible. The BF paper leaves open a number of questions, among them how sensitive to the form of the weights the J tests' properties are, and to what extent the test size inflation is a small sample problem. Both questions are addressed in the present paper, and it is found, encouragingly, that test significance levels are not greatly influenced by the form of weight matrices, while increasing sample size does appear to ameliorate the size-inflation that occurs when there is low (but not exactly zero) correlation between the regressor of a true null model and that in a false alternative model. The limiting case of zero correlation remains problematic, however, but may be of very limited practical significance.

The next section describes the spatial models between which the J-type tests are designed to discriminate, defines the moment-based estimators, and the test statistics. In Section 3 the bootstrap is introduced, Section 4 presents the experimental evidence, and Section 5 concludes.

2 The models, estimators, and the J-type tests

2.1 Null and alternative models

Following Kelejian (2008), the SARAR(1,1) model set-up is adopted, and it is set out below as in BF. Under the null hypothesis, Model$_0$ is true:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}_0\boldsymbol{\beta}_0+\lambda_0\mathbf{W}_0\mathbf{Y}+\mathbf{U}_0 \\ \mathbf{U}_0 &= \rho_0\mathbf{M}_0\mathbf{U}_0+\boldsymbol{\varepsilon}_0. \end{aligned} \tag{1}$$

Here, the matrix of exogenous variables, $\mathbf{X}_0,(n\times k_0)$ and the dependent variable, $\mathbf{Y},(n\times 1)$ are each measured without error, the $n\times n$ weight matrices, $\mathbf{W}_0$ and $\mathbf{M}_0$ are fixed *a priori*, and the unobserved shock vector, $\boldsymbol{\varepsilon}_0\sim IID(0,\sigma_0^2\mathbf{I}_n)$ independent of the exogenous regressors, $\mathbf{X}_0$. The parameters to be estimated are the slope coefficients, $\boldsymbol{\beta}_0$, the spatial lag and error coefficients, λ_0, and ρ_0, and the variance, σ_0^2. Under the alternative, the data are generated by a similar structure, Model$_1$:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}_1\boldsymbol{\beta}_1 + \lambda_1\mathbf{W}_1\mathbf{Y} + \mathbf{U}_1 \\ \mathbf{U}_1 &= \rho_1\mathbf{M}_1\mathbf{U}_1 + \boldsymbol{\varepsilon}_1. \end{aligned} \tag{2}$$

In fact, Kelejian allows for some finite number, $g \geq 1$, of alternatives of the same type in which Model$_0$ is not nested, but only the case, $g = 1$ is considered here.

2.2 Likelihood-based moment conditions

Suppose, in addition, that in (1) the disturbance, $\mathbf{v}_0$ is Normally distributed, and that both $\mathbf{W}_0$ and $\mathbf{M}_0$ arise from the row-standardisation of symmetric matrices of non-negative elements and that $w_{ii} = m_{ii} = 0$ $(i = 1, ..., n)$. Both matrices therefore have real eigenvalues. Under these conditions, when it exists, the log-likelihood may be written

$$\begin{aligned} \ln L_0 &= \frac{-n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}[\mathbf{C}_0\mathbf{Y} - \mathbf{X}_0\boldsymbol{\beta}]'\mathbf{B}_0'\mathbf{B}_0[\mathbf{C}_0\mathbf{Y} - \mathbf{X}_0\boldsymbol{\beta}] \\ & \quad + \ln|\mathbf{C}_0| + \ln|\mathbf{B}_0| \end{aligned} \tag{3}$$

where $\mathbf{C}_0 = \mathbf{I} - \lambda\mathbf{W}_0$, $\mathbf{B}_0 = \mathbf{I} - \rho\mathbf{M}_0$. For the reason given in LeSage and Pace (2009, p.26) the matrices $\mathbf{C}_0$ and $\mathbf{B}_0$ are required to be non-singular, with inverses that may be expressed as power series expansions in $\lambda\mathbf{W}_0$ or $\rho\mathbf{M}_0$ as the case may be. Denoting the eigenvalues of these matrices as ω_i, μ_i $(i = 1, ..., n)$ necessary and sufficient conditions are $\lambda \in (\omega_{\min}^{-1}, \omega_{\max}^{-1})$ and $\rho \in (\mu_{\min}^{-1}, \mu_{\max}^{-1})$ with the row standardisation assumed here implying that $\omega_{\max}^{-1} = \mu_{\max}^{-1} = 1$. Notice that in numerical maximization of (3) the solutions for both λ and ρ will be driven away from the invertibility boundary by the behaviour of the determinants comprising the Jacobian as the boundary is approached. The score vector corresponding to (3) is:

$$\begin{aligned} \frac{\partial \ln L_0}{\partial \boldsymbol{\beta}} &= \frac{1}{\sigma^2}\mathbf{X}_0'\mathbf{B}_0'\mathbf{B}_0\mathbf{U}_0 \\ \frac{\partial \ln L_0}{\partial \rho} &= -\frac{1}{2\sigma^2}\mathbf{U}_0'[2\rho\mathbf{M}_0'\mathbf{M}_0 - \mathbf{M}_0 - \mathbf{M}_0']\mathbf{U}_0 - Tr[\mathbf{M}_0\mathbf{B}_0^{-1}] \\ \frac{\partial \ln L_0}{\partial \lambda} &= \frac{1}{\sigma^2}\mathbf{Y}'\mathbf{W}_0'\mathbf{B}_0'\mathbf{B}_0\mathbf{U}_0 - Tr[\mathbf{W}_0\mathbf{C}_0^{-1}] \\ \frac{\partial \ln L_0}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\mathbf{U}_0'\mathbf{B}_0'\mathbf{B}_0\mathbf{U}_0 \end{aligned} \tag{4}$$

where $\mathbf{U}_0 = [\mathbf{C}_0\mathbf{Y} - \mathbf{X}_0\boldsymbol{\beta}]$. The score, (4) can of course be written in various forms; for present purposes it is useful to write $\mathbf{B}_0\mathbf{U}_0 = \boldsymbol{\varepsilon}_0$ or $\mathbf{U}_0 = \mathbf{B}_0^{-1}\boldsymbol{\varepsilon}_0$ as appropriate and to note that $2\rho\mathbf{M}_0'\mathbf{M}_0-\mathbf{M}_0-\mathbf{M}_0' = -[\mathbf{M}_0'\mathbf{B}_0+\mathbf{B}_0'\mathbf{M}_0]$ so that

$$\begin{aligned}
\frac{\partial \ln L_0}{\partial \boldsymbol{\beta}} &= \frac{1}{\sigma^2}\mathbf{X}_0'\mathbf{B}_0'\boldsymbol{\varepsilon}_0 && (5)\\
\frac{\partial \ln L_0}{\partial \rho} &= \frac{1}{2\sigma^2}\mathbf{U}_0'[\mathbf{M}_0'\mathbf{B}_0+\mathbf{B}_0'\mathbf{M}_0]\mathbf{U}_0 - Tr[\mathbf{M}_0\mathbf{B}_0^{-1}] \\
&= \frac{1}{2\sigma^2}\boldsymbol{\varepsilon}_0'\mathbf{B}_0^{-1\prime}[\mathbf{M}_0'\mathbf{B}_0+\mathbf{B}_0'\mathbf{M}_0]\mathbf{B}_0^{-1}\boldsymbol{\varepsilon}_0 - Tr[\mathbf{M}_0\mathbf{B}_0^{-1}] \\
&= \frac{1}{\sigma^2}\boldsymbol{\varepsilon}_0'\mathbf{M}_0\mathbf{B}_0^{-1}\boldsymbol{\varepsilon}_0 - Tr[\mathbf{M}_0\mathbf{B}_0^{-1}] && (6)\\
\frac{\partial \ln L_0}{\partial \lambda} &= \frac{1}{\sigma^2}\mathbf{Y}'\mathbf{W}_0'\mathbf{B}_0'\boldsymbol{\varepsilon}_0 - Tr[\mathbf{W}_0\mathbf{C}_0^{-1}] \\
&= \frac{1}{\sigma^2}\boldsymbol{\varepsilon}_0'\mathbf{B}_0\mathbf{W}_0\mathbf{C}_0^{-1}[\mathbf{X}_0\boldsymbol{\beta}+\mathbf{B}_0^{-1}\boldsymbol{\varepsilon}_0] - Tr[\mathbf{W}_0\mathbf{C}_0^{-1}] && (7)\\
\frac{\partial \ln L_0}{\partial \sigma^2} &= -\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\boldsymbol{\varepsilon}_0'\boldsymbol{\varepsilon}_0. && (8)
\end{aligned}$$

Setting (5), to (8) equal to zero and solving the resulting system of equations of course delivers the maximum likelihood estimator of $\boldsymbol{\theta} = (\boldsymbol{\beta}', \rho, \lambda, \sigma^2)'$. On the other hand, taking expectations in (5), to (8) yields the zero vector, so that the $\boldsymbol{\theta}$ that sets the score to zero can also be thought of as a method of moments estimator. In Lee and Liu (2009), on putting the optimal instruments described at the top of their Page 13 into their Equation 4, a set of moment conditions for this model is obtained, written in their notation as $E\{\mathbf{g}(\boldsymbol{\theta})\} = E\{[\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{P}_1\boldsymbol{\varepsilon}, \mathbf{P}_2\boldsymbol{\varepsilon}]'\boldsymbol{\varepsilon}\} = \mathbf{0}$ and, in our notation,

$$\begin{aligned}
\mathbf{Q}_1 &= \mathbf{B}_0\mathbf{X}_0 && (9)\\
\mathbf{Q}_2 &= \mathbf{B}_0\mathbf{W}_0\mathbf{C}_0^{-1}\mathbf{X}_0\boldsymbol{\beta} && (10)\\
\mathbf{P}_1 &= \mathbf{B}_0\mathbf{W}_0\mathbf{C}_0^{-1}\mathbf{B}_0^{-1} - n^{-1}Tr[\mathbf{W}_0\mathbf{C}_0^{-1}].\mathbf{I}_n && (11)\\
\mathbf{P}_2 &= \mathbf{M}_0\mathbf{B}_0^{-1} - n^{-1}Tr[\mathbf{M}_0\mathbf{B}_0^{-1}].\mathbf{I}_n && (12)
\end{aligned}$$

It is easily seen that with the addition of (8) these conditions correspond to those obtained from setting the score vector equal to zero, but with (7) split into two conditions: (10) involving $\mathbf{Q}_2$ being linear in $\boldsymbol{\varepsilon}$, (11) involving $\mathbf{P}_1$ being quadratic. Thus in Lee and Liu's framework over-identification is introduced: $\mathbf{Q}_1'\boldsymbol{\varepsilon}$ has dimension $(k_0 \times 1)$, $\mathbf{Q}_2'\boldsymbol{\varepsilon}$ has dimension (1×1), $\boldsymbol{\varepsilon}'\mathbf{P}_1\boldsymbol{\varepsilon}$ has dimension (1×1) and $\boldsymbol{\varepsilon}'\mathbf{P}_2\boldsymbol{\varepsilon}$ has dimension (1×1), thus there are k_0+3 moment conditions for the k_0+2 parameters, $\boldsymbol{\beta}$,λ, ρ to which must be added (8) to determine σ^2.

Lee and Liu (2009 Proposition 5) propose to implement such moment conditions by replacing $\mathbf{P}_i$ and $\mathbf{Q}_i$ by "feasible" versions using first round $\sqrt{n}-$consistent estimators, and then minimizing $\hat{\mathbf{g}}'(\boldsymbol{\theta})\hat{\boldsymbol{\Omega}}^{-1}\hat{\mathbf{g}}(\boldsymbol{\theta})$, say, in which $\hat{\boldsymbol{\Omega}}$ is a consistent estimator of the covariance matrix of $\mathbf{g}(\boldsymbol{\theta})$. However, neither the first round estimates of ρ or λ nor the final GMM estimates need lie in the invertibility region; in their experiments, Lee and Liu reported a few cases in which the final estimates were outside this region even for a sample size of 490 and true values of 0.4 for both λ and ρ (Lee and Liu 2009, Endnote 29). For smaller sample sizes or more extreme ρ or λ values one would expect to see many more such cases, which would be a problem in the present context. It is not necessary to introduce over-identification, however, and thus the potential problems introduced by a multi-stage approach may be avoided by fully solving for the MLE obtained by setting (5), to (8) equal to zero. This full Normal likelihood-based approach will therefore be adopted in the present paper.

Kelejian's J-tests specialised to the case, $g = 1$, are now described; the following is an abridgement of the relevant section of BF.

2.3 Kelejian's J-tests for the case g=1

Kelejian (2008) presents J-type tests based on estimates constructed using the generalised two-stage least squares estimator of Kelejian and Prucha (1998), combined with the generalised moments estimator of Kelejian and Prucha (1999). Suppose $g = 1$, so there is a single alternative to the null model. Write $\mathbf{X}_0 = [\mathbf{X}_{01} \vdots \mathbf{X}_{02}]$ and $\mathbf{X}_1 = [\mathbf{X}_{11} \vdots \mathbf{X}_{12}]$ in which $\mathbf{X}_{01} = \mathbf{X}_{11} = \mathbf{1} = [1, 1, ..., 1]'$ and $\mathbf{X}_{02}, \mathbf{X}_{12}$ are the remaining non-constant regressors in the two models. Further, write $\mathbf{Z}_0 = [\mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{W}_0\mathbf{Y}]$, $\boldsymbol{\gamma}_0 = [\boldsymbol{\beta}_0', \lambda_0]'$, $\mathbf{Z}_1 = [\mathbf{X}_{11} \vdots \mathbf{X}_{12} \vdots \mathbf{W}_1\mathbf{Y}]$, and $\boldsymbol{\gamma}_1 = [\boldsymbol{\beta}_1', \lambda_1]'$ so the null model is

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}_0\boldsymbol{\beta}_0 + \lambda_0\mathbf{W}_0\mathbf{Y} + \mathbf{U}_0 \\ &= \mathbf{Z}_0\boldsymbol{\gamma}_0 + \mathbf{U}_0. \end{aligned} \tag{13}$$

The necessary calculations can be described in four steps.

Step 1. Estimate $\boldsymbol{\gamma}_0$ and $\boldsymbol{\gamma}_1$ by instrumental variables (IV). Define the matrix, $\mathbf{L}_{0,r} = [\mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{W}_0\mathbf{X}_{02} .. \vdots \mathbf{W}_0^r\mathbf{X}_{02}]$, for some small integer r, and construct a matrix of instruments, $\mathbf{H}_{0,r} = [\mathbf{L}_{0,r} \vdots \mathbf{M}_0\mathbf{L}_{0,r}]_{LI}$, in which the subscript, LI denotes a spanning set of linearly independent columns. Define the associated projection matrix, $\mathbf{P}_{0,r} = \mathbf{H}_{0,r}(\mathbf{H}_{0,r}'\mathbf{H}_{0,r})^{-1}\mathbf{H}_{0,r}'$, leading to the IV estimator,

$$\hat{\boldsymbol{\gamma}}_{0r} = [\mathbf{Z}_0'\mathbf{P}_{0,r}\mathbf{Z}_0]^{-1}\mathbf{Z}_0'\mathbf{P}_{0,r}\mathbf{Y}.$$

Similarly, for Model$_1$ defining $\mathbf{L}_{1,r} = [\mathbf{X}_{11} \vdots \mathbf{X}_{12} \vdots \mathbf{W}_1\mathbf{X}_{12} .. \vdots \mathbf{W}_1^r\mathbf{X}_{12}]$, and $\mathbf{H}_{1,r} = [\mathbf{L}_{1,r} \vdots \mathbf{M}_1\mathbf{L}_{1,r}]_{LI}$, gives the projector, $\mathbf{P}_{1,r}$, and IV estimator,

$$\hat{\boldsymbol{\gamma}}_{1r} = [\mathbf{Z}_1'\mathbf{P}_{1,r}\mathbf{Z}_1]^{-1}\mathbf{Z}_1'\mathbf{P}_{1,r}\mathbf{Y}.$$

For convenience, we also define the hybrid matrix,

$$\mathbf{L}_{01,r} = [[\mathbf{X}_0 \vdots \mathbf{X}_1] \vdots \mathbf{W}_0[\mathbf{X}_0 \vdots \mathbf{X}_1] .. \vdots \mathbf{W}_0^r[\mathbf{X}_0 \vdots \mathbf{X}_1]]_{LI}$$

and introduce $\mathbf{H}_{01,r} = [\mathbf{L}_{01,r} \vdots \mathbf{M}_0\mathbf{L}_{01,r}]_{LI}$, the latter corresponding to Kelejian's $\mathbf{A}_n$.

Step 2. Estimate ρ_0 and ρ_1 by the non-linear GMM method of Kelejian and Prucha (1999) (KP). Define the vector of residuals from IV estimation of the null model as, $\hat{\mathbf{U}}_0 = \mathbf{Y} - \mathbf{Z}_0\hat{\boldsymbol{\gamma}}_0$. The KP estimator of ρ_0 is defined as follows. Let $\mathbf{D}$ be a fixed $n \times n$ matrix, then $\boldsymbol{\varepsilon}_0$ in (1) satisfies the second moment condition:$E\{n^{-1}\boldsymbol{\varepsilon}_0'\mathbf{D}\boldsymbol{\varepsilon}_0\} = \sigma^2 n^{-1}Tr\{\mathbf{D}\}$. Taking $\mathbf{D}$ to be respectively the identity, $\mathbf{I}$, or $\mathbf{M}_0'\mathbf{M}_0$, or $\mathbf{M}_0'$ yields the three conditions, (i) $E\{n^{-1}\boldsymbol{\varepsilon}_0'\boldsymbol{\varepsilon}_0\} = \sigma^2$, (ii) $E\{n^{-1}\boldsymbol{\varepsilon}_0'\mathbf{M}_0'\mathbf{M}_0\boldsymbol{\varepsilon}_0\} = \sigma^2 n^{-1}Tr\{\mathbf{M}_0'\mathbf{M}_0\}$, (iii) $E\{n^{-1}\boldsymbol{\varepsilon}_0'\mathbf{M}_0'\boldsymbol{\varepsilon}_0\} = \sigma^2 n^{-1}Tr\{\mathbf{M}_0'\} = 0$, which are easily converted into statements about the moments of $\mathbf{U}_0$ using the fact that $\boldsymbol{\varepsilon}_0 = \mathbf{B}_0\mathbf{U}_0$. This leads to the set of three non-linear equations in σ^2, ρ_0 and ρ_0^2 :

$$\begin{aligned} E\{n^{-1}\mathbf{U}_0'\mathbf{B}_0'\mathbf{B_0}\mathbf{U}_0\} - \sigma^2 &= 0 \\ E\{n^{-1}\mathbf{U}_0'\mathbf{B}_0'\mathbf{M}_\mathbf{0}'\mathbf{M_0}\mathbf{B}_0\mathbf{U}_0\} - \sigma^2 n^{-1}Tr\{\mathbf{M}_0'\mathbf{M}_0\} &= 0 \\ E\{n^{-1}\mathbf{U}_0'\mathbf{B}_0'\mathbf{M}_\mathbf{0}'\mathbf{B_0}\mathbf{U}_0\} &= 0 \end{aligned}$$

The estimates $\hat{\rho}$ and $\hat{\sigma}^2$ are obtained by replacing these three expectations with sample averages and minimizing the sum of squares of the three sample moments obtained when $\mathbf{U}_0$ is replaced by the residual, $\hat{\mathbf{U}}_0$.

Defining the vector of residuals from the alternative, $\hat{\mathbf{U}}_1 = \mathbf{Y} - \mathbf{Z}_1\hat{\boldsymbol{\gamma}}_1$ we estimate ρ in similar fashion to get $\hat{\rho}_1$, say.

Step 3. Estimate spatially lag-transformed (in feasible GLS form) regressions by IV. Using $\hat{\rho}_0$ from Step 2, construct the transformed regression

$$\begin{aligned} (\mathbf{I}_n - \hat{\rho}_0\mathbf{M}_0)\mathbf{Y} &= (\mathbf{I}_n - \hat{\rho}_0\mathbf{M}_0)(\mathbf{Z}_0\boldsymbol{\gamma}_0 + \mathbf{U}_0) \\ \mathbf{Y}^*(\hat{\rho}_0) &= \mathbf{Z}_0^*(\hat{\rho}_0)\boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}^*(\hat{\rho}_0) \quad \text{say} \end{aligned} \tag{14}$$

and estimate this equation by IV using the same instruments as before, $\mathbf{H}_{0,r}$; the result is the generalised spatial 2SLS procedure suggested in Kelejian and Prucha (1998) that yields, say,

$$\mathbf{Y}^*(\hat{\rho}_0) = \mathbf{Z}_0^*(\hat{\rho}_0)\hat{\boldsymbol{\gamma}}_0(\hat{\rho}_0) + \hat{\boldsymbol{\varepsilon}}^*(\hat{\rho}_0). \tag{15}$$

Use the residual vector, $\hat{\boldsymbol{\varepsilon}}^*(\hat{\rho}_0)$, to estimate the variance of the shocks, $\hat{\sigma}_0^2 = \hat{\boldsymbol{\varepsilon}}^*(\hat{\rho}_0)'\hat{\boldsymbol{\varepsilon}}^*(\hat{\rho}_0)/n$.

Similarly, using $\hat{\rho}_1$ from Step 2, construct the alternative spatially lag-transformed regression

$$\begin{aligned} (\mathbf{I}_n - \hat{\rho}_1\mathbf{M}_1)\mathbf{Y} &= (\mathbf{I}_n - \hat{\rho}_1\mathbf{M}_1)(\mathbf{Z}_1\boldsymbol{\gamma}_1 + \mathbf{U}_1) \\ \mathbf{Y}^*(\hat{\rho}_1) &= \mathbf{Z}_1^*(\hat{\rho}_1)\boldsymbol{\gamma}_1 + \boldsymbol{\varepsilon}^*(\hat{\rho}_1) \quad \text{say} \end{aligned} \tag{16}$$

and estimate it by IV using the instruments, $\mathbf{H}_{1,r}$ to obtain

$$\mathbf{Y}^*(\hat{\rho}_1) = \mathbf{Z}_1^*(\hat{\rho}_1)\hat{\boldsymbol{\gamma}}_1(\hat{\rho}_1) + \hat{\boldsymbol{\varepsilon}}^*(\hat{\rho}_1). \tag{17}$$

Let $\hat{\mathbf{Y}}^*(\hat{\rho}_1)$ denote the fitted value from (17). The RHS of (14) is now augmented to generate a test of the hypothesis that Model_0 is true.

Kelejian defines two tests:
Step 4a.(conjectured χ_1^2 version)
Using the fitted value from (17), set up the augmented equation

$$\begin{aligned}\mathbf{Y}^*(\hat{\rho}_0) &= \mathbf{Z}_0^*(\hat{\rho}_0)\boldsymbol{\gamma}_0 + \hat{\mathbf{Y}}^*(\hat{\rho}_1)\delta + \boldsymbol{\varepsilon}^*(\hat{\rho}_0) \\ &= \mathbf{Z}^{**}\boldsymbol{\gamma}^{**}+\boldsymbol{\varepsilon}^{**}\end{aligned} \tag{18}$$

and the augmented matrix of instruments

$$\mathbf{H}_r^{**} = [\mathbf{H}_{0,r} \vdots \mathbf{H}_{01,r}]_{LI} \tag{19}$$

with projection matrix $\mathbf{P}_r^{**}$, say, obtaining the IV estimator

$$\hat{\boldsymbol{\gamma}}^{**} = (\mathbf{Z}'^{**}\mathbf{P}_r^{**}\mathbf{Z}^{**})^{-1}\mathbf{Z}'^{**}\mathbf{P}_r^{**}\mathbf{Y}^*(\hat{\rho}_0)$$

with estimated asymptotic covariance matrix, $\hat{\mathbf{V}} = \hat{\sigma}_0^2(\mathbf{Z}'^{**}\mathbf{P}_r^{**}\mathbf{Z}^{**})^{-1}$, which is used to extract a Wald test statistic for $\delta = 0$ in (18) in the usual way. That is, when Model$_0$ is true, letting l denote the number of elements in $\boldsymbol{\gamma}^{**}$, so that $\hat{\boldsymbol{\gamma}}^{**}(l)$ is the last estimated coefficient, and $\hat{\mathbf{V}}(l,l)$ its estimated variance, it is conjectured that,

$$\frac{(\hat{\boldsymbol{\gamma}}^{**}(l))^2}{\hat{\mathbf{V}}(l,l)} \rightarrow^d \chi^2_{(1)}. \tag{20}$$

This remains a conjecture - see the first Remark under Kelejian's Equation 9, though the experiments reported by BF and those below indicate that (20) is a good approximation in most of the parameter space even in samples of modest size. The specification of $\mathbf{H}_{01,r}$ in the instrument set, (19), is as given by Kelejian (2008); however, in the experiments of BF it was found that test power improved dramatically in some cases if $\mathbf{H}_{1,r}$ was used in place of $\mathbf{H}_{01,r}$ at this point.

Step 4b.(χ_2^2 version)

Use the first step estimates, $\hat{\boldsymbol{\gamma}}_1$, to augment the RHS of (14) with both $\mathbf{Z}_1\hat{\boldsymbol{\gamma}}_1$ and $\mathbf{M}_1\mathbf{Z}_1\hat{\boldsymbol{\gamma}}_1$, in place of the single forecast value, $\hat{\mathbf{Y}}^*(\hat{\rho}_1)$, and augment the instrument vector as before. Following the same line of development leads to a statistic that is asymptotically $\chi^2_{(2)}$. That is, now estimate the equation

$$\begin{aligned}\mathbf{Y}^*(\hat{\rho}_0) &= \mathbf{Z}_0^*(\hat{\rho}_0)\boldsymbol{\gamma}_0 + \mathbf{Z}_1\hat{\boldsymbol{\gamma}}_1\delta_1 + \mathbf{M}_1\mathbf{Z}_1\hat{\boldsymbol{\gamma}}_1\delta_2 + \boldsymbol{\varepsilon}^\dagger(\hat{\rho}_0) \\ &= \mathbf{Z}^\dagger\boldsymbol{\gamma}^\dagger+\boldsymbol{\varepsilon}^\dagger\end{aligned} \tag{21}$$

using the instruments, $\mathbf{H}_r^{**}$, as above, obtaining the IV estimator,

$$\hat{\boldsymbol{\gamma}}^\dagger = (\mathbf{Z}'^\dagger\mathbf{P}_r^{**}\mathbf{Z}^\dagger)^{-1}\mathbf{Z}'^\dagger\mathbf{P}_r^{**}\mathbf{Y}^*(\hat{\rho}_0),$$

with estimated asymptotic covariance matrix, $\hat{\mathbf{V}}^\dagger = \hat{\sigma}_0^2(\mathbf{Z}'^\dagger\mathbf{P}_r^{**}\mathbf{Z}^\dagger)^{-1}$. Writing the matrix that selects the final two elements of $\boldsymbol{\gamma}^\dagger$ in the usual way as $\mathbf{R} = \left[\mathbf{0}...\mathbf{0}\vdots\mathbf{I}_2\right]$, the hypothesis to be tested is $H_0 : \mathbf{R}\boldsymbol{\gamma}^\dagger = \mathbf{0}$ and a Wald test statistic is

$$\hat{\boldsymbol{\gamma}}'^\dagger\mathbf{R}'[\mathbf{R}\hat{\mathbf{V}}^\dagger\mathbf{R}']^{-1}\mathbf{R}\hat{\boldsymbol{\gamma}}^\dagger \rightarrow^d \chi^2_{(2)}. \tag{22}$$

Kelejian proves (22) while the alternative one degree-of-freedom form, (20), is introduced in a remark that also raises the question of the relative efficiency of the two tests, commented on below.

2.4 Likelihood-based versions of the J-tests for g=1

Notice, first of all, that the parameter estimates obtained in Step 3 above from the two competing models could also be obtained by either the Lee and Liu GMM estimator or by maximum likelihood, or indeed by any consistent estimator. Adopting the maximum likelihood estimator, in place of $\hat{\mathbf{Y}}^*(\hat{\rho}_1)$ obtained from IV estimation of (17), we can construct the fitted value,

$$\tilde{\mathbf{Y}}^*(\tilde{\rho}_1) = \mathbf{Z}_1^*(\tilde{\rho}_1)\tilde{\boldsymbol{\gamma}}_1 \tag{23}$$

in which

$$\mathbf{Z}_1^*(\tilde{\rho}_1) = (\mathbf{I}_n - \tilde{\rho}_1\mathbf{M}_1)\mathbf{Z}_1 = \tilde{\mathbf{B}}_1\mathbf{Z}_1 \tag{24}$$

$$\tilde{\boldsymbol{\gamma}}_1 = [\tilde{\boldsymbol{\beta}}_1^{'}, \tilde{\lambda}_1]' \tag{25}$$

where $\tilde{\rho}_1, \tilde{\lambda}_1$ and $\tilde{\boldsymbol{\beta}}_1^{'}$ are the ML estimators obtained from Model 1. Similarly, define

$$\mathbf{Y}^*(\tilde{\rho}_0) = (\mathbf{I}_n - \tilde{\rho}_0\mathbf{M}_0)\mathbf{Y} \tag{26}$$

$$\mathbf{Z}_0^*(\tilde{\rho}_0) = (\mathbf{I}_n - \tilde{\rho}_0\mathbf{M}_0)\mathbf{Z}_0. \tag{27}$$

The ML-based test statistics

Step 4a, above, is implemented with, in place of (18), the augmented equation

$$\mathbf{Y}^*(\tilde{\rho}_0) = \mathbf{Z}_0^*(\tilde{\rho}_0)\boldsymbol{\gamma}_0 + \tilde{\mathbf{Y}}^*(\tilde{\rho}_1)\delta + \boldsymbol{\varepsilon}^{**}. \tag{28}$$

The test statistic for the restriction, $\delta = 0$ is then constructed from IV estimation of $\boldsymbol{\gamma}_0$ and δ in (28) similarly to Step 4a. To construct the test in Step 4b, both $\mathbf{Z}_1\tilde{\boldsymbol{\gamma}}_1$ and $\mathbf{M}_1\mathbf{Z}_1\tilde{\boldsymbol{\gamma}}_1$, are included in place of the single forecast value, $\tilde{\mathbf{Y}}^*(\tilde{\rho}_1)$, in the augmented equation and

$$\mathbf{Y}^*(\tilde{\rho}_0) = \mathbf{Z}_0^*(\tilde{\rho}_0)\boldsymbol{\gamma}_0 + \mathbf{Z}_1\tilde{\boldsymbol{\gamma}}_1\delta_1 + \mathbf{M}_1\mathbf{Z}_1\tilde{\boldsymbol{\gamma}}_1\delta_2 + \boldsymbol{\varepsilon}^{\dagger} \tag{29}$$

is estimated by IV similarly to Step 4b, above, and the test statistic is constructed in the same way.

3 The bootstrap

It was shown in BF that by resampling it was possible to gain control over empirical significance levels of the tests as originally formulated, in many cases. Here, it is anticipated that the same will be true of the reformulated tests, but the expectation is also that the severe inflation of significance levels for some parameter combinations that could not previously be corrected will now be either eliminated or much reduced.

3.1 The simple resampling scheme:

Compute the J test statistics as above, then

(i) Using the whitened residuals from Model$_0$ as the building block, draw a random sample using sampling with replacement; call this random sample, $\mathbf{e}^*$

(ii) Using $\tilde{\rho}_0$ the MLE of ρ from Model$_0$, generate

$$\mathbf{u}^* = [\mathbf{I} - \hat{\rho}_0 \mathbf{M}_0]^{-1} \mathbf{e}^*$$

(iii) Recall

$$\tilde{\boldsymbol{\gamma}}_0 = [\tilde{\boldsymbol{\beta}}_0^{'}, \tilde{\lambda}_0]'$$

and generate

$$\mathbf{Y}^* = [\mathbf{I} - \tilde{\lambda}_0 \mathbf{W}_0]^{-1} (\mathbf{X}_0 \tilde{\boldsymbol{\beta}}_0 + \mathbf{u}^*)$$

(iv) Calculate the J statistic using the $\mathbf{Y}^*$ sample

(v) Repeat (ii)-(iv) the designated number of times, m, to create a sample from the bootstrap distribution of the relevant J statistic.

(vi) If the proportion of the m bootstrap replicates that exceed the observed J statistic is less than the chosen significance level, reject the null hypothesis at that level.

4 Experimental Results

Four questions are considered. (1) Will estimating the spatial parameters by maximum likelihood produce test statistics less prone to the severe inflation of significance levels identified for certain parts of the parameter space in BF? To answer this the experiments of BF are repeated but with the tests based on the MLE as set out in Section 2.4, and the results compared. (2) Can the shortcomings of the bootstrap identified in BF be fixed by estimating the spatial parameters, λ and ρ, by maximum likelihood? This is addressed by comparing test performance on the extreme cases previously identified. (3) How important is the form of the weights matrices? To provide evidence on this point, two different sets of weights for each of two sets of 25 regions are employed, the regions being arranged in either a 5×5 square or a continuous ring. (4) How much do results improve with increasing sample size? The cases for which the bootstrapped tests are most liberal are repeated for larger samples, with the weights for $n = 400$ being represented by a block diagonal matrix formed of two copies of the weight matrices used for $n = 200$.

The experimental set-up is the same as in BF except that use of the MLE constrains both λ and ρ to the invertible region. The results reported here are based on tests with nominal significance level 5% and with $m = 99$, $s = 5000$. The standard error of the estimated significance levels is therefore approximately 0.003. As in BF, the case in which the two models have different regressors but the same spatial weights is designated Case 1, and the case in which they have the same regressors but different spatial weights is Case 2. More complicated cases are not investigated here.

Case 1. This is implemented with a single explanatory variable other than the constant, that is, the regressors are $\mathbf{X}_0 = [\mathbf{X}_{01} \vdots \mathbf{X}_{02}]$ where $\mathbf{X}_{01} = \mathbf{1}$, the vector, $[1, 1, ..., 1]'$and $\mathbf{X}_{02}$ is a draw from $N(\mathbf{0}, \mathbf{I}_n)$, and the two spatial weight matrices are equal, $\mathbf{M}_0 = \mathbf{W}_0$, while the alternative has the same spatial structure, $\mathbf{M}_1 = \mathbf{W}_1 = \mathbf{W}_0$, but the explanatory variable, $\mathbf{X}_{02}$ is replaced by another that is in general correlated with it, constructed in the experiments as, $\mathbf{X}_{12} = \rho_x \mathbf{X}_{02} + (1 - \rho_x^2)^{1/2} \times N(\mathbf{0}, \mathbf{I})$ for various ρ_x values, including zero.

Case 2. This is implemented by having the explanatory variables, $\mathbf{X}_0$ and $\mathbf{X}_1$, the same in the two models ($\rho_x = 1$), but the spatial structures differ, so that $\mathbf{W}_1 \neq \mathbf{W}_0$ and $\mathbf{M}_1 \neq \mathbf{M}_0$; for simplicity, $\mathbf{M}_0 = \mathbf{W}_0$ and $\mathbf{M}_1 = \mathbf{W}_1 \neq \mathbf{W}_0$. Setting $\mathbf{M}_i = \mathbf{W}_i$ here causes no loss of identification because of the presence of the non-constant regressors, $\mathbf{X}_{i2}$.

4.1 The set up

For comparability with BF, two spatial frameworks are used for the main results, the 26 counties of Ireland, with weight matrix as employed in Cliff and Ord (1973, p. 164), and a set of 200 EU NUTS-2 regions with weight matrix $\mathbf{W}_0$ based on a matrix of 1s and 0s denoting contiguous and non-contiguous regions respectively, subsequently normalised so that rows sum to 1, as used by Fingleton (2007). Under Case 2, the 26 county alternative weight matrix, $\mathbf{W}_1$, is defined by replacing the non-zero elements of row i of the corresponding $\mathbf{W}_0$ by $n_{i.}^{-1}$ the reciprocal of the number of non-zero entries in the i^{th} row. For the 200 EU regions, with $w_{ij} = d_{ij}^{-2}$ for $d_{ij} \leq 300km$ and $d_{ij} = 0$ otherwise, where d_{ij} is the straight line (Euclidean) distance between regions i and j, the weights are defined as $\mathbf{W}_{1ij} = \frac{w_{ij}}{\sum_j w_{ij}}$. Thus the tests are having to discriminate between quite similar weight matrices. For the auxiliary experiments, two matrices with $n = 25$ are obtained from a 5×5 square, with rook's case or queen's case contiguity weights, again row normalised, and two more from a closed ribbon of 25 regions with either equal weights of 0.5 on each of two neighbours, or linearly declining weights over four neighbours on each side of 0.2, 0.15, 0.1, 0.05. Finally, to study larger n, matrices are built from diagonal blocks each equal to one of the matrices used for $n = 200$.

For Case 1 the instrument set uses $r \in (0, 1, 2)$ that is, a minimal, intermediate, and a rich set. In Case 2 only $r \in (1, 2)$ is relevant for the two-degree-of-freedom test, while the one-degree-of-freedom test is computable with $r = 0$. The number of replications is as given in the tables. The following parameter values are used: $(\rho_0, \lambda_0) \in (0.0, 0.3, 0.6, 0.9, 0.95)^2$ and $(\rho_0, \lambda_0) = (\rho_1, \lambda_1)$ so that empirical significance levels and powers reflect solely differences between the explanatory variables (Case 1) or weight matrices (Case 2); in Case 1 the explanatory variables observed for region i, $\mathbf{X}_{02i}$ and $\mathbf{X}_{12i}$ $(i = 1, ..., n)$ are drawn from a bivariate Normal distribution with variances unity and correlation $\rho_x \in (-0.5, 0.0, 0.5, 0.9, 0.95)$; the shocks are independent standard Normal, $\mathbf{v}_0 \sim IIDN(0, \mathbf{I}_n)$ and similarly $\mathbf{v}_1$ in each case. The matrices, $(\mathbf{I} - \lambda_i \mathbf{W}_i)$ and $(\mathbf{I} - \rho_i \mathbf{M}_i)$ $i = 0, 1$ are non-singular at the chosen parameter values.

4.2 Performance

This section describes the performance of the two forms of test statistic, which here have either 1 or 2 degrees of freedom, when referred to critical values from the corresponding χ^2 distribution, designated as the "Asymptotic test" and when referred to a bootstrap distribution, the "Simple bootstrap test".

4.2.1 Case 1

Summary results for the 375 sets of empirical size and power estimates of the asymptotic tests for Case 1 are given in Table 1A for the small sample size, $n = 26$. Columns 3-6 are based on $s = 40,000$ replications and columns 7-10 on $s = 5000$ replications.

Table 1A
Case 1 $N = 26$ Cliff-Ord weights, asymptotic test

		GTSLS based				MLE based			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.043	.040	.076	.019	.06	.06	.10	.04
$1d.f.$	1	.072	.067	.188	.043	.07	.06	.17	.04
$1d.f.$	2	.092	.079	.251	.053	.08	.07	.22	.04
$2d.f.$	0	.030	.027	.047	.016	.04	.04	.07	.03
$2d.f.$	1	.057	.054	.144	.030	.05	.05	.12	.03
$2d.f.$	2	.072	.063	.205	.043	.07	.06	.17	.03
Power	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.41	.39	.76	.10	.71	.90	.96	.27
$1d.f.$	1	.58	.51	.87	.31	.75	.94	.99	.30
$1d.f.$	2	.67	.60	.93	.33	.76	.96	.99	.29
$2d.f.$	0	.33	.30	.67	.06	.63	.83	.93	.17
$2d.f.$	1	.53	.53	.80	.21	.68	.88	.97	.20
$2d.f.$	2	.65	.71	.89	.22	.69	.90	.98	.20

For the minimal instrument set, with $r = 0$, there is evidence that the MLE-based tests are slightly more liberal, but a lot more powerful than their GTSLS-based counterparts; for $r = 1$ or 2 the size differences go in the opposite direction, but there are significant power advantages to using the MLE-based procedure. Table 1B shows the bootstrap working very well here, with effective overall size-control and no power loss.

Table 1B
Case 1 $N = 26$ Simple MLE-based bootstrap test

Size	r	Mean	Med	Max	Min
$1d.f.$	0	.05	.05	.07	.04
$1d.f.$	1	.05	.05	.08	.04
$1d.f.$	2	.05	.05	.08	.04
$2d.f.$	0	.05	.05	.07	.04
$2d.f.$	1	.05	.05	.08	.04
$2d.f.$	2	.05	.05	.08	.04
Power	r	Mean	Med	Max	Min
$1d.f.$	0	.70	.87	.96	.25
$1d.f.$	1	.72	.90	.98	.26
$1d.f.$	2	.72	.90	.98	.25
$2d.f.$	0	.65	.84	.94	.19
$2d.f.$	1	.66	.84	.97	.19
$2d.f.$	2	.65	.82	.97	.19

Table 1AA demonstrates very similar results for the MLE method applied to the $N = 25$ element ring with two equally-weighted neighbours. The bootstrap works even better here, and the better power of the 1 $d.f.$ tests is again clear.

Table 1AA
Case 1 $N = 25$ Ring form, two neighbours

		Asymptotic MLE				Bootstrap MLE			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.06	.06	.08	.04	.05	.05	.06	.04
$1d.f.$	1	.07	.06	.11	.04	.05	.05	.06	.04
$1d.f.$	2	.07	.07	.14	.04	.05	.05	.06	.04
$2d.f.$	0	.04	.04	.07	.02	.05	.05	.06	.04
$2d.f.$	1	.05	.05	.11	.02	.05	.05	.06	.04
$2d.f.$	2	.07	.06	.16	.03	.05	.05	.07	.04
Power	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.69	.79	.98	.25	.66	.75	.98	.24
$1d.f.$	1	.73	.90	.99	.25	.68	.84	.98	.23
$1d.f.$	2	.74	.92	.99	.27	.68	.83	.99	.23
$2d.f.$	0	.59	.70	.95	.15	.61	.70	.98	.19
$2d.f.$	1	.64	.80	.97	.16	.61	.74	.98	.17
$2d.f.$	2	.66	.84	.98	.19	.60	.72	.98	.17

Results for the MLE-based tests using the $N = 200$ EU NUTS regional weights appear in Table 1C. Evidently the bootstrap gives a slight improvement over the asymptotic test here, but the most significant differences between these results and the results obtained using the GTSLS-based tests are revealed in Tables 4 and 5 which focus on the extremes, discussed below.

Table 1C
Case 1 $N = 200$ EU NUTS

		Asymptotic MLE				Bootstrap MLE			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.05	.05	.14	.03	.05	.05	.09	.04
$1d.f.$	1	.06	.05	.22	.04	.05	.05	.14	.04
$1d.f.$	2	.06	.05	.24	.03	.05	.05	.14	.04
$2d.f.$	0	.05	.05	.11	.03	.05	.05	.09	.04
$2d.f.$	1	.06	.05	.19	.03	.05	.05	.14	.04
$2d.f.$	2	.06	.05	.21	.03	.05	.05	.14	.04

Powers in Table 1C are all in excess of .95

Table 2 lists the upper 10 and lower 10 order statistics of the sample of empirical sizes of the 1 degree-of-freedom test and the corresponding values of r, ρ_x, ρ_0 and λ_0 and Table 3 gives similar information for the 2 degree-of-freedom test. For the parameter values in the left half of the tables, using the 95% quantiles of the respective Chi-square distribution as critical values evidently gives a poor approximation to the true 95% quantiles of the sampling distributions of the J-type statistics for this particular weight matrix and r values. Notice that in neither table does $r = 0$ appear in the upper extremes.

Table 2: Extreme empirical sizes, $N = 26$, 1 degree-of-freedom test, Case 1
using χ_1^2 critical value, $s = 40,000$ for cols 1 and 9, $s = 5000$ for cols 3 and 11
using bootstrap critical value, $s = 5,000$, $m = 99$

Upper extreme sizes[a]								Lower extreme sizes[a]							
GTSLS		MLE		parameters				GTSLS		MLE		parameters			
χ_1^2	BS	χ_1^2	BS	r	ρ_x	λ_0	ρ_0	χ_1^2	BS	χ_1^2	BS	r	ρ_x	λ_0	ρ_0
.25	.09	.22	.08	2	0	0	.95	.02	.05	.04	.05	0	.95	.95	.90
.25	.08	.19	.07	2	0	0	.90	.02	.05	.05	.05	0	.95	.90	.95
.19	.05	.17	.07	2	0	.3	.90	.02	.05	.04	.04	0	-.5	.95	.90
.18	.08	.17	.08	1	0	0	.95	.02	.06	.05	.05	0	-.5	.90	.95
.18	.05	.18	.07	2	0	.3	.95	.02	.06	.05	.05	0	.5	.90	.95
.18	.07	.15	.07	1	0	0	.90	.02	.05	.05	.05	0	.95	.95	.95
.18	.07	.12	.05	2	0	0	.60	.02	.06	.05	.04	0	0	.95	.95
.17	.07	.14	.08	2	.5	0	.95	.02	.05	.04	.04	0	.5	.95	.95
.17	.07	.12	.06	2	-.5	0	.95	.02	.05	.04	.04	0	-.5	.95	.95
.16	.06	.12	.06	2	.5	0	.90	.02	.05	.04	.04	0	.90	.95	.95

a: sizes in descending order using the χ^2 critical values
BS: simple bootstrap size

Table 3: Extreme empirical sizes, $N = 26$, 2 degree-of-freedom test, Case 1
using χ^2_2 critical value, $s = 40000$ for cols 1 and 9, $s = 5000$ for cols 3 and 11;
using bootstrap critical value, $s = 5000$, $m = 99$

Upper extreme sizes^a								Lower extreme sizes^a							
GTSLS		MLE		parameters				GTSLS		MLE		parameters			
χ^2_2	BS	χ^2_2	BS	r	ρ_x	λ_0	ρ_0	χ^2_2	BS	χ^2_2	BS	r	ρ_x	λ_0	ρ_0
.20	.09	.17	.08	2	0	0	.95	.02	.05	.03	.04	0	.95	.95	.90
.19	.08	.16	.08	2	0	0	.90	.02	.05	.04	.05	0	.95	.90	.95
.15	.08	.15	.08	2	.5	0	.95	.02	.05	.04	.05	0	-.5	.90	.95
.15	.08	.15	.07	2	-.5	0	.95	.02	.05	.03	.04	0	.5	.90	.95
.15	.05	.12	.06	2	0	.3	.90	.02	.05	.03	.05	0	0	.95	.95
.14	.08	.12	.08	1	0	0	.95	.02	.05	.03	.04	0	.5	.95	.95
.14	.07	.13	.06	2	.5	0	.90	.02	.05	.03	.04	0	-.5	.95	.90
.14	.05	.12	.07	2	0	.3	.95	.02	.05	.03	.04	0	.95	.95	.95
.14	.08	.13	.07	2	-.5	0	.90	.02	.05	.03	.04	0	-.5	.95	.95
.13	.07	.10	.06	1	0	0	.90	.02	.05	.03	.04	0	.90	.95	.95

a: sizes in descending order using the χ^2 critical values
BS: simple bootstrap size

Table 4: Extreme empirical sizes, $n = 200$ lattice, 1 degree-of-freedom test,
Case 1
using χ^2_1 critical value, $s = 40000$ for cols 1 and 9, $s = 5000$ for cols 3 and 11;
using bootstrap critical value, $s = 5000$, $m = 99$

Upper extreme sizes^a								Lower extreme sizes^a							
GTSLS		MLE		parameters				GTSLS		MLE		parameters			
χ^2_1	BS	χ^2_1	BS	r	ρ_x	λ_0	ρ_0	χ^2_1	BS	χ^2_1	BS	r	ρ_x	λ_0	ρ_0
.64	.45	.24	.14	2	0	0	0.95	.03	.05	.04	.05	0	-.5	.90	.90
.56	.41	.20	.11	2	0	.3	0.95	.03	.05	.04	.05	0	-.5	.95	.90
.51	.37	.22	.14	1	0	0	0.95	.03	.04	.05	.05	0	.5	.90	.90
.47	.30	.14	.05	2	0	0	0.90	.03	.05	.03	.04	0	0	.95	.90
.44	.35	.18	.10	1	0	.3	0.95	.05	.04	.04	.05	0	-.5	.90	.95
.41	.29	.15	.06	2	0	.3	0.90	.03	.05	.04	.05	0	-.5	.95	.95
.35	.23	.13	.06	1	0	0	0.90	.03	.05	.04	.05	0	0	.90	.95
.34	.27	.14	.09	2	-.5	0	0.95	.03	.04	.04	.05	0	0	.95	.95
.32	.28	.14	.09	2	.5	0	0.95	.03	.05	.04	.05	0	.5	.90	.95
.31	.22	.13	.07	1	0	.3	0.90	.03	.04	.04	.05	0	.5	.95	.95

a : GTSLS sizes in descending order using the χ^2 critical values
BS: simple bootstrap size

Table 5: Extreme empirical sizes, $N = 200$, 2 degree-of-freedom test, Case 1 using χ^2_2 critical value, $s = 40000$ for cols 1 and 9, $s = 5000$ for cols 3 and 11; using bootstrap critical value, $s = 5000$, $m = 99$

Upper extreme sizes[a]								Lower extreme sizes[a]							
GTSLS		MLE		parameters				GTSLS		MLE		parameters			
χ^2_2	BS	χ^2_2	BS	r	ρ_x	λ_0	ρ_0	χ^2_2	BS	χ^2_2	BS	r	ρ_x	λ_0	ρ_0
.60	.47	.21	.14	2	0	0	.95	.02	.04	.03	.04	0	.9	.95	0.90
.51	.40	.20	.13	2	-.5	0	.95	.02	.04	.04	.05	0	-.5	.90	0.95
.50	.41	.16	.11	2	0	.3	.95	.02	.05	.04	.05	0	-.5	.95	0.95
.49	.41	.20	.13	2	.5	0	.95	.02	.05	.04	.05	0	0	.90	0.95
.47	.38	.20	.14	1	0	0	.95	.02	.04	.03	.04	0	.5	.95	0.90
.41	.35	.20	.13	2	-.5	.3	.95	.02	.05	.04	.05	0	.95	.90	0.95
.41	.34	.12	.06	2	0	0	.90	.02	.04	.03	.05	0	.95	.95	0.95
.40	.34	.14	.10	2	.5	.3	.95	.02	.05	.04	.05	0	.5	.95	0.95
.40	.34	.18	.13	1	-.5	0	.95	.02	.04	.04	.05	0	.9	.90	0.95
.40	.34	.18	.13	1	.5	0	.95	.02	.05	.04	.05	0	.9	.95	0.95

a: GTSLS sizes in descending order using the χ^2 critical values
BS: simple bootstrap size

Table 5A
Case 1 $N = 400$ (2 diagonal blocks of EU NUTS)
most over-sized cases from $N = 200$ only

		Asymptotic MLE				Bootstrap MLE			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.048	.047	.054	.044	.049	.049	.052	.044
$1d.f.$	1	.053	.052	.060	.048	.050	.050	.057	.047
$1d.f.$	2	.055	.056	.063	.045	.052	.053	.055	.046
$2d.f.$	0	.050	.050	.054	.046	.049	.050	.054	.046
$2d.f.$	1	.064	.064	.076	.052	.051	.051	.056	.046
$2d.f.$	2	.075	.079	.095	.049	.051	.051	.057	.043

$\rho_x \in (-0.5, 0.0, 0.5)$
$\lambda \in (0.0, 0.3)$
$\rho \in (0.9, 0.95)$
All powers in the table are equal to 1.0 and are therefore not shown.

The upper extreme sizes in Tables 4 and 5 recorded for the GTSLS-based tests are much more liberal than is the case for the smaller sample size, while the corresponding empirical sizes of the MLE-based tests are of the same order of magnitude as for the smaller sample. In this case, it is very clear that basing the tests on parameter estimates constrained to the invertibility region has dramatically improved performance. However, it is also clear that even after bootstrapping the test statistics they remain quite liberal for certain parameter combinations. For the reasons described in BF, this problem is clearly related to a lack of precision of the parameter estimates obtained under the null hypothesis that feeds through into the augmented regression from which the test statistics are derived. The problem cannot be fully eliminated using a simple bootstrap because the bootstrap itself relies on these same imprecise parameter estimates. At this stage the best we can do seems to be to use a minimal instrument set, $r = 0$, together with estimation by MLE. However, Table 5A summarises a small experiment that highlights the region of the parameter space in which the results for $N = 200$ are particularly unsatisfactory. It appears that the greater precision of the parameter estimates obtained from the larger sample has indeed largely eliminated the size inflation problem for these cases. However, it is still necessary to exercise some caution in applying the J test to models in which regressors are very weakly correlated. To illustrate, Table 5B (not yet available) summarises results obtained by replacing the regressor in the alternative model by the residual from its projection onto the regressor in Model_0. Thus in the sample, the new regressors in the two models are exactly orthogonal.

Table 5B
Case 1 $N = 400$ (2 diagonal blocks of EU NUTS)
Regressors in Model_0 and Model_1 exactly orthogonal

		Asymptotic MLE				Bootstrap MLE			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0								
$1d.f.$	1								
$1d.f.$	2								
$2d.f.$	0								
$2d.f.$	1								
$2d.f.$	2								

Finally, we note that the powers obtained in these experiments are most encouraging, especially for the larger samples, indicating that the tests have high discriminatory power.

4.2.2 Case 2

For the reason given in BF the instruments, $\mathbf{H}_{1,r}$ are used in (19) in place of $\mathbf{H}_{01,r}$ for the estimation of the augmented equation. Summary statistics for the 25 (ρ, λ) parameter combinations, obtained using MLE, are in the final four columns of Tables 6c and 6d, and are each based on 5000 replications unlike columns 3-6, reproduced from BF, that are based on 10,000 replications. Although the general pattern is clear, it should be noted that the standard errors of the estimated proportions reported in these tables are 0.002 for $s = 10000$ and 0.003 for $s = 5000$, and the standard error of the difference is approximately 0.0037 so that only differences of the order of 0.01 should be regarded as individually significant. That use of the MLE in the first steps of the procedure is delivering a closer match between nominal and empirical significance levels is immediately obvious. However, it remains the case that there is some size-inflation present, and the reduction in such inflation comes at the cost of greater conservatism for some parameter combinations - compare columns 6 and 10.

Table 6A

Case 2 $N = 26$. asymptotic test

		GTSLS based				MLE based			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0					.05	.05	.08	.04
$1d.f.$	1	.06	.06	.11	.04	.05	.05	.09	.03
$1d.f.$	2	.08	.08	.13	.05	.06	.05	.08	.04
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.06	.05	.12	.04	.05	.05	.08	.03
$2d.f.$	2	.07	.07	.14	.04	.05	.05	.09	.02
Power	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0					.12	.12	.23	.05
$1d.f.$	1	.21	.22	.34	.06	.16	.15	.28	.05
$1d.f.$	2	.32	.33	.50	.08	.21	.18	.36	.07
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.15	.15	.23	.05	.12	.11	.19	.05
$2d.f.$	2	.26	.26	.42	.07	.17	.14	.30	.06

Table 6B

Case 2 $N = 26$ Simple MLE-based bootstrap test

Size	r	Mean	Med	Max	Min
$1d.f.$	0	.05	.05	.06	.04
$1d.f.$	1	.05	.05	.07	.04
$1d.f.$	2	.05	.05	.06	.04
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.05	.05	.07	.04
$2d.f.$	2	.05	.05	.06	.04
Power	r	Mean	Med	Max	Min
$1d.f.$	0	.12	.11	.21	.05
$1d.f.$	1	.15	.14	.26	.04
$1d.f.$	2	.19	.16	.35	.05
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.12	.10	.20	.05
$2d.f.$	2	.16	.13	.29	.05

Table 6BB

Case 2 $N = 25$. Queen's case 5×5 grid vs ring with two neighbours

		Asymptotic MLE				Bootstrap MLE			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.06	.06	.07	.04	.05	.05	.06	.04
Power	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0	.20	.13	.53	.06	.18	.13	.51	.05

Table 6C

Case 2 $N = 200$. asymptotic test

		GTSLS based				MLE based			
Size	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0					.05	.05	.06	.04
$1d.f.$	1	.08	.05	.23	.04	.05	.05	.09	.04
$1d.f.$	2	.09	.06	.28	.03	.05	.05	.10	.03
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.10	.05	.34	.04	.05	.05	.13	.03
$2d.f.$	2	.11	.06	.38	.04	.05	.05	.14	.03
Power	r	Mean	Med	Max	Min	Mean	Med	Max	Min
$1d.f.$	0					.44	.32	.94	.05
$1d.f.$	1	.45	.44	.97	.05	.48	.43	.98	.06
$1d.f.$	2	.53	.57	.99	.06	.52	.47	.99	.06
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.47	.50	.96	.05	.50	.50	.95	.05
$2d.f.$	2	.56	.67	.97	.06	.57	.68	.98	.05

Table 6D

Case 2 $N = 200$ Simple MLE-based bootstrap test

Size	r	Mean	Med	Max	Min
$1d.f.$	0	.05	.05	.06	.04
$1d.f.$	1	.05	.05	.08	.04
$1d.f.$	2	.05	.05	.07	.04
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.05	.05	.10	.04
$2d.f.$	2	.05	.05	.10	.04
Power	r	Mean	Med	Max	Min
$1d.f.$	0	.43	.34	.93	.05
$1d.f.$	1	.46	.41	.97	.05
$1d.f.$	2	.50	.45	.98	.05
$2d.f.$	0	n.a.	n.a.	n.a.	n.a.
$2d.f.$	1	.48	.48	.95	.05
$2d.f.$	2	.54	.61	.97	.05

As in Case 1, notice that the most liberal parameter combinations are again *more liberal* for the sample size 200 than for the sample size 26. Since the proportions in columns 7-10 of these two tables are each based on 5000 replications, the standard error of the difference is approximately 0.0044, so again, as a rule of thumb individual differences of the order of 0.01 are significant. Again, evidently, the MLE-based statistics are much better sized than those computed using the GMM/GTSLS estimator; also, there is no power loss as a result of the tighter size control.

The upper and lower 10 empirical significance levels for $n = 26$ are given in Tables 7 and 8 in which $(\lambda_0, \rho_0) = (\lambda_1, \rho_1) \in (0.0, 0.3, 0.6, 0.9, 0.95)^2$. Similarly, the upper and lower 10 empirical significance levels for $n = 200$ are given in Tables 9 and 10. While the size distortions of the asymptotic tests are much smaller than for Case 1, for each sample size the pattern is similar, with the tests being liberal when λ is small and ρ large, and more-or-less correctly sized when λ is large, or when both λ and ρ are small. For the smaller sample the empirical sizes of both forms of the 1 and 2 *d.f.* tests are similar, while both forms of the 2 *d.f.* tests are slightly more liberal for the larger sample.

Table 7: Extreme empirical sizes, $n = 26$ lattice, 1 degree-of-freedom test, $r = 1, 2$ Case 2

Upper extreme sizes[a]							Lower extreme sizes[a]						
GTSLS		MLE		parameters			GTSLS		MLE		parameters		
χ^2_1	BS	χ^2_1	BS	r	λ_0	ρ_0	χ^2_1	BS	χ^2_1	BS	r	λ_0	ρ_0
.13	.09	.07	.06	2	0.0	.95	.05	.05	.05	.05	1	.90	.60
.11	.08	.08	.06	2	0.0	.90	.05	.05	.05	.05	1	.90	0.0
.11	.09	.07	.06	2	.30	.95	.05	.05	.05	.04	1	.90	.30
.10	.08	.06	.05	2	.30	.90	.05	.05	.04	.04	1	.95	.60
.10	.09	.09	.07	1	0.0	.95	.04	.06	.04	.05	1	.95	.90
.10	.08	.08	.06	1	0.0	.90	.04	.05	.04	.04	1	.95	.30
.09	.08	.07	.06	1	.30	.95	.04	.05	.04	.05	1	.90	.95
.09	.08	.07	.06	1	.30	.90	.04	.05	.04	.04	1	.95	0.0
.09	.07	.05	.05	2	.60	.90	.04	.05	.04	.05	1	.90	.90
.09	.06	.07	.05	2	.30	.60	.04	.05	.03	.04	1	.95	.95

a: in descending order of GTSLS size using the χ^2 critical values

Table 8: Extreme empirical sizes, $n = 26$ lattice, 2 degree-of-freedom test, $r = 1, 2$ Case 2

Upper extreme sizes[a]							Lower extreme sizes[a]						
GTSLS		MLE		parameters			GTSLS		MLE		parameters		
χ^2_2	BS	χ^2_2	BS	r	λ_0	ρ_0	χ^2_2	BS	χ^2_2	BS	r	λ_0	ρ_0
.14	.10	.09	.06	2	0.0	.95	.04	.04	.05	.05	2	.90	0.0
.12	.10	.08	.06	2	0.0	.90	.04	.06	.03	.04	1	.95	.90
.11	.11	.08	.07	1	0.0	.95	.04	.06	.04	.05	1	.90	.90
.11	.10	.08	.06	1	0.0	.90	.04	.06	.04	.05	1	.90	.95
.11	.09	.07	.06	2	0.3	.95	.04	.05	.04	.04	2	.95	0.0
.11	.09	.07	.06	2	0.3	.90	.04	.05	.04	.05	1	.90	.30
.10	.10	.07	.06	1	0.3	.95	.04	.05	.04	.04	1	.95	0.0
.09	.09	.06	.06	1	0.3	.90	.04	.05	.03	.04	1	.95	.95
.08	.06	.06	.05	2	0.0	.60	.04	.05	.03	.04	1	.95	.60
.08	.06	.06	.05	2	0.3	.60	.04	.04	.03	.04	1	.95	.30

a: in descending order of GTSLS size using the χ^2 critical values

Table 9: Extreme empirical sizes, $n = 200$ lattice, 1 degree-of-freedom test, Case 2

Upper extreme sizes[a]							Lower extreme sizes[a]						
GTSLS		MLE		parameters			GTSLS		MLE		parameters		
χ^2_1	BS[b]	χ^2_1	BS[b]	r	λ_0	ρ_0	χ^2_1	BS[b]	χ^2_1	BS[b]	r	λ_0	ρ_0
.28	.23	.10	.07	2	0.0	.95	.05	.05	.06	.05	1	0.6	0.0
.23	.20	.09	.08	1	0.0	.95	.05	.04	.05	.05	2	0.0	0.6
.22	.18	.08	.07	2	0.3	.95	.05	.05	.05	.05	1	0.9	0.3
.20	.17	.08	.06	1	0.3	.95	.05	.05	.04	.04	1	0.9	0.6
.18	.15	.06	.05	2	0.3	.90	.04	.05	.04	.05	2	0.9	0.9
.17	.14	.07	.05	2	0.0	.90	.04	.05	.05	.05	1	0.9	0.0
.14	.13	.07	.06	1	0.3	.90	.04	.05	.04	.04	1	0.9	0.9
.14	.12	.06	.05	1	0.0	.90	.04	.05	.04	.05	1	0.9	.95
.12	.11	.06	.06	1	0.6	.95	.04	.04	.04	.05	2	0.9	0.6
.11	.10	.05	.05	2	0.6	.90	.03	.05	.03	.04	2	0.9	.95

a,b: See Table 2

Table 10: Extreme empirical sizes, $n = 200$ lattice, 2 degree-of-freedom test, Case 2

Upper extreme sizes[a]							Lower extreme sizes[a]						
GTSLS		MLE		parameters			GTSLS		MLE		parameters		
χ^2_2	BS[b]	χ^2_2	BS[b]	r	λ_0	ρ_0	χ^2_2	BS[b]	χ^2_2	BS[b]	r	λ_0	ρ_0
.38	.31	.14	.10	2	0.0	.95	.05	.05	.04	.05	1	0.6	0.6
.34	.29	.13	.10	1	0.0	.95	.05	.05	.05	.05	1	0.3	0.0
.29	.24	.10	.09	2	0.3	.95	.05	.05	.03	.04	1	0.9	0.6
.28	.25	.10	.08	1	0.3	.95	.05	.05	.04	.04	1	0.3	0.6
.26	.20	.07	.05	2	0.0	.90	.05	.05	.04	.05	1	0.9	0.3
.23	.18	.07	.05	2	0.3	.90	.04	.06	.03	.05	2	0.9	0.9
.22	.19	.07	.06	1	0.0	.90	.04	.04	.04	.05	2	0.9	0.6
.20	.18	.06	.06	1	0.3	.90	.04	.05	.03	.05	1	0.9	0.9
.14	.14	.05	.06	2	0.6	.95	.04	.05	.03	.05	2	0.9	.95
.14	.14	.07	.07	1	0.6	.95	.04	.05	.03	.05	1	0.9	.95

a,b: See Table 2

That the simple bootstrap is superior to the asymptotic tests in this setting is not obvious from the rather marginal improvements visible in Tables 7 - 10. However, as in Case 1, we see that the size distortions are worse for the $N = 200$ weight matrices than for the $N = 26$ matrices. To see if increasing the sample size further would bring the extreme empirical significance levels further down towards the nominal 5%, the extreme cases were run with the weight matrix comprising two diagonal blocks, each of which was equal to the NUTS 200 matrix. The results, for the 1 d.f. test, show that the increased sample size eliminates the problem evident in the left side of the top four rows of Table 9:

Table 9A
Case 2, $N = 400$ (2 copies of the NUTS 200 weights as diagonal blocks)
1 d.f. test, MLE-based

Parameters			Asy Test		BS Test	
r	ρ	λ	size	pow	size	pow
0	0.95	0.0	.057	.31	.050	.30
1	0.95	0.0	.048	.31	.046	.30
2	0.95	0.0	.052	.29	.051	.28
0	0.95	0.3	.052	.37	.051	.35
1	0.95	0.3	.044	.38	.044	.38
2	0.95	0.3	.047	.38	.047	.36

The effect of instrument choice is not as simple as in Case 1. First of all, the minimal set, obtained by setting $r = 0$, is not rich enough to permit calculation of the two-degree-of-freedom test in this case: when $\mathbf{X}_{02}$ and $\mathbf{X}_{12}$ coincide we have $\mathbf{L}_{0,r} = [\mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{W}_0\mathbf{X}_{02} .. \vdots \mathbf{W}_0^r\mathbf{X}_{02}]$, $\mathbf{H}_{0,r} = [\mathbf{L}_{0,r} \vdots \mathbf{M}_0\mathbf{L}_{0,r}]_{LI}$ while $\mathbf{L}_{1,r} = [\mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{W}_1\mathbf{X}_{02} .. \vdots \mathbf{W}_1^r\mathbf{X}_{02}]$, and $\mathbf{H}_{1,r} = [\mathbf{L}_{1,r} \vdots \mathbf{M}_1\mathbf{L}_{1,r}]_{LI}$; further, setting $\mathbf{X}_{01} = \mathbf{1}$, $\mathbf{M}_0 = \mathbf{W}_0$ and $\mathbf{M}_1 = \mathbf{W}_1$, noting that our row normalisation gives $\mathbf{W}_j\mathbf{1} = \mathbf{1}$, $j = 0, 1$ the above instrument matrices reduce to $\mathbf{H}_{0,0} = [\mathbf{1} \vdots \mathbf{X}_{02} \vdots \mathbf{W}_0\mathbf{X}_{02}]$ and $\mathbf{H}_{1,0} = [\mathbf{1} \vdots \mathbf{X}_{02} \vdots \mathbf{W}_1\mathbf{X}_{02}]$ so that $\mathbf{H}_0^{**} = [\mathbf{H}_{0,0} \vdots \mathbf{H}_{1,0}]_{LI} = [\mathbf{1} \vdots \mathbf{X}_{02} \vdots \mathbf{W}_0\mathbf{X}_{02} \vdots \mathbf{W}_1\mathbf{X}_{02}]$. The dimension of $\boldsymbol{\gamma}^{**}$ is 4, but that of $\boldsymbol{\gamma}^{\dagger}$ is 5, so that the 2 degree of freedom test cannot be calculated with $r = 0$. However, the one-degree-of-freedom test *can* be calculated with this minimal instrument set, and doing so gives the results in Tables 11 and 12.

Table 11: Extreme empirical sizes, $n = 26$, 1 degree-of-freedom test, $r = 0$
Case 2

Upper 5 extreme sizes[a]						Lower 5 extreme sizes[a]					
GTSLS		MLE		Parameters		GTSLS		MLE		Parameters	
χ_1^2	BS	χ_1^2	BS	λ_0	ρ_0	χ_1^2	BS	χ_1^2	BS	λ_0	ρ_0
.06	.06	.05	.05	0.0	0.0	.02	.06	.04	.04	.95	.30
.06	.05	.06	.06	0.3	0.0	.02	.07	.04	.04	.90	.90
.05	.06	.05	.05	0.6	0.0	.02	.07	.04	.05	.90	.95
.05	.05	.06	.05	0.3	0.3	.02	.06	.04	.04	.95	.90
.05	.05	.06	.05	0.0	0.3	.02	.07	.04	.05	.95	.95

a: in descending order of GTSLS size

Table 12: Extreme empirical sizes, $n = 200$, 1 degree-of-freedom test, $r = 0$
Case 2

Upper 5 extreme sizes[a]						Lower 5 extreme sizes[a]					
GTSLS		MLE		Parameters		GTSLS		MLE		Parameters	
χ_1^2	BS	χ_1^2	BS	λ_0	ρ_0	χ_1^2	BS	χ_1^2	BS	λ_0	ρ_0
.07	.06	.05	.05	0.0	.95	.04	.05	.04	.05	.95	.60
.05	.06	.05	.05	.90	0.0	.03	.05	.04	.05	.90	.90
.05	.05	.05	.05	0.0	0.0	.03	.05	.04	.05	.95	.90
.05	.05	.06	.06	.60	0.0	.03	.05	.04	.05	.90	.95
.05	.05	.05	.05	0.0	.30	.02	.05	.04	.05	.95	.95

a: in descending order of GTSLS size

To assess the impact, if any, of reversing the roles of the two weight matrices, the simulations for $r = 0$ and the 1 degree-of-freedom test for the $N = 200$ sample size were repeated with the modified NUTS weights as the null. Only random differences in empirical significance levels attributable to sampling variation, and no significant variation in power were observed. This result cannot be taken as definitive, obviously, but there is no reason to suggest that it will be untypical.

Evidently, for Case 2 with either $N = 26$ or $N = 200$, or $N = 400$, taking $r = 0$ and using the 1-degree-of-freedom test we achieve approximately correct significance levels with all four methods. The MLE-based asymptotic test is slightly less conservative at the lower end than the GTSLS-based test, but both these tests are essentially fully corrected by the bootstrap.

5 Conclusions

In broad terms:

(i) the asymptotic tests based on the MLE have virtually correct significance levels provided the dimension of the instrument set is as small as possible, and are superior to the tests using the generalised methods of moment estimators

Concentrating therefore on the MLE-based tests:

(ii) in Case 1 there is little to choose between the one and two-degree-of-freedom statistics though the 1 df test has slightly more power

(iii) in Case 2 the fact that the 1 df test can be calculated with a smaller instrument set gives it a clear practical advantage

(iv) the bootstrap appears to deliver correct significance levels when the minimal instrument set is used in virtually all cases, and in all cases tested with $N = 400$, except one (to follow......).

(v) if the bootstrap is not used, then reasonable size control can be achieved by using the minimal instrument set.

At this moment it may be worth acknowledging the inherent limitations of this kind of numerical exercise. Ideally, one wants an exact algebraic account of the dependence of test performance on parameter values, regressor characteristics, weight matrices and sample size, but faced with the difficulty of providing such an analysis, a numerical approach is the next best thing. It must not be forgotten, however, that numerical evidence cannot prove the absence of problems, but it can reveal them. Some comfort can be taken if major problems are not revealed by an attempt to find them, of course. Our experiments relate to the case of a single alternative model, so $g = 1$, and a single non-constant explanatory variable, and either different weight matrices or different regressors, but not both. The weights for the main experiments were taken from real examples that have been used in empirical research. In small samples the tests are clearly subject to implementation problems in some parts of the parameter space that are not entirely eliminated by use of a simple bootstrap resampling procedure, and these cases merit further study.

References

[1] Burridge, P., Fingleton, B. (2010) Bootstrap Inference in Spatial Econometrics: the J-test. Spatial Economic Analysis, 5, 1, 93-119.

[2] Cliff, A.D., Ord, J.K. 1973. Spatial Autocorrelation. Pion. London.

[3] Fingleton, B. 2007. Testing the “new economic geography” : a comparative analysis based on EU regional data, Chapter 3 pp 70-97 in New Directions in Economic Geography (Ed B Fingleton)

[4] Davidson, R., MacKinnon, J.G. 1981. Several tests for model specification in the presence of alternative hypotheses. Econometrica 49, 781-793.

[5] Kelejian, H.H. 2008. A Spatial J-test for model specification against a single or a set of non-nested alternatives. Letters in Spatial and Resource Sciences 1, 3-11.

[6] Kelejian, H.H., Prucha, I.R. 1998. A generalised 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

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

[8] LeSage, J., and R.K. Pace 2009. Introduction to Spatial Econometrics. CRC Press.